2010

MARKET STRUCTURE AND MORTGAGE PRICING: THE ROLE OF INFORMATION IN FIRM AND CONSUMER BEHAVIOR

Abdullah A. Al-Bahrani
University of Kentucky, a.albahrani@uky.edu

ABSTRACT OF DISSERTATION

Abdullah A. Al-Bahrani

The Graduate School

University of Kentucky

2010

MARKET STRUCTURE AND MORTGAGE PRICING: THE ROLE OF INFORMATION IN FIRM AND CONSUMER BEHAVIOR

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Business and Economics
at the University of Kentucky

By
Abdullah A. Al-Bahrani

Lexington, Kentucky

Director: Dr. Frank Scott, Professor of Economics

Lexington, Kentucky

2010

ABSTRACT OF DISSERTATION

MARKET STRUCTURE AND MORTGAGE PRICING: THE ROLE OF INFORMATION ON FIRM AND CONSUMER BEHAVIOR

This dissertation analyzes information, market structure, and firm pricing strategies. I begin the dissertation with an analysis of the market structure of the mortgage industry. I find that the configuration of the mortgage market at its present state is vastly different than its historical structure. The reduction in the cost of transmitting information has increased the collaborative environment and facilitated the dis-integration of the supply chain. Generally, the mortgage industry has been successful at reducing principal-agent problems and minimizing asymmetric information concerns that arise in segmented markets.

In the first essay I provide a theoretical explanation of the effect of the internet on market outcomes. Search models assume that the reduction in search frictions would lead to competitive markets. However, I argue that gatekeepers operating in online markets may create an anticompetitive effect, in addition to reducing the consumers' search cost. Therefore, the conduct of the gatekeeper can cause prices in online markets to be higher than in retail markets and provide online firms with larger profits.

In the second essay "I empirically examine the role of the internet and Internet Comparison Search sites in reducing consumer search costs and their effects on the prices consumers pay for mortgages. Additionally, I expand the study to test for the effects of the internet on firm profits. Using a unique data set, I examine a mortgage firm's pricing strategies and profits in online and retail markets, and find evidence of market power in online markets that do not exist in retail markets. The presumed benefits to the consumer from the reduction of search cost are offset by the anticompetitive environment in online markets.

In the final essay, I examine a mortgage firm's portfolio choice. I investigate the loan characteristics that affect the firm's decision to retain mortgages as part of its own portfolio. I find that the decision to retain loans as a lender is driven by unobservable qualities. The firm does sort loans by quality, but it also prices non-brokered loans lower based on unobservable qualities. The sorting behavior suggests that asymmetric information exists between the lender and the secondary market.

KEYWORDS: Asymmetric information, mortgage, internet, market structure, pricing strategies.

Abdullah Al-Bahrani

Date

MARKET STRUCTURE AND MORTGAGE PRICING: THE ROLE OF INFORMATION ON FIRM AND CONSUMER BEHAVIOR

By

Abdullah A. Al-Bahrani

Director of Dissertation

Director of Graduate Studies

Date

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Abdullah A. Al-Bahrani

The Graduate School

University of Kentucky

2010

MARKET STRUCTURE AND MORTGAGE PRICING: THE ROLE OF INFORMATION ON FIRM AND CONSUMER BEHAVIOR

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Business and Economics
at the University of Kentucky

By
Abdullah A. Al-Bahrani

Lexington, Kentucky

Director: Dr. Frank Scott, Professor of Economics

Lexington, Kentucky

2010

Acknowledgements

I am thankful to my advisor, Frank Scott, whose encouragement, guidance and support from the initial to the final stage enabled me to complete this dissertation. I also benefited greatly from my dissertation committee, Adib Bagh, Chris Bollinger, and Don Mullineaux. Finally, I thank my family; this dissertation would not be possible without their love and support.

Table of Contents

List of Tables

List of Figures

1. CHAPTER 1: Introduction to Dissertation

Homeownership is often cited as an essential component of living the American dream. This may be attributed to several reasons, including consumer preferences, asset portfolio diversification,[1] or as a response to the favorable tax treatment of owner-occupied housing.[2] Homeownership also comprises a large portion of consumption expenditure and remains the single most important asset in homeowners' portfolios. In the decade from 1996 to 2005, the homeownership rate in the United States increased from 65.4 percent to 68.9 percent . Although the benefits of homeownership are mostly pecuniary in nature, researchers also have identified non-pecuniary benefits; for example, homeowners are more likely to develop management and problem-solving skills (Green and White 1997). Homeowners are also more likely to become members of nonprofit organizations, solve local problems and confer other external benefits on society. The gains associated with homeownership are not limited to the individual but extend to the local economy, as DiPasquale and Glaeser (1999) propose. Thus, homeownership is a significant asset to the individual, his/her neighbors, and the local and national economy as a whole.

According to the US Census, the median income in 2009 was $50,303, while the average price of a house in the America was $282,000. The large disparity between income and home the cost of a home places an enormous financial burden on the consumer and thus requires vibrant credit markets to help facilitate homeownership (Campbell and

[1] For a review of literature on Portfolio choice see Carcuru,S., J. Heaton, et al. (2004). "Heterogeneity and Portfolio Choice: Theory and Evidence."

[2] Poterba, J. and T. Sinai (2008). "Income Tax Provision Affecting Owner-occupied Housing: Revenue Costs and Incentive Effects." NBER Working Paper 14253

Cocco 2003). Despite the recent collapse of financial institutions and the mortgage market, current and prospective homeowners have access to a large number of mortgage suppliers and an extensive array of mortgage programs. The mortgage market is an important factor in attaining the "American Dream," and understanding the market and its components can aid policy makers, lenders, and consumers in making homeownership a reality.

In this dissertation, I provide a detailed examination of the mortgage industry and the role of informational frictions in determining the structure and performance of the mortgage market. In Chapter 2, I provide an analysis of the market structure of the mortgage industry. This work also discusses the institutional details and the complex structure of the mortgage market. I find that the configuration of the mortgage market, in its present state, is vastly different than its historical structure due to the reduction of information costs. The innovation of the internet and local area networks have reduced the cost of transmitting information and increased the ability of firms to collaborate. As a result, the mortgage market has disintegrated into specialized firms, where each firm provides a single component and collaborates with other components of the supply chain to produce a final product. However, the vertical disintegration of the supply chain has introduced principal-agent problems. Asymmetric information concerns are present mainly in the broker segment of the mortgage market. The incentives and objectives of the broker are incompatible with the overall objective of the secondary market. Consequently, I find that more complete contracts are needed to realign the broker and market objectives.

Chapter 3, "A Model of Electronic Commerce in the Presence of a Gatekeeper," provides a theoretical explanation of the effect of the internet on market outcomes. Em-

pirical studies examining the effect of the internet have provided inconsistent results. Search models assume that the reduction in search frictions would lead to more competitive markets. However, I argue that gatekeepers operating in online markets may create an anticompetitive effect in addition to reducing the consumers' search cost. Therefore, the conduct of the gatekeeper can cause prices in online markets to be higher than in retail markets and provide online firms with larger profits. When testing for the effect of the internet on market outcomes, it is imperative to consider the conduct of the gatekeeper in the market and test for both prices and profits.

In Chapter 4, "Joint Determination of Consumer Search Behavior and Mortgage Pricing," I examine the role of the Internet and Internet Comparison Search sites in reducing consumer search costs, and the effects these searches have on the prices consumers pay for mortgages. Additionally, given the results from the third chapter, I expand the study to test for firm profits. Using a unique data set, I examine a mortgage firm's pricing strategies in online and retail markets, and I uncover evidence of market power in online markets that does not exist in retail markets. To control for consumers' selection into market type, I use a switching regression and find that selection into online commerce is random and thus the estimation method is reduced to a pooled OLS. I find that online and retail consumers pay the same price on average for a mortgage. After controlling for variation in marginal cost, I find that the firm earns higher profits in online markets relative to the retail market. In addition to reducing the cost of search to the consumer, the availability of information in online markets allows firms to observe their competitors' pricing, and therefore reduces the firms' cost of monitoring. Both consumers and firms benefit from the availability of information, and consequently the presumed benefits to the con-

sumer from the reduction of search costs are offset by the anticompetitive environment created by Internet Comparison Search sites in online markets. For some mortgage types, the anticompetitive effect dominates the search cost effect, which leads to higher prices in online markets.

In the final chapter, I examine a mortgage firm's portfolio choice. Using the unique data set from the previous chapter, I investigate the loan characteristics that affect the firm's decision to retain mortgages as part of its portfolio. Mortgage lenders can use their own funds to finance a mortgage or may broker the loan using other lenders' financing. I estimate the quality differences between lender-originated loans and brokered loans using Oaxaca decomposition. I am able to decompose differences in prices between the broker and lender market into portions explained by the observed data and unexplained pricing, which I attribute to asymmetric information. Since the interest rate of a mortgage is the price of its underlying risk, any variation in price is due to risk characteristics.

I find that non-brokered loans are, on average, priced three basis points lower than brokered loans. However, the difference is statistically insignificant. By decomposing the difference, it is possible to disaggregate the disparity in pricing into differences due to observable risk characteristics and differences due to unobserved characteristics. Non-brokered loans are priced on average 18 points lower than brokered loans because of unobserved qualities. Consequently, it is possible to assume that the lower unexplained price for non-brokered loans is in response to better unobserved characteristics of lender originated loans. Although this supports the presence of asymmetric information in the mortgage market, the unobserved difference cannot be attributed to asymmetric informa-

tion with certainty. It is possible that the difference arises because of the lower cost of funds when the firm operates as a lender. Therefore, the lower prices in the lending market may be due to the firm passing on the reduction in cost to the consumer.

2. CHAPTER 2: The Mortgage Process and the Vertical Disintegration of the Supply Chain

2.1 Introduction to the Mortgage Market

The mortgage market, as it presently exists, encompasses numerous institutions and individuals who contribute to the provision of the financial product. This chapter serves as a detailed look into the structure of the mortgage market, while providing the institutional details and explaining the services provided by the primary and secondary mortgage markets. My approach is to follow a mortgage through the stages of its "life," beginning with the consumer's decision to obtain a mortgage, the origination, underwriting, the financing of the mortgage, and ending with the securitization process.

2.1.1 *The Mortgage Decision*

The first step in this chain of activities begins with a consumer's decision to acquire a mortgage. A consumer, or borrower, in search for a mortgage is either in the process of purchasing a new house or refinancing an existing mortgage. Purchase consumers demand funds to finance the acquisition of a new home. Alternatively, the motive to refinance an existing loan is more complex and may be due to the consumer's desire to obtain a lower interest rate and/or adjust the term of the loan (rate or term refinance), to consolidate non-tax-deductible interest payments into a tax deductible mortgage, or to receive cash for any other purpose (cash out refinance).

Between 2001 and 2006, the demand for mortgages increased drastically as market interest rates fell below their historical averages. Figure 2.1 illustrates that refinance

applications, measured by the Mortgage Bankers Association (MBA) refinance index, increased after market interest rates reached their lowest point in March 2003. As market interest rates reached historic lows, homeowners began to refinance their existing mortgages and exchange them for lower rate mortgages. Given the decline in the price of mortgage financing, homeownership became more affordable and consequently assisted in the increase of homeownership rates during the same period.

2.1.2 *Consumer Search Process*

Upon deciding to obtain a mortgage, the consumer must locate a firm to finance the mortgage. The vast array of possible products and seemingly differentiated programs make the search process complicated. Consumers searching for a mortgage find that acquiring product information is exceedingly difficult (LaCour-Little 2000). Before the internet and electronic commerce, consumers searching for a mortgage product would be required to contact and physically visit the mortgage supplier. However, the internet has made the search process less onerous, and it has made information about mortgage products and programs more accessible to the consumer. Beginning with the consumers' search process, Figure 2.2 depicts the steps of a mortgage as it progresses through the supply chain.

The consumers' search begins with the decision to either conduct retail or an electronic search process. LaCour-Little (2000) surveyed the role of technology in mortgage financing between 1990 and 2000 and predicted that the internet would reduce the cost of search to the consumer and the cost of operating a mortgage firm. At this point in time, it was expected that the reduction of frictions associated with the acquisition and transmis-

sion of information would increase the consumer surplus and reduce the price consumers pay for a mortgage.

Although the costs of search have been reduced, the empirical evidence provided in Chapter Four indicates that the lower search cost in online markets is not associated with lower prices, as theory suggests. Online search through firms like Lendingtree.com does not necessarily increase the competitiveness of the market and may have the opposite effect of increasing the suppliers' market power. Additionally, the availability of information about mortgage products and programs does not fully eliminate the cost of acquiring a mortgage because the consumer is not aware of the underlying guidelines associated with the products or which firm supplies the product. To become more informed of each firm's guidelines and mortgage underwriting process, the consumer must submit an application to the mortgage supplier. This implies that the gains from the reduction of search costs due are not fully realized because mortgage suppliers have resisted making the information totally available to the consumer.

2.2 Primary Market

The mortgage market has two separate but connected parts, simply referred to as the primary market and the secondary market. The initial phase of the loan, where consumers and mortgage suppliers transact, discuss the consumer's loan preferences, and eventually originate the mortgage, is recognized as the primary market. After new loans are originated, the primary market supplies the secondary market with the loans, which are in turn traded by investors.

2.2.1 *Mortgage Choices*

During the origination process, the consumer and originator agree on the mortgage product and program that best meets the consumer's needs. Mortgage products can be complex, vary in many dimensions, and are supplied by different types of mortgage originators. In general, mortgage products are usually a variation of one of two types: fixed or adjustable rate mortgages (ARM). The choice of mortgage product depends on whether the consumer meets the required guidelines, the consumers' preferences, and expectation of future interest rate movements. The fixed rate program has a preset interest rate and offers the consumer a constant monthly payment for the life of the loan. Adjustable rate mortgages are fixed during the introductory period, usually two to five years, and adjust thereafter. With a fixed rate product, the consumer can be certain that future payments would remain unchanged even if the market environment does change. During the introductory phase, the interest rate on ARM products is lower than the interest rate on comparable fixed rate mortgages. When deciding between fixed or adjustable rate mortgages, the consumer is trading off between the certainty of future payments and the lower initial price of ARM products during the introductory period.

The lender charges a lower interest rate for ARM products to compensate the consumer for bearing the interest rate risk and absorbing some of the lender's risk exposure. When the mortgage interest rate is fixed, the lender's profit margin is uncertain and would fluctuate with the movements in the market interest rates; however, profit margins for ARM products are constant because the interest on the mortgage adjusts as the market interest rate moves. Therefore, ARM products allow the lender to transfer the risk expo-

sure to the consumer. In contrast, when market interest rates fall, consumers with fixed rate mortgages would have a higher interest rate relative to the market. From the consumer perspective, ARM products become more attractive when future interest rates are expected to decrease as they would allow the consumer to benefit from the interest rate movement. Fixed rate mortgages are preferred when market interest rates are expected to increase; historically consumers' mortgage choices have been in response to the expectation of market interest rate movements. However, during the recent period of low interest rates, consumers shifted toward ARM products, while historic trends would suggest an increase in the demand for fixed rate mortgages. The recent shift towards ARM products is due to the product's low initial interest rate and consequently its affordability.[3]

The Case-Shiller home price index, a measure of national home prices, increased dramatically between 2001 and 2007, as indicated in Figure 2.3. The unprecedented appreciation of home prices resulted in an increase in consumer demand for mortgage refinancing and an increase in mortgage loan amounts. Due to the drastic appreciation rates, homeowners experienced an increase in their net worth and began to refinance their mortgages to extract the available equity from their homes and consequently realize the gains. Alternatively, consumers purchasing new homes were required to either provide a larger down payment or demand larger loan amounts to finance the purchase of more expensive homes. Yet, growth in income was deficient in comparison to home appreciation rates, which further exacerbated growth of loan amounts during 2001 to 2006. Two important measures of home affordability, which lenders examine during the loan under-

[3] Campbell et al. 2003 suggest that if consumers are risk-averse, fixed rate mortgage products are favorable for individuals with large mortgages, volatile income, high default cost, or immobile homeowners.

writing process, are the *housing payment-to-income* ratio and the *total debt-to-income* ratio. The *payment-to-income* ratio is a measure of the fraction of the gross income allocated to housing, and the *debt-to-income* ratio measures the amount of income allocated to all liabilities. Since loan amounts were increasing faster than income, both measures of affordability increased. To minimize the required monthly payment, homeowners shifted towards products with lower interest rates. ARM products offered consumers lower initial payments compared to fixed mortgages, but in doing so, the consumer absorbed the interest rate risk. Although historic trends suggested that consumers preferred to lock in low rates when market interest rates were low, the unprecedented home appreciation made ARM products more appealing.

Follain (1990) provides a comprehensive description of the mortgage choice decision and details several factors that consumers consider when comparing mortgage products. In addition to choosing the mortgage type, analogous to a firm's decision of the debt to equity level to maintain, a borrower must decide on the loan-to-value ratio. As the loan-to-value ratio increases, the amount of equity the consumer has invested in the property decreases. From a lender's perspective, the amount invested by the consumer serves as a proxy for consumer's attachment to the property. Consumers without an investment in the property are more likely to walk away or default on the loan.[4] As the loan-to-value ratio increases, the expected probability of default increases. The choices made by the consumers are important to the lender because the lender's central objective is to attempt to identify the likelihood a consumer will successfully repay the mortgage.

[4] The literature on strategic default suggests that consumers will only default on a mortgage when they have negative equity (when the value of a home is less than the outstanding mortgage). Other researchers have argued that equity and default are negatively correlated, even if equity is positive. See Order, R. V. (2000). "The U.S Mortgage Market: A Model of Dueling Charters." Journal of Housing Research **11**(2): 233-256.

2.2.2 *Mortgage Originators*

During the search for a mortgage product, consumers are also implicitly choosing the mortgage originator. Before the collapse of the Savings and Loans Industry in 1989, options for mortgage suppliers were limited and centralized. Borrowers' searches led them to their depository institution, such as a Savings and Loans or community bank. The depository institutions financed mortgages based on available deposits. In the customary process, the mortgage provider was responsible for originating the loan, underwriting the risk, financing the loan, and finally servicing the loan until its termination.[5] In the traditional structure, the mortgage originator retains the loan in its own asset portfolio and is responsible for it throughout the life of the loan.

Contrary to its initial structure, the current state of the mortgage market involves many specialized firms which have vertically dis-integrated and unbundled the services provided by the thrifts. Advancements in the secondary market have influenced the structural evolution of the mortgage market, but have done so in conjunction with or in response to regulatory changes in the banking industry. Before the 1990s, thrifts managed the entire mortgage process mainly due to the comparative advantage acquired through regulation. As a commitment towards increasing homeownership rates, government regulation reduced capital requirements and provided lower tax rates for firms investing in mortgages. Moreover, beneficial treatment of thrifts reduced their cost of borrowing from depositors (Regulation Q), and gave them a reduction in depository insurance premiums (Follain and Zorn 1990).

[5] Termination refers to default, prepayment, or maturity of the loan.

In 1989, the collapse of Savings and Loans ushered in the Financial Institutions Reform, Recovery and Enforcement Act (FIRREA). The Act eliminated the policies that favored thrifts and led to the centralization of the mortgage process. With equal treatment, new firms began to compete with traditional mortgage originators. Due to economies of scope, advances in technology, and the securitization market,[6] the unbundling of mortgage services soon followed. The vertical dis-integration of the mortgage process has created an assortment of firms that specialize in loan origination, underwriting, financing, or loan servicing.

Mortgage originators typically fall into one of two comprehensive categories; they are classified as lenders or third party originators (TPOs). The distinction between the two groups is based on the mortgage originators' services, intensity of specialization, and the degree of dependence on secondary market financing to supply mortgages.

Lenders are firms that originate loans, and either sell their loans to the secondary market or maintain them in their own portfolio. The lender classification is further segregated into portfolio lenders, mortgage bankers, and wholesale lenders. In addition to originating the loan, a portfolio lender funds its own loans, services the loans, and most importantly maintains the loan in its own portfolio. Their production model most resembles the thrifts' mortgage process. Usually they are depository institutions, and therefore the financing of new mortgages is achieved through consumer deposit. This allows the portfolio lender to provide the consumer with unique loan programs and often advertise special loan products that do not conform to the guidelines imposed by the secondary market. In the case that portfolio lenders need to liquidate their unique loans, they must ser-

[6] Securitization process is discussed in further detail in the next section.

vice the loan for a year before it can be eligible for sale into the secondary market. Banks and other depository institutions, like Washington Mutual, are considered portfolio lenders.

Firms that finance the loan and immediately sell it to the secondary market are referred to as mortgage bankers. Like portfolio lenders, mortgage bankers may also be depository institutions, making dissimilarity between mortgage bankers and portfolio lenders less apparent. The distinction between the two types of operations is that mortgage bankers usually unbundle mortgage servicing rights to other firms. Additionally, because mortgage bankers immediately sell their loans to the secondary market, they abide by the secondary market guidelines. In the case that a mortgage banker does not collect deposits, mortgages are financed through a warehouse line of credit. For firms operating by means of credit lines, it is imperative for them to remain liquid and thus sell loans instantaneously to the secondary market to facilitate the origination of new loans. Countrywide Financial and Wells Fargo Home Loans operate as mortgage bankers; however, both Countrywide Financial and Wells Fargo Home Loans also boast wholesale lending divisions.

Wholesale lenders are mortgage bankers or portfolio lenders that, in addition to having their own retail division, obtain some of their business through third party originators (TPOs), specifically mortgage brokers. The difference between the lender classifications is a function of who provides the origination services and the level of integration between the lender and the originator. When the origination and underwriting services are vertically integrated, the firm operates as a mortgage banker or a portfolio lender. On the

other hand, wholesale lenders disintegrate the origination and underwriting process and allow specialized independent brokers to originate the loan. This provides the lender with the ability to specialize in underwriting and mortgage financing services. This in turn reduces their operating costs by allowing them to divest from the origination process. Theoretically, a firm can operate solely as a wholesale lender, in which case it is a buyer of origination services and specializes in underwriting loans.

The wholesale lender maintains ownership over the underwriting services, financing the loan and sale of the loan to the secondary market. As part of the unbundling process in the mortgage industry, third party originators have specialized in the mortgage origination services. The independent broker assists the consumer in finding the product that meets their preferences by accessing the products and guidelines of a multitude of wholesale lenders. The existence and purpose of the independent broker is to reduce the transaction cost incurred by the consumer due to searching for a product that meets their needs. In recent years, mortgage brokers have become a popular method of consumer search. They have seen their market share increase from 52% in 1997 (Lamalfa 1998) to 68% of the origination market in 2004.[7]

The increase in mortgage broker utilization would suggest that there are benefits to the consumer for choosing a mortgage broker. According to LaMalfa (1998), the popularity of mortgage brokers is due to their lower cost structure and incentive-based pay. The performance pay attracts the best and most ambitious loan officers. Although loan officers in the mortgage brokers system are agents to the consumer, they do not have a

[7] http://www.mortgagebankers.org/files/Bulletin/InternalResource/44664_September2006-ResidentialMortgageOriginationChannels.pdf

fiduciary responsibility to the borrower. This is further highlighted by the conflict of interest in the broker's sources of revenue. Brokers can obtain income through origination fees (commission) collected from the consumer for the service of matching them with a lender; they can also obtain income through yield spread premiums (YSP) paid by wholesale lender to the broker. The yield spread premium is created by the lender as an incentive to the broker to charge an interest rate higher than required by the minimum guidelines.

Wholesale lenders can further unbundle their services by relinquishing both origination and underwriting services to the TPO. The firm providing origination and underwriting services operates as a *correspondent* for the wholesale lender. In this case, the correspondent[8] is responsible for originating the loan and underwriting it so that it meets the wholesale lender's guidelines. The wholesale lender specializes in financing the mortgage and buys the asset immediately from the correspondent at the closing of the loan. Thereafter, the wholesale lender can either retain the loan in its own portfolio (portfolio lender) or sell it to the secondary market (mortgage banker).

The categorical definitions for loan originators are not exclusive and firms can operate under different types. Firms operating in the market can easily transition within the lender types, continuously altering which mortgage services to make or buy in response to market conditions. Traditionally, firms have remained separated within the larger classification subsets, either operating as a lender or third party originator. However, recently a new firm model has developed which allows a firm to operate across the

[8] Also referred to as table funders.

lender and broker boundary. The new class of firm operates as a hybrid; it can choose to operate either as a broker (TPO) or as a mortgage banker.

The development of the hybrid lender is a result of the increased liquidity in the secondary market. Lenders previously incurred a high cost of establishing the source of funds to finance mortgages. The increase in demand for mortgage securities in the secondary market increased the supply of mortgage loans, which led to a reduction in lending standards (Dell'Ariccia, Igan et al. 2009). The decline in the cost of lending allowed existing TPOs to develop mortgage banking divisions and consequently merge portions of the process that were previously unbundled. Unlike lenders, the hybrid firm can be a lender while also retaining the option to act as a broker or a correspondent. The possibility of operating as either a lender or a broker raises questions about the conditions under which the firm chooses one form of operation over the other. A discussion of these conditions is the topic of Chapter Five.

The extensive assortment of lenders and originators is the byproduct of the specialization that occurred due to the increased ability of the market to coordinate the flow of information between the firms. The advent of the internet and local area networks reduced the frictions and transaction cost in the market (LaCour-Little 2000). The mortgage broker model of operations is only possible because loan officers are able to access and survey wholesale lenders pricing effortlessly and in real time. The ability of a broker to coordinate with the lender facilitates the vertical dis-integration of the process. Lenders are able to specialize in the financing of mortgages while allowing other firms to specialize in origination services. Although the reduction of market frictions was initiated by

changes in the regulatory environment, it was enhanced by the growth of the secondary market and the reduction in the cost of information. The ability to dis-integrate the vertical chain of supply within the primary market would not be possible if firms could not transact and share information readily (Jacobides 2003).

2.3 Secondary Market

Although the market for mortgage suppliers is large and segmented, the main objective of the mortgage process is to connect the consumers with the end investor. The lender is simply an intermediary in the process that funds mortgages using financing from the credit markets. Therefore, they merely serve as a link between demanders of funds and the suppliers of capital in the secondary market. With the exception of the portfolio lender, which retains the mortgage on its own balance sheet, all other lenders utilize secondary market financing to fund their operations.

The secondary market is structured so that it increases liquidity in the housing markets and reduces the lender's risk exposure. The presence of investors willing to fund mortgage loans alleviates the need for lenders to collect deposits to finance their loans. Moreover, for depository institutions, the secondary market reduces the risk associated with funding long term mortgages via short term deposits. A vibrant secondary market reduces the capital requirements necessary to start up a lending firm. The establishment of the secondary market has reduced the barriers to entry, and thus increased competition in the mortgage lending market.

2.3.1 *Securitizers*

The secondary market has existed since the establishment of the Federal National Mortgage Association (FNMA , or Fannie Mae) in 1938. Although it has existed since then, the role of the secondary market was minimal at its inception. Fannie Mae's primary responsibility was to create a market for loans insured by the Federal Housing Administration (FHA) and the Veterans Administration (VA). This allowed mortgage lenders to increase their origination of FHA and VA loans because they knew the mortgages would be purchased by Fannie Mae. In 1968, Fannie Mae's charter was adjusted to allow it to purchase conventional loans and FHA/VA loans became the responsibility of the Government National Mortgage Associations (GNMA or Ginnie Mae). Additionally, in 1970 the Federal Home Loan Mortgage Corporation (FHLMC or Freddie Mac) was established to purchase assets from Savings and Loans.

Fannie Mae and Freddie Mac are chartered as Government Sponsored Enterprises (GSE), which implies that although they are not government owned, they receive preferential treatment (McDonald and Thornton 2008). They are exempt from state and local taxes and have access to a $2.25 billion line of credit from the Treasury. This preferential treatment is due to the alignment of the GSE's mission with the government's objective of making homeownership more attainable. The GSEs approach to making homeownership more accessible integrates the mortgage markets with the capital markets.

As of 2006, the two GSEs had an outstanding debt of $4.47 trillion, almost equal to the $4.84 trillion of publicly held government debt (Poole 2007). Furthermore, their role in the mortgage securitization process has increased from 8% in 1981 to 50% of all

originated loans in 2003 (Jaffee 2010). Their main activity is to pool mortgage credit and create mortgage backed securities (MBS), which in turn are sold to institutional and non-institutional investors. The securitization process allows the lender to transfer the credit risk to the GSE, while the GSE receives the stream of mortgage payments and a guaranteed fee (for insuring the mortgage).

During the growth of the subprime segment of the mortgage market from 2005 to 2007, the market share of loans securitized by the GSEs declined. As part of the charter, GSEs have limits on the credit risk they can incur and the types of loans eligible for purchase. Loans which meet the GSE securitization guidelines are commonly referred to as conforming or agency loans. As part of the securitization process, the GSEs insure the mortgage backed securities investor against default. The agencies treat a mortgage default as prepayment of the loan, and therefore they are responsible for paying the investors if the loan were to default. To maintain an optimal level of credit risk, the GSEs restrict the loan-to-value ratio to 80% and a maximum loan amount (adjusted annually for inflation and housing appreciation).[9] Loans above the 80% loan-to-value threshold are still available for GSE purchase if the consumer obtains private mortgage insurance (PMI) to cover the amount exceeding 80%.

Loans that do not meet GSE criteria are *nonconforming* (subprime) to the guidelines and thus are sold to private label firms. Lehman Brothers[10], Salomon Brothers, General Electric, and insurance firms are some of the Securitizers in the private label

[9] 2009 conforming loan limit was $417,000 for a single family home, an increase of 65% from 2000. Alaska, Hawaii, Virgin Islands and Guam are considered high cost areas. The conforming limit in high cost areas is 50% higher.

[10] On September 15th 2008, Lehman Brothers Holding Inc. filed for chapter 11 bankruptcy. http://www.lehman.com/press/pdf_2008/091508_lbhi_chapter11_announce.pdf

market. These firms' services create a secondary market for loans that are not covered due to the guidelines imposed by the GSEs. With the exception of the provision of insurance, the private label securitization procedure is similar to the GSE's.

2.3.2 *Securitization Process*

During the process of a loan, it is originated, underwritten, funded, and eventually sold to the GSE. If the loan does not meet the conforming guidelines, it is sold to private label firms. The GSEs securitize loans through either the swap program or the cash program. The most common form of securitization is the swap program, where mortgage backed securities are traded to the lender for a pool of mortgages. The value of the mortgage backed security is net of guarantee fees paid to the agency and the cost of servicing (Ambrose, LaCour-Little et al. 2005). The exchange provides the lending firm with a liquid asset that can either be retained by the firm or sold to investors. In contrast, the cash program requires that the agencies combine smaller pools of loans purchased from multiple lenders and issue securities backed by the entire pool. Because the lender is paid in cash, it enables the company to realize profits from the sale in the current period.

Private label firms are at a competitive disadvantage to the GSE because their loans are not insured and do not have the implicit backing of the government. Consequently, mortgage backed securities issued by private label firms are riskier and usually priced 40 basis points[11] higher than comparable GSE securities (Passmore, Sparks et al. 2002). However, this estimate is misleading because in addition to not being insured, private label securities include riskier loans to begin with.

[11] A basis point is equivalent to $1/100^{th}$ of a percent. A 40 basis point difference is equal to 0.40%

Private label firms have three methods to reallocate and reduce the risk exposure to the final investor. First, they can separate the pool into tranches of senior and subordinate classes. The highest rated tranches (senior) usually have the least amount of risk exposure, but receive the lowest yield. Alternatively, the lowest rated tranche bears most of the credit risk and has the highest yield. The lowest tranche is the least desirable, usually described as toxic waste, and is difficult to sell. Consequently, toxic waste is mostly retained by the issuer, sold to unregulated institutions, or repackaged into collateralized mortgage obligations. Investors of mortgage backed securities or collateralized mortgage obligations do not have information about the risk of individual loans in the pool of mortgages. Investors acquire information about risk of the asset through the seniority of the tranche.[12]

The second method of risk mitigation is the overcollateralization of the mortgage security. In this case, each security is backed by mortgages whose values exceed the value of the securities. If a single loan defaults, the payment stream to the investor remains current. Nevertheless, the private firms will continue to incur some of the credit risk because a portion of the mortgages in the pool are unsold. Finally, the private label firm can purchase insurance from banks, insurance firms or the GSEs against the risk of default. Furthermore, derivative securities, such as collateralized mortgage obligations, were also established to further diversify the risk of default. Collateralized mortgage obligations pool the tranches of several mortgage backed securities to reallocate and diversify the tranches to create new securities.

[12] According to Downing et al (2005), the GSEs purposefully restrict information about the underlying assets to eliminate the lemons problem from the secondary market.

The development of the GSE and private label securitization has increased liquidity in the mortgage market. In addition to the liquidity gains, the involvement of the secondary market has also contributed to the segmentation of the upstream process. It has allowed origination services to be separated from the financing process; also, it has permitted the separation of loan servicing activities from the origination process. It is no longer essential for the lender to collect the monthly payments after the sale of the loan. Most lenders have divested loan servicing from their operations and sold the rights to other firms that specialize in collecting payments from the borrower. In return, the servicing firms receive the servicing fee, usually 25 basis points of the principle loan amount (LaCour-Little and Chun 1999).

The largest contribution of the secondary market to the mortgage industry is that it has standardized mortgage programs and made the mortgage process more transparent[13]. Lenders, originators, and underwriters processing mortgages can access the secondary market's publicly available guidelines. The transparency of the guidelines eliminates asset-specificity issues and reduces the potential of the holdup problem throughout the supply chain (Williamson 1981). Additionally, both Fannie Mae and Freddie Mac have further reduced the transaction cost to the lender by reducing the frictions in acquiring a loan purchase approval. Fannie Mae's Desktop Underwriter (DU) and Freddie Mac's Loan Prospector (LP) were developed to automate the credit scoring and decision making process. Lenders and originators can input the loan information in either the DU or LP software and gain immediate response from the GSEs. After the inception of the

[13] There are variations across states in mortgage documentation and programs due to variations in state mortgage regulation.

DU and LP programs in 1995, non-agency Securitizers provided their own automated credit scoring programs.

The automated underwriting system provides a risk rating for each loan, details the minimum requirements under which the Securitizers would purchase the loan from the lender, and sets the minimum price to accept for the risk (interest rate). For GSE systems the possible risk ratings are "accept," "reject," and "caution." Accept means that the loan will be purchased by the GSE pending verification of the information entered in the system. If the loan is marginally unacceptable and requires "quality redeeming" features, the loan receives a "refer risk" rating. In this case the lender needs to provide more information to justify financing the mortgage.[14] A loan with a caution rating will not be eligible for purchase by the GSEs. For non-agency loans, the applicable risk ratings are either eligible or ineligible for purchase.

The automated systems reduce the uncertainty associated with the mortgage process. Firms allocating resources to a loan are assured that their investment will be rewarded. Mortgage originators can focus on originating loans without concern about the financing, and lenders can finance the loan knowing that, if all criteria are met, the secondary market will purchase the loan. These innovations in the secondary market have drastically impacted the structure of the mortgage market.

The interaction of the growth of the secondary market with the technological evaluation has allowed the disintegration of the mortgage process and has facilitated the unbundling of services. A segmented market is able to specialize in the provision of a

[14] Information about the consumer's wealth accumulation or other assets may be required.

service and coordinate their actions when trading with firms down the supply chain. The market today is vastly different than the centralized mortgage market that existed previously.

2.4 Asymmetric Information in the Mortgage Market

Although technology has increased the flow of information, it does not guarantee that the information is either perfect or complete. The recent near-collapse of the mortgage market has increased skepticism about the quality of information transmitted in the mortgage market. The improved coordination and transparency between the market participants may have reduced the holdup problem, yet asymmetric information issues still exist and play a role in market transactions and the industry performance.

The segments of the supply chain are separated by the boundaries of adjoining firms. Although the coordination of firms across the boundaries has allowed them to specialize, there are concerns that information is not transmitted completely. Additionally, firms do not confront a common incentive structure, which usually leads to principal-agent problems. The segmentation and unbundling of the mortgage process increases the probability that a single firm does not distribute all available information to other participants in process.

2.4.1 *Consumer-Lender Relationship.*

The first relationship that may exhibit asymmetrical information is at the underwriting phase; it is the relationship between the consumer and lender. Once a consumer and lender have agreed on the mortgage choice, the lender underwrites the risk associated

with the loan and confirms that it meets the secondary market guidelines. To price the risk of the loan, during the underwriting process the lender combines all loan information to generate an application score (the automated system),[15] or origination score (Avery et al, 2000). In addition, the underwriter's objective is to verify that the information provided by the consumer is sufficient and acceptable to the secondary market. The origination score is assumed to be a more objective measure and provides the lender with a methodical approach to measure and price the risk of default. Lenders put forth a tremendous amount of effort to accurately measure the consumer's default likelihood and to ensure that the loan meets the guidelines. Consequently, lenders base their lending decisions on observable information provided by the consumer. The quality of the consumer is said to be high if the combined attributes of the loan suggest that the consumer is less likely to default. The borrower can assist the lender in assessing the likelihood of default and reduce uncertainty by making available complete information beyond the minimum amount required. The additional information is usually unobserved, but valuable to the lender and the secondary market. The unobserved quality measures can be in the form of wealth information, likelihood of continued employment, and debt information not appearing on the credit report.

The origination score is the combined score from the 3-C's. The *capacity* to pay the loan is measured by the debt-to-income ratio and employment history of the consumer. The value of the *collateral* is a function of both the value of the home and the loan-to-value ratio. Finally, the borrower's *credit* worthiness is signaled by the payment behavior on current and past credit, which is provided by the information in the credit report.

[15] Accept, reject, caution for the GSE loans, or eligible and ineligible for Private Label.

These three factors provide the minimum information needed for the firm to successfully decide if it will fund the loan through GSE or private label financing.

While the consumer knows his/her true quality, the lender is attempting to estimate the consumer's likelihood of default based on the information provided. Thus, the relationship between the consumer and the lender exhibits asymmetric information. Specifically, the consumer has more information about his/her likelihood of continued employment, while the lender has to approximate the borrower's future employment based on historical employment stability. Most lenders require that the consumer provide two years of continuous employment history, usually supported by W-2 documentation. Additionally, the consumer credit report does not provide the true measure of credit history or current outstanding debt. The credit report only includes information that is reported to the repository firms; any private debts and payment history on monthly utilities are excluded. The perceived default likelihood would decrease as the borrower's net worth increases; however, firms do not gain information about the consumer's savings and retirement accounts unless the consumer volunteers the information or is explicitly required by the automated system's conditional approval.

In the same context, if the consumer owns any other real estate, the lender will become informed only if there is a mortgage associated with the property and it appears on the credit report. If the consumer owns an investment property or a second home and has either paid the mortgage in full or purchased the property using their own assets, the lender will not observe any of the information. The unobserved information has the potential of adjusting the quality of the loan, but is usually omitted because the consumer is

unwilling to furnish the information or is unaware that the information is useful to the lender; also, this information is not required by the automated approval system.

2.4.2 *Originator-Lender Relationship*

Another instance where information may not be completely transmitted occurs between the originator and lender. This is especially the case when origination services are provided by a third party originator, and the objectives of the originator and lender do not align. The originator's service is to match the consumer with the lender, whereas the lender's objective is to finance a loan that will be of high quality.

The lender's reputation is at risk if it provides the secondary market with loans that default or prepay too frequently. If the lender supplies poor quality loans, the secondary market may restrict future purchases on recourse basis. The reputational threat would require due diligence by the lender to ensure that the probability of default or prepayment is minimized.

The originator does not absorb the cost of default, and would not be as meticulous about gathering complete information from the consumer. Moreover, if the originator has information that reduces the quality of the loan but is not required by secondary market guidelines, the originator is not required to furnish the information. The originator's objective is to process the loan and receive the commission and yield spread premium for their service. Neither revenue source is conditional on the loan performance.

The principal-agent problem reduces the quality of information that flows from originator to lender. If a loan receives an "accept" risk grading, then the originator does

not need to supply any additional documentation, even if it is available. This is especially important when the additional information may change the risk rating of the loan. For instance, if a borrower supplies the originator with a 401k statement to verify net worth but the loan guidelines do not require the verification of assets, the broker would not be inclined to submit that information to the lender. However, 401k statements can reveal information about the consumer's liquidity and can adjust the quality of the loan. A consumer who has taken a loan from the employer against the value of the 401k account would represent a higher risk than a consumer without a loan. To accurately measure the risk of default, it would be beneficial to the lender to have the 401k statement.

Studies examining the performance of loans originated by third party originators have found evidence of the presence of asymmetric information in the relationship between originator and lender. Alexander et al (2002) propose that the principal-agent problem evolves between TPO and lender because TPOs are compensated for the origination of the loan but do not confront the cost of poor loan performance. The empirical results confirm that loans originated by TPOs are more likely to default in comparison to loans originated by the lender. The higher cost of default by TPO loans was initially absorbed by the Securitizers. Since TPO and lender-originated loans were equally priced, market inefficiency resulted. To correct the market inefficiency and to reflect their higher risk of default, Securitizers eventually increased the cost of funding TPO loans.

LaCour-Little et al (1999) also examine the principal-agent problems that may transpire between the TPO and Lender. However, they study the differences in early repayment between TPO and lender loans. Prepayment is costly to the investor because it

alters the schedule of payments to the mortgage backed security. The probability of prepayment always exists, but it increases when market interest rates decrease, as refinancing becomes more favorable. Therefore, risk of prepayment is already priced by the secondary market; however, the authors find that loans originated by TPOs are more likely to prepay after controlling for other factors of termination risk. They associate the higher probabilities of prepayment to the incentive for TPOs to "churn" the consumer.

The contractual agreement between TPO and lender restricts the TPO from soliciting the consumer after the loan is closed. However, broker compensation is based upon the number of transactions and the volume of business the broker produces, and herein lies the principal-agent problem. The contract is too costly to monitor because the consumer is free to prepay the loan at anytime, and the lender does not observe the reasons for prepayment. Consequently, mortgage brokers are not bound by the contract and solicit consumers to refinance as often as possible.

There is evidence that the principal-agent problem restricts the amount of information transmitted from the originator to the lender and eventually to the secondary market. The efficiency gains acquired from specialization are mitigated due to the conflict of interest and the lack of accountability by originators.

2.4.3 *Lender-Secondary Market Relationship*

The relationship between lender and the secondary market is more defined, and asymmetric information concerns are minimized due to the contracts and incentive mechanisms. The ability of the secondary market to impose restriction on lenders ensures

that lenders are held accountable for their loan quality. The reputation effect maintains coordination and aligns the incentives.

The strength of relationship between lender and the secondary market can be tested using variations in the behavior of portfolio lenders. Ambrose et al (2005) examine the difference in default rates for loans sold to the secondary market and those retained on portfolio. Given the choice to sell or retain loans on portfolio, the reputation effect would predict that loans sold to the secondary market would be of higher quality. In fact, the results of the study support the reputation effect hypothesis because loans sold to the secondary market are less likely to default than those retained in the firms' portfolios. Therefore, firms are retaining the poor quality loans while selling the high quality loans. This would suggest that asymmetric information concerns between lender and the secondary market are negligible.

It is important to clarify that the data used in Ambrose (2005) is provided by a bank. Although a portfolio lender, banks have additional regulatory considerations that need to be accounted for. The authors cannot exclude that the variation in the bank's behavior is due to capital arbitrage opportunity that surfaces in response to the Basel Accords. Specifically, the Basel Accords (Basel I) requires that banks hold 8% of each loan as capital, and the reserve requirement does not vary with risk of the loans.[16] Therefore, for low risk loans the reserve requirement is too high, which creates the incentive to sell low risk loans and retain high risk loans in portfolio.

[16] In 2004 the Basel Accords where modified (Basel II) to allow for variations in the reserve requirements based on the loan quality.

The principal-agent problem between the lending segment and the secondary market is minimized due to the existence of complete contracts and a well established incentive structure. The main issue for the lender is in estimating the consumers' default risk with unobserved qualities. The revelation of the hidden qualities will not reduce the amount of loans supplied, but it will allow lenders to adjust the price of the loans to reflect the true quality.

2.5 Conclusion

The growth of the secondary market, in conjunction with the technological advancements in the 1990s, has influenced the unbundling of services in the mortgage market. Specialization and trade within the market have created efficiency gains. Coordination between the originating, underwriting, and financing service providers is necessary for information to flow through the supply chain. Although the market is complex and segmented, the standardization of mortgage products has made consumers more aware of the available products in the marketplace. Consumers are easily matched with programs, and funding is easily accessible. All factors have contributed in making homeownership more feasible.

The recent near-collapse of the mortgage market would suggest that the fundamental structure of the market may not be viable. The segmentation of the market may be credited with an increased level of efficiency, yet it is due to this that the quality of information can diminish. Unbundling of mortgage services must be accompanied with more complete contracts and alignment of incentives.

Third party originators' lack of accountability and the incentives they confront do not match the remainder of the market. A market which has an objective to price and allocate risk efficiently cannot have agents who do not incur the cost of mispricing risk. Correcting the incentive structure is critical and necessary for the market to operate efficiently. Without the correction, the existence of the TPO market is questionable. To align the incentives in the market, either TPOs will be forced to be accountable for loan performance or they will vertically integrate with the lender.

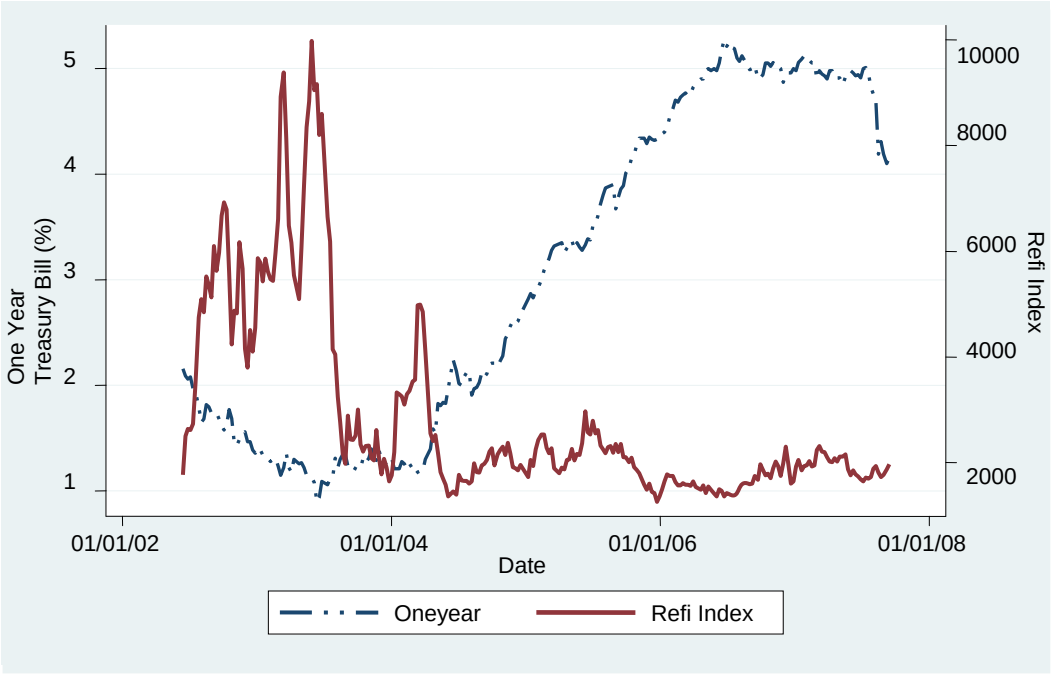

Figure 2.1 One Year Treasury and MBA Refinance Index from 2002-2008:
Mortgage Bankers Association weekly survey Refinance Index a measure of refinance application volume. The survey covers approximately 50 percent of all U.S. retail residential mortgage originations, and has been conducted weekly since 1990. Respondents include mortgage bankers, commercial banks and thrifts. Base period and value for all indexes is March 16, 1990=100. One Year Treasury Rates are collected from yahoo finance http://finance.yahoo.com/

Figure 2.2: The Mortgage Process and Market Segments

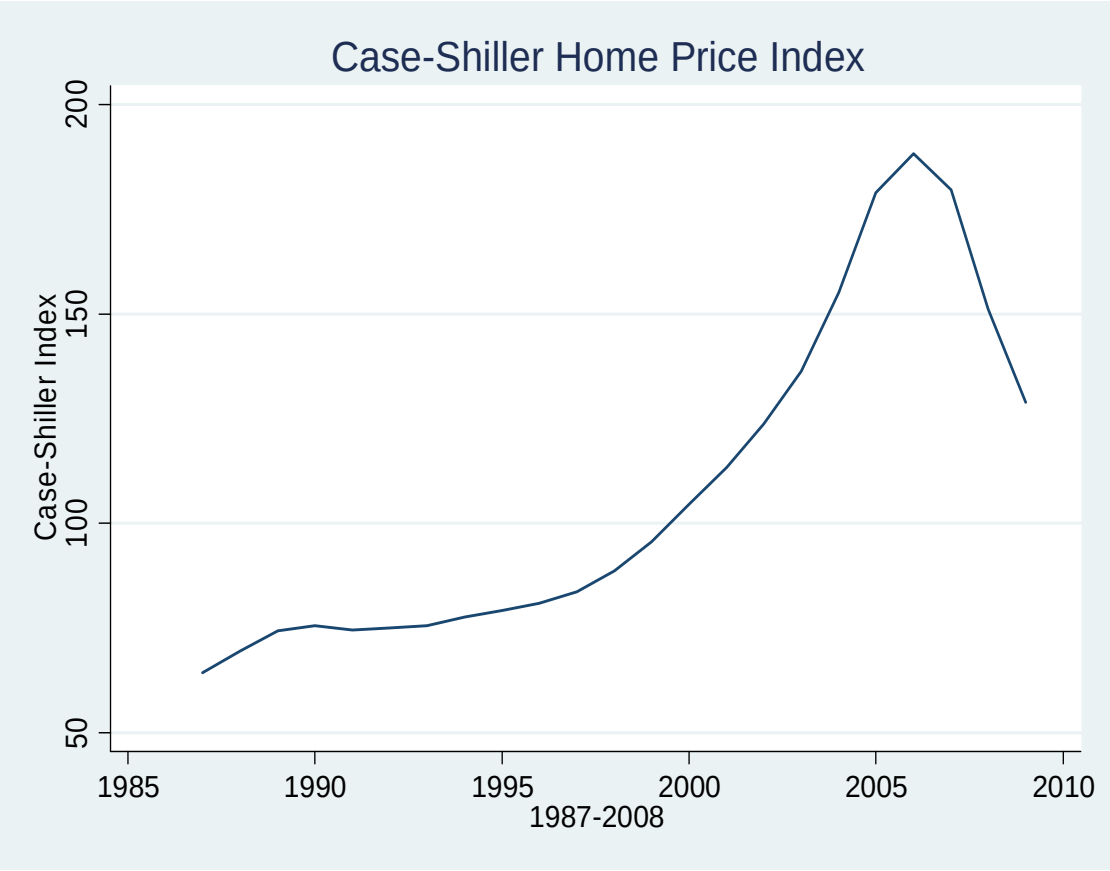

Figure 2.3: Composite Case-Shiller Home Price Index from 1985-2009
The S&P/Case-Shiller Home Price Indices (non-seasonally adjusted) are calculated monthly using a three-month moving average and published with a two month lag. Base year (January, 2000=100).

3. CHAPTER 3: A Model of Electronic Commerce in the Presence of a Gatekeeper

3.1 Introduction

The focus of this chapter is to examine the relationship between the consumer and originator of the mortgage. The objective is to develop a model which helps explain consumers' search processes and firms' pricing strategies in response to electronic commerce in the mortgage market. At the initial phase of the mortgage decision, and after the consumer has decided on obtaining a mortgage, the consumer must decide how to search for a mortgage product. Previously, consumers were required to physically visit their banks or other retail establishments to inquire about mortgage products and pricing. However, today, most consumers have access to the Internet and can now search through electronic markets.

The development of the Internet was expected to facilitate consumer search by eliminating transaction frictions, which would lead to more competitive markets. The drastic improvements in the availability of information to the consumer, in an environment of lower search costs, would create Bertrand competition in the market. To survive, firms are forced to compete by lowering prices toward the marginal cost of production. According to Stigler (1961) and Stahl (1989), it is expected that prices in online markets (low search cost and better-informed consumers) would be less dispersed and priced lower than traditional brick and mortar markets (high search cost and less-informed consumers). However, empirical studies which examine price differences between online and

retail markets are inconclusive and do not provide a clear consensus about the effects of the Internet on market prices[17].

While search cost models predict lower prices in online markets, they often assume that the consumer is able to observe all price quotes available on Internet Comparison Search (ICS) sites. The online aggregator websites (gatekeepers) specialize in providing information to the consumer, which eliminates a significant amount of searches. At a single site, consumers can view all available prices rather than visit each retailer's website or retail office. The benefits to the consumer in the form of convenience and lower prices have been previously defined by search cost models. However, the outlined benefits to the consumer imply that firms should be less willing to join the ICS sites due to the lower profits associated with increased information. A casual observation suggests that online markets are saturated with firms willing to provide pricing information to consumers through the aggregator.

Baye and Morgan (2002) provide a model of online markets and establish the equilibrium outcome in the presence of an information gatekeeper. The gatekeeper reduces search cost and increases information transparency, which allows the consumer to obtain a lower price. In response, firms reduce their prices to attract more customers. The equilibrium outcome is that online prices are lower than retail prices, and the gatekeeper restricts membership to the online market by charging an advertising fee to the firms. Consumers, however, are allowed to search for "free" in the online market.

[17] See Bailey (1998), Lee (1998), and Brown and Goolsbee (2002).

The premise for firm participation in online markets has been based solely on expanding demand, as firms are willing to join the gatekeeper's network for the opportunity to compete and attract a larger consumer base. Although online firms receive lower profit margins per unit, they are able to increase overall revenue—and thus also earn higher profits. I argue that a gatekeeper may purposefully reduce the competitive environment of online markets and thus attract firms interested in increasing their market power. The gatekeeper attracts customers through a promise of lower search costs and accessibility to more information; it provides the member firms with a flow of captive consumers who believe they are observing the lowest prices in the market.

Lendingtree.com is an online gatekeeper that operates in the mortgage market. The firm's market purpose is to reduce the consumers' search cost by matching them with mortgage suppliers. Their slogan asserts that "when banks compete, you win," which suggests consumers benefit by searching for mortgage prices on their site. However, LendingTree is an intermediary that also provides firms on its network with a flow of consumers who were previously unattainable due to spatial limitations. Moreover, LendingTree restricts the amount of price quotes the consumer observes by reducing the amount of firms competing to a fraction of the available firms on the network. When firm participation is restricted, the equilibrium prices may increase and firms may receive higher profits, which are unattainable without the presence of the gatekeeper. This setting differs from previous models, which focused on gatekeepers (or shopbots) that provide the consumer with price quotes from all firms participating on the network. This change has neglected the gatekeeper's anticompetitive conduct.

In this dissertation, I provide an alternative explanation of the online market, where a gatekeeper restricts a firm's membership through price and thus only a small part of all firms will participate on the network. Additionally, consumers searching for price information will only observe a share of the available quotes. The effect on market outcomes will therefore depend on the gatekeeper's conduct, and whether the search cost effect dominates the anticompetitive effect.

3.2 The Consumer and Search Process

For example, let there be two types of consumers shopping for a homogenous good. However, the consumers are heterogeneous in their cost of search for the good. The first type of consumers have zero search cost (i.e zero opportunity cost of time), such that they are willing to search over all available firms to gain price information. The second type of consumer has high search costs (high opportunity cost of time), and will search for price information at a subset of the available firms. The subscript *i* denotes the consumer type, where *i=1* is the consumer with zero search costs and *i=2* denotes the consumer with high search costs.

3.2.1 *Consumer Search Without Internet*

Each consumer has a demand function $D(p)$. The consumer's surplus when observing a price P is given by eq 3.1:

$$\boldsymbol{CS(P) = \int_b^P D(x)dx} \tag{3.1}$$

The decision to conduct an additional unit of search depends on the expected consumer surplus from the search. Given a price P, the expected consumer surplus from an additional unit of search will also depend on the distribution of price in the economy, given by F(p). Therefore, the expected consumer surplus of an additional unit of search is given by integrating eq 3.2 with respect to the cumulative price distribution F(p):

$$\boldsymbol{ECS(P) = \int_b^p CS(p)dF(p)} \tag{3.2}$$

Here, b is the lower bound of the price distribution. If the expected gains to consumer surplus exceed the marginal cost of search, $ECS(P) \geq C_i$, then the consumer will conduct an additional unit of search. Otherwise, the consumer will stop and purchase from the firm with the lowest observed price.

For a given price p', the expected gains of an additional unit of search for the consumers with zero search cost will exceed the expected gains for high search cost consumers. Therefore, type one consumers will undertake more search than type two consumers. Let S denote the amount of search for each consumer type:

$$\mathbf{S_1(p') > S_2(p')} \tag{3.3}$$

Zero search cost consumers will visit all firms to gain price information and purchase from the firm with the lowest price. High cost consumers will search for the lowest price up to the point where the gains from additional search equal the cost of search.

3.3 Firm Behavior without the Internet

The firms know that there are two types of consumers that exist in the market, where μ fraction are type one consumers and (1- μ) are type two consumers. The firm cannot identify the individual consumer type but knows their proportion in population. Let there be N firms, all with constant marginal cost of production. The objective of the firm is to maximize expected profits subject to consumer types and demand. Firms do not observe consumer types and cannot price discriminate based on observable information. The firm will maximize expected profit and a single price will prevail in equilibrium. For simplicity I normalize the measure of consumers to unity and let $\mathrm{pD(p)} = \mathrm{R(p)}$. Firm i's maximization problem is given in Stahl (1989) and is represented as:

$$\mathbf{E}\boldsymbol{\pi_i}(\boldsymbol{p}, \boldsymbol{F}(\boldsymbol{p})) = \frac{(\mathbf{1}-\boldsymbol{\mu})\mathbf{R(p)}}{\boldsymbol{N}} + (\boldsymbol{\mu})\big(\mathbf{1} - \boldsymbol{F}(\boldsymbol{p})\big)^{\boldsymbol{N-1}}\mathbf{R(p)} \qquad (3.4)$$

The first term represents the fraction of high search cost consumers, which are shared by N firms in the market. The second term represents the revenue from the zero search cost consumers, where the firm obtains their business if their price is lower than N-1 other firms. Finally, the firm has constant marginal cost $\delta = 0$.

The equilibrium is given by the cumulative price distribution that is a solution to:

$$\int_b^r \boldsymbol{CS}(\boldsymbol{p})\boldsymbol{dF}(\boldsymbol{p}) = \boldsymbol{C_i} \qquad (3.5)$$

and:

$$\boldsymbol{F}(\boldsymbol{p}) = \mathbf{1} - [\left(\frac{\mathbf{1}-\boldsymbol{\mu}}{\boldsymbol{N\mu}}\right)\left(\frac{\boldsymbol{R}(\boldsymbol{P_r})}{\boldsymbol{R}(\boldsymbol{p})} - \mathbf{1}\right)]^{\frac{\mathbf{1}}{\boldsymbol{N-1}}}, \qquad (3.6)$$

where $R(P_r) = min\{r, P^m\}$; r denotes the reservation price while P^m is the monopoly price. As the search cost decreases, the price distribution converges towards the Bertrand outcome, with price equal to marginal cost.

3.3.1 *Firm Advertising Decision*

The firm's decision to advertise is a function of the advertising fee set by the gatekeeper and the fraction of firms the gatekeeper allows the consumer to sample. If the gatekeeper drastically restricts the number of firms able to compete, market power increases for the firms that are allowed to compete but reduces the probability the firm is selected as one of the competitors. If the gatekeeper allows all firms to supply a quote, the market power of the firm's advertising falls but the probability that the firm's quote is observed by the consumer increases. Therefore, there is an optimal level of restriction that will maximize firm participation. Simply, the firm will advertise only if the expected profits in online markets are equal or larger than the profits in retail markets. Let π_i^G denote profits through the gatekeeper (online) and π_i^R denote the profit in retail markets. For a firm to advertise it must be that:

$$\pi_i^G \geq \pi_i^R \tag{3.7}$$

3.3.2 *Firm Pricing with the Internet*

Assume that the online market has a small number of gatekeepers. Each of the N firms can choose to advertise by paying an advertising fee. In the mortgage market there are few gatekeepers that provide price comparison services. The market is dominated by

LendingTree[18], but others like Lowermybills.com and Nextag.com also provide the same service. Consequently, each firm can choose to advertise in any or all of the gatekeepers' sites. Since gatekeepers charge a fee for firms to advertise, a total $n \leq N$ firms will join the gatekeepers' network. Furthermore, the gatekeeper further restricts the number of quotes the consumer observes to $K < n \leq N$.

Given the decision to shop online or offline, a fraction μ of consumers will decide to shop offline and $(1-\mu)$ will shop through online markets. It can be assumed that consumers that choose to shop in retail markets will conduct their entire search within the offline market, and the fraction $(1-\mu)$ online consumers will search through gatekeeper markets only. In both offline and online markets, zero and positive search cost consumers exist. Let μ_1 represent the fraction of offline consumers with zero search cost, and μ_2 denotes the fraction of online consumers with zero search cost.

Consumers deciding to search through gatekeeper sites are interested in reducing the cost of their search. Searching online allows consumers to do this because they observe more quotes per unit of search. The fraction of consumers in online markets $(1-\mu_2)$ with positive search cost will pay an effective cost of C_i^G for each unit of search. Thus, we can assume that the search cost in online markets is less than the search cost in retail markets:

$$C_i^G < C_i^R.$$

[18] Chadwick Martin Bailey, Inc. conducted a study for LendingTree in June 2001 which measured the Total Brand awareness for firms in the online lending market. The study found that LendingTree has 2.5 times the awareness level of direct online competitors and exceeds the awareness of the top retail bank in the study.
http://www.thefreelibrary.com/LendingTree+Extends+No.+1+Brand+Awareness+Position+in+Online+Lending.-a078011256

The described online and retail markets are two separated markets which may have different prices. Each market will have an equilibrium price distribution as described in equations (3.5) and (3.6). For the retail market, the equilibrium is given by the price distribution $F(p)$that solves:

$$\int_b^r \boldsymbol{CS(p)dF(p) = C_i^R,} \tag{3.8}$$

and:

$$\boldsymbol{F(p) = 1 - [\left(\frac{\mu-\mu_1}{N\mu_1}\right)\left(\frac{R(P_r)}{R(p)} - 1\right)]^{\frac{1}{N-1}},} \tag{3.9}$$

For the online market, the equilibrium is given by the price distribution $G(p)$ that solves:

$$\int_b^r \boldsymbol{CS(p)dG(p) = C_i^G,} \tag{3.10}$$

and:

$$\boldsymbol{G(p) = 1 - [\left(\frac{(1-\mu)-\mu_2}{N\mu_2}\right)\left(\frac{R(P_r)}{R(p)} - 1\right)]^{\frac{1}{K-1}},} \tag{3.11}$$

The difference between the two solutions will depend on the parameters $\boldsymbol{C_i^G, C_i^R, \mu, \mu_1, \mu_2, N}$ and $\boldsymbol{K}$. Given that $\boldsymbol{C_i^G < C_i^R}$ **,** the price in retail markets will be higher than prices in online markets. On the other hand, the number of firms competing in online markets is less than the number of firms competing in retail market, $\boldsymbol{K < N.}$ The lower number of firms competing in the online market will cause price in online markets to exceed the price in retail markets. Consequently, the effect of online commerce on market outcomes is ambiguous and will depend on the gatekeeper's conduct. If the gatekeeper

creates a market where the search cost effect dominates the anticompetitive effect (small K), then prices in online markets will be lower than prices in retail markets. Alternatively, prices in online markets will be higher if search cost effect is dominated by the effect of a small number of competitors.

3.4 Gatekeeper Choice Variables

The gatekeeper will influence the market outcomes given its choice variables. The gatekeeper maximizes profits with respect to the subscription fee, the advertising fee, and the fraction of firms it allows to compete for each consumer (K). The gatekeeper's objective is to provide consumers and firms with benefits to joining the site, but extract as much of the gains as possible. Consumers will find it beneficial to join online search if their effective search costs are reduced.

The number of firms that choose to participate is a function of the advertising fee, the number of consumers shopping through the gatekeeper site, and the fraction of firms the gatekeeper allows to compete per quote (K). The gatekeeper needs to create an incentive for firms to participate and must choose the optimal (K) to insure firms receive higher profits. Additionally, the gatekeeper's advertising fee will extract profits from the firm. The fee can be a linear price or a two-part tariff.[19] The profit to the firm must be larger or equal through the gatekeeper relative to retail market profits.

[19] LendingTree charges multiple fees: a onetime membership fee, a "quote fee" paid for the right to quote a price, and finally a "closing fee" paid by the firm with the winning bid.

3.5 Conclusion

Search models assume that the Internet would facilitate trade by eliminating transaction frictions and consequently lead to more competitive markets. The drastic improvements in the availability of information to the consumer, in an environment of lower search costs, would create Bertrand competition in the market. This would force firms to compete by lowering prices toward the marginal cost of production.

Empirical studies examining price differences between online and retail markets have been inconclusive and do not provide a clear consensus about the effects of the Internet on market prices. In this chapter, I provide an alternative explanation for the inconsistent empirical results. The competitive outcomes predicted due to lower search cost on the Internet are sensitive to the assumptions that the online consumer can observe all prices. However, when a gatekeeper is present, the gains from reducing search cost to the consumer may be offset by the gatekeeper's conduct. When a gatekeeper restricts the quotes observable to the consumer, an anticompetitive effect exist. If the anticompetitive effect dominates the search cost effect, prices in online market will be higher than in retail markets.

Therefore, the inconsistency in empirical results when examining the effect of the Internet on prices may be due to variations in gatekeeper conduct. An empirical test examining prices and profits in online versus retail markets would assist in identifying which effect dominates, the search cost effect or the anticompetitive effect.

4. CHAPTER 4: Joint Determination of Consumer Search Behavior and Mortgage Pricing

4.1 Introduction

To accurately measure the effect of the internet on prices, two simultaneous effects must be accounted for. The first is the gains that are associated with lower search costs and an increase in availability of information to the consumer. On the other hand, firms may be able to collude because they can identify with certainty their direct competition and these competitors' quoted prices. In a collusive environment, the transparency of pricing information reduces the firms' cost of monitoring its competitors. Thus the efficiency gains in online markets associated with lower search cost may be diluted by the gatekeepers' conduct. When comparing price differentials in online and retail markets, empirical research has attributed any price difference solely to differences in consumer information; yet, the online and retail market structures differ in more dimensions.

The literature has assumed that consumers are essentially identical and heterogeneous only in their search costs and consequently their information sets. Each consumer's search method is random in nature, and thus consumer characteristics do not influence their choice of market type. Additionally, firms are assumed to have homogenous cost structures regardless of their market of operations. This suggests that online firms are identical to retail firms and differ in their pricing strategies only. Although the assumption helps simplify the model and can attribute price difference to consumer information, it is impractical in an empirical setting. Assuming away differences in cost across firms

also eliminates the possible scenario that firms price at their marginal cost; prices would then be dispersed according to variation in marginal costs.

I will present evidence of pricing differentials in online and offline markets for a single firm. Using micro-level data from a mortgage firm, I am able to examine the firm's pricing strategies for their online customers relative to their retail customers. By examining a single firm I am able to eliminate firm heterogeneity as the reason for any price differences and can reveal any differences in pricing strategies associated with the consumers' information sets and market structure. I examine the joint determination of consumers' choices between online and offline markets and the price they pay for a mortgage. This construct allows consumers' characteristics to determine their method of search. The estimation method will allow for the interdependence between mortgage interest rates, profits, and the consumers' market choices. I control for the interest rate and closing-cost trade off, which eliminates the prevailing endogeneity problem in mortgage pricing research. This is possible in this study because of the availability of a unique data set which contains cost and revenue information and provides micro-level information about one particular firm's operations and consumer characteristics.

Using a switching model to control for consumer selection into market type, I find two results: First, consumers' unobservable characteristics do not influence their choice of commerce which supports previous literatures assumption that consumer choice to conduct searches online is a random process. Second, once consumers have made their choice and after controlling for observables, online customers pay the same price on average as their retail counterparts. Therefore, I find that access to less-costly information

in online markets has not led to lower prices of mortgages (rates). However, the results reveal the presence of an anticompetitive effect that offsets the reduced search cost effect. In a number of circumstances, and for a subset of loan types, the anticompetitive effect exceeds the search cost effect. Firms operating in the gatekeeper format are more able to price discriminate in online markets by loan duration and loan type, whereas the retail firm cannot differentiate its price with respect to consumer characteristics. Therefore the gains to the consumer from lower search costs are eliminated by the anticompetitive market structure created by the gatekeeper in this market. When controlling for the variation in marginal cost between online and retail markets, I find that the firm earns higher profits in online markets. This further highlights the presence of an anti-competitive effect in online market, which I argue is due to the dynamic nature of online markets and the gatekeeper's conduct. These results indicate that market structure plays a role and must be considered in further empirical studies when comparing online and retail price markets.

This chapter will progress as follows: Section 2 details the previous literature focusing on search costs and market structure. Section 3 includes background on the mortgage firm and internet comparison sites (ICS). Section 4 includes the empirical model used for estimation. Section 5 discusses the data and summary statistics. Section 6 includes the results. The chapter concludes in Section 7.

4.2 Literature

4.2.1 *Search Theory*

Neoclassical economics suggests that competitive markets with no externalities operate efficiently and that, because of the availability of information, a single price will

equilibrate both supply and demand. These perfectly competitive markets are characterized by the law of one price, which means that all goods that are homogenous are priced identically and consumers are fully aware of the market-clearing price. This strict assumption that all agents operate with full and perfect information helps simplify many complex economic models and thus is useful in many applications. Yet, theoretically recognized, perfectly competitive markets are rarely observed in daily consumer transactions (Stigler 1957).

Markets seldom have a single price that satisfies demand and supply, and usually we observe a dispersion of prices for what is characterized as a homogenous good. One can easily justify this anomaly as an outcome of a market that is in disequilibrium (Stiglitz 1979). However, price dispersion is overwhelmingly present in markets and thus raises questions about the underlying assumptions of the model. Economists interested in the phenomenon of price dispersion have cited two possible reasons. First, the assumption of homogeneity is hard to establish. Goods sold in markets may not differ in the basic characteristics of the product and may seem homogenous; however, firm attributes may add (or detract) value to the product. Stigler (1961) expresses that absolute homogeneity of goods is nearly impossible when consumers include service by the firm or any other supplier characteristics. More important to the topic of this paper is the assumption of full and perfect information. A consumer entering a new market may not know the prevailing price in the market. To gain this information, the consumer must partake in a search process.

The lack of information about the product or firm characteristics, as suggested by Stigler, violates the neoclassical assumption. Thus, prices in the economy may in fact be dispersed and consequently deviate from the expected uniform price outcome. The economics literature has examined the consequence of violating the perfect information assumption.

Theoretical literature on price dispersion often credits Stigler (1961) for establishing the literature of costly information in his seminal paper, "The Economics of Information." A model of ignorance is first established by relaxing the Neoclassical assumption of full information. The model considers consumers in search of a good who are unaware of the price of the product in the market, and the assumption holds even if we assume the law of one price. Consumers must undertake the costly search for the price that meets their reservation price. The amount of search necessary is a function of individual characteristics, specifically the opportunity cost of time and the fraction of income spent on the good. Consumers are heterogeneous in the cost of search for information, and consequently firms can vary their prices according to this variability in consumer search costs. In the market, prices would be dispersed and would therefore violate the law of one price.

As consumers search, they gain information about the price dispersion and since consumers buy at the lowest offered price, this induces firms to lower their price and thereby reduces the dispersion of prices. Consumers have a higher incentive to search for price information for goods that are purchased frequently, by incurring the search costs only once and reducing their purchase price for every transaction thereafter. Thus we expect prices to be more dispersed in a durable goods market since consumers purchase this

type of good less frequently. Sorensen (2000) examines the price of prescription drugs in two different markets and observes that drugs purchased often are more likely to have a lower price. He attributes this to the consumer's attempt to minimize the price of frequently purchased goods, as search theory predicts.

Stigler finds that expected returns to additional search decrease with the amount of search. Also, the expected savings from searching will increase with the price dispersion observed in the markets. He states that there is an optimal amount of search, and it is possibly sequential in nature but leaves this point for others to explore. Finally, he forecasts that, because of consumers' desire to reduce search costs, firms will emerge whose sole purpose is to facilitate the dissemination of information to both buyers and sellers in the market.

A conclusion provided by Stigler's model is that price uniformity is only possible when search costs are zero. If all consumers can search without any costs, then they have an incentive to shop all prices and will choose the lowest price. Firms respond by reducing their prices to a uniform price. In equilibrium the perfectly competitive price will prevail. If a uniform price exists and is above marginal costs, then each firm has an incentive to reduce its price marginally and attract the entire market. Thus all firms will have normal profits in equilibrium with a price equal to marginal cost. Stigler proposes that it is unprofitable for firms to eliminate all dispersion in the market and may seek to create some information uncertainty. Recent research by Carlin (2009) has focused on information complexity as a strategies of keeping price above marginal cost. Carlin shows that as consumers become more informed, the firms will increase the cost of information by add-

ing complexity to their pricing structure. Consumers are therefore less likely to become informed, and their ignorance is a source of market power to the firms, as Scitovsky (1950) proposed. Ellison and Ellison (2009) also provide a model and some evidence of how firms in online markets are more likely to add friction in consumers' search for information.

Hong and Shum (2006) established a model that allows estimation of search costs when observing a price distribution for a homogenous good. Their model is used to estimate the search costs associated with two types of search processes, sequential and non-sequential. A sequential search process is where the consumer, after each price observation, can decide to take the lowest price observed or continue to search for additional price quotes. A consumer employing this strategy will continue to search as long as the expected gain from finding a lower price exceeds the cost of search. Alternatively, a non-sequential search is a process where the consumer purchases from the lowest priced firm after randomly sampling a predetermined number of firms. In the context of the mortgage industry, a consumer shopping on ICS sites is using a non-sequential search process, while the retail customer's process is sequential.

The estimation procedure of Hong and Shum (2006) finds that sequential shopping rules lead to higher search costs relative to non-sequential shopping rules. However, Morgan and Manning (1985) indicate that non-sequential search is more favorable when there are positive fixed costs of searching. This is the case of searching for a mortgage. Organizing, collecting personal documentation, and completing a loan application are fixed costs. Whereas, the marginal costs of search are the cost of locating a firm, submit-

ting the application, and discussing the loan with the bank/originator. The authors argue that online price shopping meets the assumptions of non-sequential search. However their measure of search costs is sensitive to the parameters in the model.

To obtain a measure of search costs in Hong and Shum's model, it is necessary to include strict assumptions about the quantities purchased, specifically that each consumer purchases one unit of the good. A good that would meet that criterion is the market for books. However, consumers rarely purchase one book at a time. Thus calculations of search costs might be biased upward, since consumers are minimizing their search costs for the bundle of goods. Empirical tests of search costs usually do not have any information on the consumer and therefore must make assumptions about their characteristics. My mortgage data includes consumer information and thus reduces the assumptions associated with estimation.

Hong and Shum argue that a measure of search cost would be more interesting if there are detailed individual-level data sets, which would help in identifying the consumer's search costs. Additionally, the authors do not consider the possibility that consumers can move from one type of search process to the other. More precisely they assume that sequential shoppers will remain sequential indefinitely. For this study, the assumption is that shoppers do not switch from online markets to retail markets and vice versa.

Another expectation is that the reduction of search costs due to the use of ICS sites would increase the information available to consumers about the price distribution in the market. Stigler (1961) suggests that as consumers become more aware of the price distribution, firms competing for consumer business will reduce their markup to attract

the consumer.[20] Thus as consumers gain information about the underlying distribution of prices, the market price will approach the Bertrand-Nash equilibrium. Markets with informed customers will pay a uniformly lower price and firms will receive normal profits when customers experience zero search costs. On the other hand, if customers have high search costs and therefore do not actively search for the lowest price, the average price will increase and the market will observe a dispersion of prices.

Stigler's model assumes that consumers are homogenous and gain information uniformly. Stahl (1989) extends Stigler's theory to allow for both informed and "totally ignorant" consumers to exist in the market. However, in Stahl's model, consumers know the Nash equilibrium (NE) distribution of prices before they begin their search and obtain each firm's pricing information as they search. Firms set their price in response to consumers' search cost and the NE price distribution. The result of the model is that consumers' reservation price is endogenously determined, and prices are dispersed in equilibrium due to firms' mixed strategy. Additionally, Stahl's model supports both the Bertrand and Diamond equilibrium. Diamond (1971) presents a model with heterogeneous consumers, where each individual has a varying cost of search, but if all consumers have a positive cost of search, the market price will be the monopoly price. In that case each firm has an incentive to increase its price by an amount equal to the cost of search to the consumer with the lowest cost. If each firm follows these rules, then in equilibrium one price exists at the monopoly price. Stahl's model yields the Diamond equilibrium: When the population with zero search costs is set to equal 0 (all consumers are totally ignorant), the equi-

[20] Markup margins are referred to as yield spread premiums or overages.

librium will yield the monopoly price outcome; when it is set to 1 (all consumers are fully informed), the equilibrium result is the Bertrand outcome.

In the context of this research, customers applying through retail locations are perceived to have higher search costs. Individuals applying online use the internet to reduce their cost of search. The firm predicts that, because of the high cost of search, customers will not incur the cost of visiting multiple retail locations. Higher search cost increases the probability of the quoted price being the lowest price observed by the customer, thus increasing the incentive for the firm to quote a price higher than the competitive price. The firm can gain a customer's business as long as the price offered is marginally below the client's reservation price.

4.2.2 *The Internet as a Search Cost Reducer*

Empirical research has been inconclusive in determining whether the internet has lowered the price of online products relative to retail prices. By investigating the car market, Lee (1998) finds that car prices online are higher relative to retail prices, while Morton et al. (2005) find that new cars sold on online markets are priced 2.2% lower than in retail markets. Lee concludes that the quality differences of second-hand cars in online markets exceed that of cars sold at retail dealerships and as a result is capturing the price differential due to quality. He argues that the dispersion in prices is due to the heterogeneity of the good. Moreover, he attributes the higher prices to online customers' willingness to pay a premium for the convenience of purchasing a car without having to leave the comfort of their homes.

Bailey (1998) examines the market for books, CDs, and software and finds that online prices are higher relative to prices at retail stores. He attributes the higher prices to the infancy of online markets and predicted that lower prices will emerge as online commerce flourishes and consumers use online markets more frequently to gain information. Bailey's research differs from Lee's by examining a market of homogenous goods. In addition to the concerns due to quality, Bailey cites inefficiency in Lee's study associated with the online market examined. Lee's study examined online auctions for second-hand cars, specifically. A characteristic of an auction is that the good is sold to the consumer with the highest willingness to pay, and therefore auction prices are likely to be higher in general when compared with sales between two individuals.

Brown and Goolsbee (2002) examine the relationship between the price of term life insurance and a measure of each state's internet usage as an instrument for search activity and information. They find that an increase in internet usage decreased price dispersion as well as average price of term life insurance. They also measure the premiums before and after the development of the internet and find that the premiums after the internet are lower and less dispersed. They conclude that the internet reduces search costs and therefore reduces the market price and increases consumer surplus.

4.2.3 *The Internet as an Anti-competitive Market*

The view of online markets as anticompetitive environments has been theoretically represented but mostly neglected in the empirical literature. The overwhelming belief that the internet is a benefit to the consumer has predominated. Baye and Morgan (2002) developed a model that incorporates the presence of a gatekeeper in online markets. They

find that although the gatekeeper can successfully limit its membership, the severe Bertrand competition will lead to lower prices in the online market. However, their model centers on a gatekeeper that does not limit the available quotes to the consumer and provides the consumer with pricing information from all of the members on the network. They also provide the optimal fee structure for the gatekeeper, such that the fee should be imposed only on firms and not consumers. The rationale is that firms will be willing to pay for the flow of consumers that are provided by the gatekeeper. Their approach simplifies the gatekeeper's role to an advertiser for the network of firms and disregards any market structure implications.

Chen et al. (2002) developed a model where "infomediaries" refer consumers to retailers. They find that the intermediary benefits the consumers by facilitating the flow of information. A firm's profits are U-shaped with respect to the intermediaries' reach, and its profits increase as a result of the increase in demand attributed to the firm's ability to gain consumers who had not been accessible to them previously. However, since the firm is encroaching on the retail competitor's clientele, eventually it will create a competitive effect, where prices and profits are reduced for both the retail and online firms. Finally, when there are equal fractions of consumers in both the internet and retail shopping markets, the firm can price discriminate by providing different prices in online and retail markets.

An alternative approach is to consider the dynamic environment the gatekeeper creates. Campbell et al. (2005) provides a model that takes into account the dynamic nature of competition in online markets. The approach that search models take has been

static in nature and measures price variation with respect to a difference in search costs. When including firm behavior and their ability to monitor each other's prices, even when observations are imperfect, it is easier for firms to collude. Thus the welfare gains due to lower search costs are eliminated by the anticompetitive market. When firms cannot monitor each other's prices and if the fraction of shoppers increases, average prices fall as predicted by search theory.

Given the mixed findings when examining the effect of the internet on prices, recent empirical studies have begun to hypothesize and test for market power and profitability inonline markets. Hitt and Frei (2002) study the differences of the characteristics and behaviors of consumers in the online and retail market. Their study focuses on the banking sector and finds that online consumers are more profitable to the bank due to unobservable characteristics. Although their study does not specifically examine prices, their findings support the existence of a more profitable online market relative to the retail market. They attribute the higher profits to the loyalty of online consumers, higher balances and their greater propensity to adopt new banking services. This evidence further supports the hypothesis that consumers' behavior in online markets provides the firms additional market power.

A study by Viswanathan et al. (2007) explores the role of gatekeepers in online markets and provides an alternative explanation for the inconsistent empirical findings. They distinguish the types of gatekeepers by the information provided to the consumer. Gatekeepers can be divided into two categories; those that provide product comparisons and those that allow for price comparisons. Examining the automobile industry, they find

that consumers pay a higher price when the consumer searches online to compare products (i.e comparing car models). Alternatively, when consumers are conducting searches on pricing information and are comparing prices, they pay lower prices in online markets relative to the retail market. This study is the first to highlight the differences among gatekeepers and how the differences might impact the market outcome. Since gatekeepers differ in the information they provide, information serves as a mechanism for firms to segment the market and price discriminate accordingly. Thus, not all online searches will lead to consumers paying lower prices relative to the retail market. The price consumers pay will ultimately depend on the type of information they search for.

4.3 The Mortgage Industry: Date Source, Markets and Consumers

4.3.1 *The Mortgage Industry and Data Source*

According to the Mortgage Bankers Association, in 2009 total mortgage transactions totaled $2.1 trillion dollars.[21] Outstanding mortgage debt accounts for $10 trillion, and represents roughly one quarter of the total outstanding debt in the United States. Home equity accounts for one fifth of total household wealth accumulation (Belsky and Prakken 2004). However large and important, the mortgage industry remains a market characterized by asymmetric information. Loan brokers have access to multiple loan programs and wholesale pricing rates that the customer does not observe. The market has a high degree of product differentiation, which yields higher search costs to the consumer. As predicted by Stigler (1961), new firms emerged in the market to facilitate the transmission of information between sellers and buyers. Beginning with LendingTree in 1997,

[21] In 2004, the peak in mortgage transactions, mortgage originations totaled $4 trillion dollars http://www.mbaa.org/files/Bulletin/InternalResource/73418_.pdf

several ICS sites began to bridge the information gap.[22] Customers searching for mortgage products and price information are now able to obtain it through these ICS sites, therefore creating an environment where firms compete for the clients' business.

The data used in the empirical study in this paper were acquired from a mortgage firm after it became insolvent in August 2007. The firm purged the data of any consumer personal information, and approved its use solely for academic research. The mortgage firm operated as a mortgage broker and banker headquartered in Louisville, Kentucky, and conducted business in 22 states. To expand its customer base and diversify its business beyond Kentucky, the firm joined several ICS sites. The firm classified itself as an online mortgage company in view of the fact that most of its business was obtained through the ICS sites. However, the firm did not limit its operations to ICS customers; services were offered through the main branch in Louisville, Kentucky, as well. Retail operations were supported via local advertising, telemarketing, customer "call ins," and referrals.

The firm was established in 2002, during what has become known as the "refi boom." After the first year of business, it joined the LendingTree network and eventually 18 other ICS sites. Nonetheless, LendingTree accounted for 83% of the firm's online

[22] Lowermybills.com, mortgagequote.com, banxquote.com and Bankrate.com are some of the other ICS. On May 19th 2010, Banxquote.com filed an antitrust law suit against LendingTree for price fixing and collusion with other competitors. http://finance.yahoo.com/news/Federal-Antitrust-Complaint-prnews-2134255318.html?x=0&.v=1

business. The firm was able to compete for customers who applied through the ICS sites for a mortgage in any of the 22 states.[23]

4.3.2 *The Internet*

Consumers interested in receiving a quote for a mortgage can apply online through LendingTree. LendingTree is the industry leader in online lending and mortgage comparison sites.[24] According to the LendingTree's latest 10-K filing to the Securities and Exchange Commission (SEC), their biggest competitors in the price comparison market are lending institutions entering the online market.[25] While firms like lowermybills.com, and nextag.com are in the mortgage price comparison market, they cannot compete with the brand awareness of LendingTree. A study by Chadwick Martin Bailey, Inc. in June 2001 found that LendingTree has 2.5 times the awareness level of direct online competitors and exceeds the awareness of the top retail bank in the study.

During the application process with LendingTree, customers provide personal information including social security number, requested loan amount, and loan program. This information enables LendingTree to obtain a credit score for every applicant. Each application is sent to 5 firms[26] on the network[27] that are licensed in the state the property is located in. LendingTree has a complex predictive modeling system that matches lenders and customers to ensure the highest probability of success. This matching process takes into account customers' credit score, the loan program requested, the lenders' cus-

[23] The Firm Operated in AK, CA, CO, CT, FL, GA, IL, IN, KY, MA, MD, ME, MI, MN, MO, NC, OH, TN, TX, VA, WI, and WY
[24] http://www.mortgagemag.com/news/2009/0316/1000010309070.htm
[25] http://investing.businessweek.com/research/stocks/financials/drawFiling.asp?formType=10-K
[26] Loan applications were sent to 4 firms up to February 1, 2004.
[27] As of 10/23/07 there were 303 firms on the LendingTree Network. As of 9/16/09, there were 216 firms. Some firms operate nationally, while others operate in selected states.

tomer service performance, and the lenders' past success with the requested loan program and credit score.

A competing firm is able to provide up to 3 distinct quotes to each customer referred by LendingTree. The three quotes can be differentiated by loan type (fixed, adjustable rate mortgage, or home equity line), duration, and/or closing costs. If the applicant meets the credit score requirements for all 5 network firms, the applicant will receive a maximum of 15 offers. However, the firms choosing to offer a quote cannot observe how many other firms are competing for the client or, in theory which firms they are competing against. However, using state licensing information and the LendingTree list of participants, it is possible to approximate which firms can compete in each market. If the applicant's credit score, or the loan requested do not meet the guidelines for available products, firms can choose not to provide a quote. It is possible that the applicant receives no offers when applying, but that is rare; it is more likely that firms will provide the customer with a quote on an alternative loan product.

Firms incur a flat fee for the opportunity to compete for the customer's business. Additionally, the firm that provides the "winning" bid on a loan and completes the transaction is required to pay a computer loan origination fee (CLO) to LendingTree from the proceeds of the loan. The fee is a function of the type of loan, loan amount, and customer credit score; at times this fee can exceed $1000. LendingTree is able to extract profits from the network of firms and thus there is the presumption that firms experience above-normal profits by joining the network. This is a contradiction to the outcomes implied in

search models, which predict lower profits in online markets due to the increase in available information to the consumer.

As part of the mortgage firm's management decision, all online loans were subject to the same fee structure regardless of which ICS site provided the application.[28] Therefore, loan originators are responsible for paying the fee to the ICS; consequently they know in advance the cost associated with online customers and are able to price loans and their commissions accordingly. Customers search online to find the lowest price and do not explicitly observe any cost of applying through the ICS site.

Loan originators are mortgage broker/banker employees who provide the service of matching the customer with the mortgage. At the majority of broker firms, loan officers are commission-based employees and they earn income from a loan in two ways. The loan originators can charge the customer a loan origination fee for the service provided or offer the customer an interest rate higher than the wholesale price in order to receive a yield spread premium (YSP). More commonly, loan originators receive income through both sources.

4.3.3 *The Retail Consumer*

Any application to the firm that is not received through an ICS site is considered a retail transaction: business is obtained through local advertising, word of mouth referral, and any direct call or email from the consumer. Retail consumers' search is a non-sequential process, where each consumer observes one price quote per unit of search.

[28] Since the majority of loans where obtained from the LendingTree network and since the LendingTree fees are highest amongst all ICS sites, the firm used LendingTree pricing for all its loans.

Additionally, the firm does not have any information about how many quotes the consumer has received prior to current application.

The only information available to the firm are the retail consumer's characteristics and the loan program requested. The firm is uncertain about the consumer's search behavior or what competitors it is in direct competition with. The firm's objective is to quote a price that is marginally lower than the consumer's reservation price. Since the reservation price is a function of the consumer's information about the price distribution in the economy, the firm is uncertain about the consumer's reservation price.[29] The firm assumes that retail consumers have a higher search cost than online consumers, since it is necessary for them to undertake additional search to gain additional information. From search theory perspective, retail consumers should receive higher prices than online consumers because of their higher search cost. However, since the firm is uncertain about the retail consumer's information set, it is possible that the firm would price its product at a lower price than search theory would predict.

4.4 Empirical Model and Specification

Theory has established that better-informed consumers should receive lower prices than those who are less informed. To test whether online customers receive a different price than offline customers, it is possible to estimate mortgage prices controlling for all default and prepayment risk correlates and including an online or retail dummy variable.

[29] Further research will examine differences in the firm's success in converting applications to closed loans in both markets. If firms are able to convert more applications in online market, this is evidence that firm is able to better predict the consumers ' reservation price

Table 3.1 includes some of the default and prepayment variables expected to affect the pricing of a mortgage:

$$\boldsymbol{i = \beta_0 + \beta_1 X_{1i} + \beta_2 D + \varepsilon} \tag{4.1}$$

The interest rate paid on a mortgage is a function of the probability of default or prepayment of the loan. Thus, the exogenous variables would include all variables that increase or decrease the probability of default or prepayment of the loan, such as credit score, the loan-to-property-value ratio, whether the loan is a purchase or refinance, and the consumer's net worth. A refinance loan, more specifically, can be considered either a cash-out loan or a rate/term loan. If the proceeds of the loan are used to pay off debt other than the mortgage, or if the borrower is liquidating the equity of the house, the loan is considered a cash-out loan. Since the customer is extracting equity from the house, a cash-out loan will have a higher default likelihood compared with a refinance loan that is acquired to adjust the loan term or loan rate only. Therefore, the borrower receives a higher interest rate on cash-out loans because of the higher probability of default. Pre-payment risk is associated with the likelihood of the consumer paying the loan in full before the maturity date. Consumers who face a higher liquidity constraint, specifically those with a higher net worth or liquid assets, are in the position to pay the mortgage before its maturity. If online consumers are more informed, they will receive lower rates than offline consumers and consequently $\beta_2 > 0$. However, the above specification suffers from several econometric issues.

The dependent variable is the price of the mortgage. Mortgages can be priced using a two-part pricing method. The first price is the interest rate, which is usually quoted in newspapers and has been used as evidence of dispersion in prices.[30]

The other price is the fee or points paid to the broker. A borrower can gain a lower interest rate if they are willing to pay higher origination fees. Any specification that does not control for the endogeneity between interest rate and points will be biased. A more-accurate dependent variable would be the total price paid by the consumer: interest rate plus points or total commission paid to the broker. To control for this endogeneity, I constructed an APR measure which takes into account the note rate plus the fees per dollar of loan.[31] The APR variable represents the revenue (in percent of loan amount) from each loan.

Each customer receives an APR, where APR is a function of the mortgage interest rate i at time t and upfront costs U.

$$\boldsymbol{APR_{it} = f(I_i(R_t, f_{it}, D(X_i), P(X_i)), U_i(X_i))} \qquad (4.2)$$

The consumer's interest rate I_i is a function of the prevailing market interest rate R at time t and the margin at which firm j receives its funds f_{jt} at time t. Additionally, the interest rate includes a default premium $D(X_i)$ and prepayment risk $P(X_i)$, which represents the additional interest rate consumer i pays for their default and prepayment risk.

[30] Baye, M. R. and J. Morgan (2002). "Information gatekeepers and price discrimination on the internet." Economics Letters **76**(1): 47-51.

[31] APR is the effective interest rate after controlling for the fees paid by the consumer. See Appendix for method of calculating the APR.

In constructing the dependent variable, I assume that the firm's cost f_j is constant through time; and, since the analysis focuses only on one firm, it is possible to place $f_{jt} = K$.[32] However, in a multi-firm analysis it is necessary to add firm dummy variables, as to control for the cost of capital and any other firm-level variations in the pricing of default and prepayment risk. The variable R_t can be replaced by the 10-year Treasury rate at time t. Some studies have included year dummy variables to control for macroeconomic conditions, although a daily measure of macroeconomic conditions is more suitable. The time at which all variables are calculated is the date when the loan is closed. The actual date the rate is assigned (lock date) is not included in the firm's data, since loan processing time may take 30 to 60 days[33]. Using the market interest rate on the day the loan is closed may introduce measurement error. Crawford and Rosenblatt (1999) found that there are differences in when consumers decide to "lock" the rate. There may be benefits to keep the rate "floating" during the loan processing if the consumer believes that market interest rates will fall in the near future. The dependent variable can now be expressed as measuring the default and prepayment risk premium of person i.

$$\boldsymbol{Default\ Risk_i + Prepayment\ Risk_i = APR_{it} - R_t = K_0 + \beta_1 X_i + \varepsilon_i} \quad (4.3)$$

$$\boldsymbol{Price_i = \beta_0 + \beta_1 X_i + \beta_2 D + \varepsilon_i} \quad (4.4)$$

Estimating equation (3.4) using ordinary least squares (OLS) will provide estimates of the coefficient for each credit risk variable. If a dummy variable D is included, it will estimate the pricing differential for online customers relative to offline customers,

[32] This is a strict assumption, and the focus of chapter 4 examines variations in f_{jt}. However for this analysis I use year fixed effects to control for variations in the cost of capital.
[33] Siemens and Benjamin (1990).

which is not associated with consumers' default or prepayment risk. More specifically, the coefficient of the dummy variable should capture any pricing variations associated with consumer information. If additional information does not provide the customer with any benefits, then the coefficient would be equal to zero. However, to ensure that omitted variable bias is not introduced, all risk variables that are considered by mortgage underwriters must be included.

Because of the extensive list of variables examined by underwriters when pricing mortgage products, empirical studies are commonly plagued with omitted variable bias. Data sets provided by firms rarely include the entire list of risk variables, and therefore researchers are often attempting to estimate the size and direction of the bias associated with limited data (Dietrich 2005). Regulatory data sets that contain national information on mortgage denials, such as the Home Mortgage Disclosure Act (HMDA), usually exclude at least one risk variable. The HMDA data are used extensively to examine racial differences in mortgage denial rates, yet the data set does not include the applicant's credit score, a major determinant of default risk. The repercussion in this research is that if an underwriting variable is omitted from the data set and if it is correlated with the consumer's market type, then it would cause the error term to be correlated with the market dummy variable and thus the coefficient on the dummy variable will be biased.

The data set utilized in this empirical study does not include several variables that are usually associated with prepayment and default risk underwriting. Specifically, consumer income, wealth, and debt information are unobserved. However, that information was available to the loan underwriters. If online consumers are wealthier than retail con-

sumers, the online dummy variable will absorb the effect wealth has on mortgage pricing and lead to a biased online coefficient. To eliminate the omitted variable bias and to construct a more homogenous set of loans, I utilize Federal National Mortgage Association (FNMA) and Federal Home Loan Mortgage Corporation (FHLMC) guidelines and limit the sample to "Agency" loans. The reduced sample includes loans with identical default and prepayment risk characteristics as priced by FNMA and FHLMC. Loans with identical default and prepayment risk will have an identical qualifying interest rate. If omitted variabales are correctly controlled for then any variation in the price paid by consumers is due to their ability to reduce their markup (overage). However, if one of the omitted variable is also correlated with consumer bargaining power and market choice then the estimation will still remain biased.

Assuming omitted variable bias is controlled for, then in this specification the coefficient of risk characteristics β_1 should be equal to zero. A positive (negative) coefficient is interpreted as an increase (decrease) in the consumer's bargaining power. Including an online dummy variable will capture any reduction in markups or overages associated with an increase in the availability of information to the consumer and the firm's market power. The default and prepayment risk price can be considered the marginal cost of providing the loan; any price above can be regarded as profits to the firm.

By comparing the variation in the firm's markup strategies, I can observe if interest rates vary depending on the clientele in each market. Specifically I will test whether firms are able to price discriminate according to the search costs incurred by the consumer. If online markets reduce the search costs and cause a reduction in information asym-

metries, customers should be able to negotiate lower prices. Therefore, I can observe if the mortgage firm considers online markets as more competitive and test whether the online market prices approach the Bertrand equilibrium, where prices and profits are lower relative to retail markets. Alternatively, if LendingTree's policies facilitate tacit collusion, then consumers will not be able to be aggressive negotiators.

Since each customer choice between online and retail firms is a function of their search costs, and not randomly assigned, the OLS estimation of the model would not yield consistent estimates of the coefficients. To correct for the endogenous market choice, I employ a switching model. This enables me to identify the interest rate for each customer if they chose the alternative market to conduct their mortgage transaction.

Consumers maximize their expected consumer surplus given their search cost and the underlying distribution of total mortgage prices (denoted by APR)

$$\boldsymbol{ECS(C) = \int_{b}^{c} CS(APR)dF(APR)} \tag{4.5}$$

Consequently, consumers will search online if the gains from online shopping exceed their reservation ρ_i.

$$\boldsymbol{Price_{if} - Price_{io} > \rho_i} \tag{4.6}$$

where $Price_{io}$ denotes the rate for individual i in online markets and $Price_{if}$ denotes the rate for the individual in offline or retail markets. Thus, the consumer shops and purchases online if the expected gains of online commerce are greater than their reservation rate ρ_i.

The reservation rate ρ_i is a function of the individual characteristics X_i and search costs C_i.

$$\boldsymbol{\rho_i = \alpha X_i + B C_i + \epsilon_{i1}} \tag{4.7}$$

The individual characteristics in the empirical test will include observable loan variables. Search costs are unobservable; therefore for identification purposes I need to include exclusion restriction. As an exclusion restriction, I include the distance of the property from the firm's office.[34] Consumers further away from the home office incur a higher search cost of visiting the firm and would therefore be more likely to apply online. Since the distance should not affect the price a consumer pays for a mortgage, it is not included in the price equation. Additionally, I use the Census data on average income by zip code as an instrument to proxy for the individual's income. It is assumed that individuals with higher income would have higher opportunity cost of searching and therefore would be more likely to shop online to reduce their search costs. Another instrument used is Census data on average education by Zip Code. More-educated consumers may be more likely to be familiar with the benefits of the internet and will shop online. Thus we can estimate:

$$\boldsymbol{C_i = \gamma_1 + \gamma_2 X_i + \gamma_3 Z_i + \epsilon_{2i}} \tag{4.8}$$

$$\boldsymbol{Price_{if} - Price_{io} > \alpha X_i + \beta(\gamma_1 + \gamma_2 X_i + \gamma_3 Z_i + \epsilon_{2i}) + \epsilon_{1i}} \tag{4.9}$$

$$\boldsymbol{Price_{if} - Price_{io} > (\alpha + \beta\gamma_2) X_i + \beta\gamma_1 + \beta\gamma_3 Z_i + \beta\varepsilon_{2i} + \epsilon_{1i}} \tag{4.10}$$

[34] Graph 3 provides state difference between online and retail loans.

Thus we can rewrite eq (4.8) as a probit model, where the consumer shops online if $\tau_i^* > 0$ and retail otherwise. Following Lee (1978) , where

$$\boldsymbol{\tau_i^* = \gamma_1 + \gamma_2 X_i + \gamma_3 Z_i + \epsilon_i} \tag{4.11}$$

$$\boldsymbol{E[Price_{io}|O] = \theta_{o1} + X_i \theta_{o2} + \sigma_{1\varepsilon}\left(\frac{f(-\varphi_i)}{1-F(-\varphi))}\right)} \tag{4.12}$$

$$\boldsymbol{E\left[Price_{if}\middle|f\right] = \theta_{f1} + X_i \theta_{f2} - \sigma_{2\varepsilon}\left(\frac{f(-\varphi_i)}{F(-\varphi))}\right)} \tag{4.13}$$

$$\boldsymbol{\tau_i = \delta_0 + \delta_1(price_{if} - Price_{io}) + \delta_2 X_i + \gamma_3 Z_i + \epsilon_i} \tag{4.14}$$

errors are distributed trivariate normal with covariance matrix

$$\Omega = \begin{bmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon o} & \sigma_{\varepsilon f} \\ \sigma_{\varepsilon o} & \sigma_o^2 & . \\ \sigma_{\varepsilon f} & . & \sigma_f^2 \end{bmatrix}$$

The switching model will allow an examination of the joint determination of the propensity to conduct online search (equation 4.11) and the effects of lower search costs on the price of mortgage for each consumer type (equation 4.12 and 4.13). The inclination to reduce search costs will depend on the gains to the consumer from online commerce, which is the latent variable τ_i (equation 4.14). Since I do not observe the price each consumer pays in both markets and observe only the price a consumer pays after their selection into the market type, I must estimate τ_i^* and control for the selection on unobservables in the price equations.

Using the inverse Mills ratio, I estimate the price equations conditional on selection into each market. Therefore equations 4.12 and 4.13 will allow the estimation of the

price conditional on the consumer selecting the online market and retail market, respectively. By assumption $\sigma_\varepsilon = 1$, therefore the covariances $\sigma_{\varepsilon o}$ and $\sigma_{\varepsilon f}$ can be rewritten as $\sigma_o \rho_{0\varepsilon}$ and $\sigma_f \rho_{f\varepsilon}$, where $\rho_{o\varepsilon}$ and $\rho_{f\varepsilon}$ are the correlation coefficients. If the error terms in the price equations are uncorrelated with the error in the propensity equation, the estimation is reduced to a pooled OLS. A switching model is efficient and provides consistent standard errors by employing the full-information maximum likelihood method (FIML) to simultaneously estimate the propensity to conduct online commerce and the price equations for each market type. Although The FIML is efficient among all simultaneous equation models, it is computationally burdensome and may provide difficulties in convergence. A Limited-information maximum likelihood (LIML) approach is usually preferred for its simplicity[35] and robustness to non-normality. An alternative method is to use the Heckman two-stage least squares (2SLS), which utilizes a LIML estimation process. While robust, the LIML is not as efficient as the FIML.

4.5 Data

The mortgage firm made available its customer database, which includes all loan applications submitted through ICS sites and retail locations. The data span five years from 2002 to 2007, where the firm received 41,054 loan applications.

This study will examine the differences in the business obtained by the firm, more specifically the difference in price paid by online customers and retail customers. Thus, I limit the sample to include only completed transactions (loans closed) and restrict the

[35] Greene, W. (2003). Econometric Analysis, Prentice Hall.

sample period to 2002 to 2006. The data for 2007[36] are interesting in their own right; however, management's decision to cease operations was made public in early 2007. To eliminate any bias introduced by this information these observations are excluded from the sample. During the time period 2002-2006, the firm's production totaled 7,977 loans.[37] Furthermore, because of the variation in mortgage pricing and omitted variable bias and the desire for a more homogenous sample, I limit the data to loans classified as "Agency" loans by FNMA or FHLMC.[38] Agency loans are expected to have an identical default and prepayment risk. The two programs do not impose any interest rate adjustment to the note rate as long as the customer is approved. Thus, an omitted variable such as income may disqualify a customer from the program; however, if a customer is approved, income does not affect the rate obtained.

Reducing the sample is assumed to eliminate any omitted variable bias that may exist from not having customer income data. Although income does not affect the interest rate the consumer qualifies for, it may be correlated with the consumer's bargaining power. In the case that income is correlated with the consumer's bargaining power and market choice, omitted variable bias will still exist in this specification. Specifically, if low income (or high debt-to-income) consumers are more likely to shop online and if bargaining power increases with income, then the coefficients of the risk characteristics would be biased upwards in the online market. For the purpose of this study, it is assumed that if any correlation exists, it is equal in both markets. Therefore the bias would be equal

[36] The firm's management announced on several occasions in 2007 that operations might end. Finally the company closed on August 31, 2007.
[37] I refer to this sample as the large sample in my discussion.
[38] I refer to this sample as the reduced sample in my discussion

across markets and will not jeopardize the analysis. The assumption is stringent and remains a weakness in this study.

There are 3,368 completed transactions in the reduced sample. Summary statistics for both samples are included in Table 4.2. Column 3 details the summary statistics for the large sample, and column 6 details the averages and variances of the reduced sample. Columns 1 and 2 compare the online and offline averages for the large sample, and columns 4 and 5 are for comparison of online and offline customers in the reduced sample.

The data consist of loan level observations. The individual variables included are: a loan type variable, and a variable indicating the consumer's market choice. The only individual-level variable included in the dataset is the consumers' middle credit score. There are three companies that report consumers' credit scores: Experian, Equifax, and Transunion. The "middle" or median credit score is used for mortgage pricing and risk underwriting. The average credit score for all completed transactions (Column 3) is 665,[39] and the average credit score for the reduced sample was 10 points higher.

Loan-level variables include information on the type of loan- whether the loans are fixed rate or adjustable rate mortgages (ARM). The ratio of loan-to-home value (LTV) represents the amount of equity the consumer has accumulated in the property. The higher the loan-to-value, the lower the equity owned and the higher the presumed risk of default. The average observation in the reduced sample had a loan amount of $139,870 and a loan-to-value of 74%.

[39] The national average credit score is 680. http://www.creditreport.com/info/credit-scores/average-credit-scores.asp

Included are dummy variables that specify whether the loan is a cash-out loan, purchase loan, or second-lien loan and a variable indicating the loan program. The variable "rate/term" is constructed using the dummy variables. If a loan is not a cash-out, purchase, or second-lien then it is a rate/term refinance. Cash-out loans represent 54% of the loans originated during the sample period. Second-lien loans are classified as such because they are loans that place the mortgagor in second position to receive any funds from foreclosure or sale of the house. Rate/term refinancing is when customers refinance to take advantage of more favorable terms, such as adjusting the duration of the loan, or more likely to take advantage of lower interest rates in the market. Although all loans in the sample have identical risk premiums, consumer credit worthiness and loan types do vary.

Using the loan program, I construct a dummy variable that takes on the value of 1 for fixed loans and 0 for ARMs or home equity line of credit (HELOC). Fixed-rate loans account for 92%; this is a drastic variation from the larger sample where fixed loans accounted for 62% of all loans. ARMs are usually associated with subprime loans, while government-insured loans are considered prime loans (Carrillo 2008). It is expected that ARM loans would be underrepresented in the reduced sample.

The market-choice variable "online" is equal to 1 if the consumer applied through an ICS site and 0 otherwise.[40] In the sampling period ICS sites accounted for 70% of the firm's business, and LendingTree accounted for 58% of the total business.

[40] Each ICS has different practices about how many firms compete for each customer's business. A more detailed examination of specific ICS practices and its impact on the average price and variances can further contribute to the search literature

More information does reduce the dispersion of prices as predicted by search theory. The variance of Price in online markets is 1.41 compared with 1.52 for the retail consumers. Figure 4.2 shows the price dispersion for both markets using kernel densities.[41] The mean and dispersion is smaller for online consumers. Table 4.3 provides summary statistics for online commerce by year and shows that consumers purchasing mortgages online paid less on average compared with offline consumers (Table 4.4). In contradiction to Bailey's (1998) hypothesis that the dispersion of prices would decrease with time and as more consumers used the internet, the variance of prices in both markets increased each year. Nonetheless, the variance of online prices is consistently lower than the variance of retail prices.

The reduced sample is representative of the large sample in most categories. The proportion of online and retail transactions remains unchanged. The reduced sample eliminates any heterogeneity due to subprime underwriting. Loans such as no income verification, stated income, and no income/no asset loans are not included because of the unobserved pricing guidelines. However, by reducing the sample I am unable to analyze the search behavior and the pricing the non-conforming consumers pay.

4.6 Results

4.6.1 *Price to the Consumer*

The switching regression simultaneously estimates the probability of shopping online and the price that consumers pay for a mortgage. The estimation method is contingent on joint normality of the error terms in the Probit and price equations. Given the hy-

[41] This is a subsample of the data which excludes second-lien loans and ARM products and includes only 30-year loans. Figure 4.1 represents the reduced sample

pothesis developed by search theory, consumers' search process is endogenous to mortgage prices observed. Search costs are unobserved but affect the consumer's choice of market, and not controlling for this selection bias will affect the inference. Therefore, including the inverse Mills ratio in the price equations will control for the selection process if it exists.

$$\boldsymbol{E[Price_{io}|O] = \theta_{o1} + X_i\theta_{o2} + \sigma_{1\varepsilon}(\frac{f(-\varphi_i)}{1-F(-\varphi))})} \quad (4.15)$$

$$\boldsymbol{E\left[Price_{if}\middle|f\right] = \theta_{f1} + X_i\theta_{f2} - \sigma_{2\varepsilon}(\frac{f(-\varphi_i)}{F(-\varphi))})} \quad (4.16)$$

$$\boldsymbol{\tau_i = \delta_0 + \delta_1(price_{if} - Price_{io}) + \delta_2 X_i + \gamma_3 Z_i + \epsilon_i} \quad (4.17)$$

Results from the switching model are reported in Table 4.5. The exclusion variable Distance is positive and significant in the probability equation (reported in column 3). The distance of the property from the retail office is not expected to affect the price the consumer pays for a mortgage; however, distance is likely to affect consumers' method of search. The predictive ability of distance supports its use as an exclusion restriction and will assist in identification. The variable measuring consumer education is negative, implying that individuals with higher education are less likely to use LendingTree or another ICS. The result is opposite of what is expected, where the hypothesis is that higher education levels are correlated with financial acumen and internet usage. Education is statistically insignificant and therefore it does not affect the probability of online com-

merce. Since the majority of homeowners in today's economy have a computer and access to the internet, internet usage and education are less likely to be correlated.[42]

All variables observed by loan officers during the application process are insignificant in the probit estimation. Credit score, loan amount, mortgage type (fixed or ARM), second-lien, and loan duration dummy variables are insignificant. Thus consumer and loan characteristics do not predict consumers' search process. The variable Purchase is negative and significant. Thus we should expect customers to shop in the retail markets when purchasing a new home. Historically, most purchase transactions involve real estate agents; this finding reflects steering by the real estate agent towards local brokers and lenders.

The results of the estimation indicate there is a variation in the propensity to shop online throughout the years of the sample. Compared with the reference year in 2003, dummy-year effects are all positive and significant at the 5% level. In 2004, online commerce was more likely than in 2003; nonetheless, the magnitude of the coefficient decreases with time, possibly due to changes in taste. Alternatively, it may suggest that as demand for mortgages decreases from its height in 2004, online demand decreased more rapidly than retail demand.

The coefficients "rho1" and "rho2" report the correlation between the error in the market choice equation and the price equation errors. Both coefficients are statistically insignificant, which implies that consumers shopping in either market do not self select

[42] The switching model fails to converge when including other exclusion restrictions such as distance squared and the average zip code income, which is a proxy for consumers' income. Furthermore, I include Census information on consumer online shopping and computer ownership as exclusion restrictions, but the results are unchanged.

into the market type. Additionally, it indicates that the estimation method can be reduced to a grouped OLS estimation. The insignificance of the correlation coefficients implies that the propensity to shop online and the price equations are in fact not endogenous.

Table 4.6 provides the results for the grouped OLS estimation. Since endogeneity does not play a role, I can compare the coefficient estimates across groups. The predictions of search theory suggest that consumers in the online market pay lower prices for each characteristic compared with the retail consumer. When testing if the coefficients are equal across price equations, I fail to reject that they are equal for 12 of the variables, including the constant term. If variables are equal across equations it implies that online and retail consumers pay equal prices for the loan features. Consequently consumer access to more information does not create a more competitive environment as theory would suggest. For loan duration variables, purchase loans, and the year 2006, there is a statistical difference in what online consumers paid compared with retail consumers.

In the online market, the firm is able to price discriminate with respect to the duration of the loan. Loan-duration dummy variables are significant and increasing in loan maturity. Alternatively in the retail markets, loan duration does not impact the price the consumer pays. Relative to the reference category, 10-year loans, online consumers pay a premium for each additional 5 years up to 25 years. There is a 97-basis-point discount for consumers who obtain a 30-year mortgage compared with 25-year mortgages. Because of the increase in market power, online firms are able to discriminate with respect to loan types. The market power is eliminated in retail markets, where information asymmetries reduce the firm's market power. In retail markets, the firm is uncertain about the consum-

er search behavior and information set. The lack of information reduces the firm's market power and allows increases the consumers bargaining power.

Consumers in search of a mortgage loan with the purpose of purchasing a home will pay 45 basis points more if they obtain their mortgage through an ICS site compared with what a retail consumer will pay. Most purchase transactions involve real estate agents, who usually have established relationships with mortgage brokers or lenders in their area. In an attempt to attract consumers of purchase loans, the firm made efforts to solicit and recruit local realtors. Therefore, search costs are lower in retail markets when a real estate assists the consumer. Moreover, the repeated interaction between realtor and lender is expected to reduce the prices realtor-assisted consumers pay. This is also reflected in the estimation of the propensity to shop online, where purchase consumers were less likely to search online. Therefore, if a consumer applies through an ICS site, they are signaling to the firms that neither they nor their realtor has a relationship with a local lender. Additionally, unlike a refinance loan, purchase loans have a specific closing date, and thus the consumer's demand is less elastic. Firms are able to exert their market power on purchase consumers applying in online markets. The signal of high search costs, in addition to inelastic demand, allows the firm to price purchase loans higher in online markets than in retail markets.

With the exception of the dummy variable for 2006, all other year dummy variables are not statistically different in both equations. Profits in 2006 decreased 47 basis points in online markets and 73 basis points in retail markets when compared with 2003. The reduction in profits per loan may be capturing the decrease in demand for mortgages

as housing prices began to decline. An alternative hypothesis is the assumption that the firm acquires its capital at a constant K may not be accurate. The firm may have been able to reduce its cost of capital by becoming a mortgage banker and consequently transferring the lower costs to the consumer.[43] Nonetheless the retail consumers paid lower prices in 2006 compared with online consumers, when controlling for all other variables. When comparing the coefficients for the year 2004, I cannot reject that in both the retail and online markets prices remained similar to 2003. Consequently, the change in LendingTree policy did not affect the firms' pricing strategies. Increasing the number of firms competing for each consumer from four to five did not have an effect on prices.

It is possible to conclude that the number of competing firms and the bidding environment created by LendingTree and other ICS firms has not reduced the average price of mortgages. The ability of consumers to observe multiple offers does not create more of a competitive environment as compared with markets where the consumer observes a single price per search. On average, prices are equal in both markets. The benefits of the reduction in search costs are apparently offset by the anticompetitive environment the gatekeeper provides. For some loan programs and types, the anticompetitive effect is larger and online consumers pay higher prices than the retail consumer. The results are unchanged when the sample is restricted to an even more homogenous product.[44]

An alternative explanation for the higher prices in online markets can be attributed to the gatekeeper fees. The fees charged to the firm by the gatekeeper increase the

[43] A mortgage lender uses its own funds to supply loans, whereas a broker must use other firms' funds. By becoming a lender, the firm has access to a lower cost of capital.
[44] I exclude second-lien loans and ARM products to generate a homogenous sample of mortgages. When the sample includes only fixed loans that are in the first-lien position, the results are unchanged.

cost of online loans. Consequently, even if prices are equal to marginal cost of production in the online market, the higher marginal cost would cause online prices to be higher relative to retail markets. Therefore testing for differences in profits would assist in either supporting or eliminating marginal cost variations as the reason for higher online prices. If online profits are larger than retail profits, then the hypothesis of the existence of an anti-competitive effect would be supported.

As suggested by Brown and Goolsbee (2002) , the possibility exists that the inception of ICS has caused a reduction in all market prices. Consequently, informed consumers and ignorant consumers pay lower prices after the internet compared with before the internet. Since the mortgage data starts in 2002, well after the establishment of the internet, it is not possible to test this hypothesis. However, this implies that all coefficients should be statistically equal, which is not evident in the results.

4.6.2 *Robustness Check*

The switching model uses the FIML, which is contingent on the normality of the error terms. Alternatively, the Heckman 2SLS uses a LIML in the estimation process; however, the tradeoff is in the form of efficiency. The estimation procedure is less onerous compared with the FIML and more likely to meet the convergence criteria. The FIML specification estimated requires at least one exclusion restriction for identification. To test whether the estimation is sensitive to the exclusion restriction, I include distance squared and zip code income as other instruments. However, the FIML does not converge when the other exclusion restrictions are included. Thus, I test whether the switching

model is sensitive to the exclusion restrictions and normality of the error term by comparing the Heckman 2SLS with the switching results.

Table 4.7 provides the estimation results for online and offline pricing. Column 1 provides the results for the pricing equation when all exclusion restrictions are included. Column 2 includes only the exclusion restrictions from the switching model. Comparing across the Heckman 2SLS specification, the coefficients are similar and maintain their significance; therefore, the specification is robust with regard to the exclusion restrictions and the specification does not suffer from weak instrumental variable bias. Comparing the results from the Heckman 2SLS and the switching regression with identical exclusion restriction, the results are unchanged. Consequently, the results derived from the switching model are not sensitive to the assumption of normality of the errors or the choice of exclusion restrictions. Therefore, I can eliminate the possibility of selection and endogeneity between market choice and mortgage prices.

Blinder-Oaxaca (1973) decomposition is often used in labor economics to study wage differentials by group. The analysis separates the difference in wages into explained and unexplained portions. Other fields of economics have also used the decomposition to explain differences between groups. In this research it is possible to use the method to separate the difference in interest rate into explained and unexplained parts.

Using the results from the pooled OLS estimation, I conduct the Oaxaca decomposition proposed by Neumark (1988). The decomposition in equation 4.18 measures the difference in prices between the retail and online markets. It assumes that there is a stan-

dard price β that should prevail and allows prices in both the retail and online market to deviate from the "true" price.

$$\boldsymbol{Price_f - Price_o = \beta(X_f - X_o) - (\beta^f - \beta)X_f + (\beta^o - \beta)X_o} \qquad (4.18)$$

The first term in the decomposition is interpreted as the difference in price due to differences in characteristics. The next two terms represent the difference in price due to preferential treatment. If the "true" price is assumed to exist in the retail market, that is $\beta^f = \beta$, then the Neumark decomposition simplifies to the Oaxaca decomposition. Thus, the Oaxaca decomposition is a specific case which can be derived from the general format represented in equation 4.18 (Neumark, 1988).

$$0 = 0.00 + 0.5 - 0.5$$

The results indicate that on average online consumers and retail consumers pay the same price for a mortgage. The result of the decomposition suggests there are no differences between group observables and therefore characteristics do not affect the price. The second and third terms indicate that the differences in online and retail markets are solely due to differences in the coefficients. In fact, the average retail consumer would increase their rate by 50 basis points if they were to shop online. Alternatively, a consumer shopping online would reduce their price by 50 basis points if they shopped in retail markets. The consumer characteristics are similar between groups and the variation in prices is due to the market structure. Therefore, market structure in the online market provides the firm with ability to price discriminate in terms of consumer characteristics.

4.6.3 *Profits to the Firm*

The analysis thus far has established that online consumers pay a higher price compared to the retail consumer. However, online consumers may be receiving higher prices due to the fee which is charged by the gatekeeper. To fully examine the profits in online markets I must account for the variation in marginal cost and the gatekeeper fee. Thus, I create a new APR measure that excludes the gatekeeper fee for online loans, where the new dependent variable will measure the profits the firm receives.[45] For retail operations the revenue the firm receives is the profit to the firm because the firm does not have to pay any additional cost. Therefore in retail operations, the firm retains all revenue, while online operations have a higher cost of originating the loan and it must be accounted for.

To estimate the profit equation I use a grouped OLS method, and compare the coefficients across markets. If the coefficients are equal across markets it indicates that the consumer receives a higher price due to the higher marginal cost imposed by the gatekeeper and profits are equal in both markets. Alternatively, if the coefficients are higher (lower), then that suggests that firms make higher (lower) profits in the presence of a gatekeeper. If the profits are higher in online markets, the presence of the gatekeeper increases the anticompetitive effect. Alternatively, if the firm receives lower profits in online markets compared to the retail market that suggests that the online market is more competitive.

[45] See Appendix for method of calculating the APR

The results of the profit equations are reported in Table 4.8. After controlling for the variation in marginal costs, I find that the presence of the gatekeeper creates both a search cost effect and an anticompetitive effect depending on the market segment. On average the profits to the firm are equal across market type, and thus the anticompetitive effect and search cost effect are equal and offset each other. For fixed loans and loans originated in 2004, the search cost effect dominates the anti-competitive effect. The firm received lower profits in online transactions with LendingTree compared to retail transactions.

For loans originated in 2006, Purchase loans, and Loan Duration variables, the Firm received higher profits in the online market compared to the retail market. Consequently, for those market segments, the anti-competitive effect dominates the search cost effect. The presence of the gatekeeper and information presents firms the opportunity to tacitly collude and increases their market power. Alternatively, the results suggest that the gatekeeper ‘s pricing strategy can be adjusted to extract more profits from the firms. The gatekeeper can increase its profits by increasing fees on loans originated in 2006, purchase loans, and for the loan duration variables.

4.7 Conclusion

In this research I have examined the effects of the internet and ICS such as LendingTree on market prices and firm profits. Economic theory predicts that ICS sites help lower search costs and will increase consumption of information by the consumer. As consumers become aware of the price distribution in the economy, they are able to choose the lowest-priced firm. Therefore, the introduction of online markets is expected

to cause a reduction in prices. However, I examine another effect of online markets, which I call the anticompetitive effect. Online markets with a gatekeeper that limits the amount of information the consumer receives provide firms with market power that is not observed in retail markets.

I test my hypothesis by examining firm profits and whether informed consumers (online consumers) pay lower prices for mortgage products than uninformed consumers (retail consumers). Mortgage pricing varies across lenders and consumer risk characteristics, and therefore I establish an approach to eliminate the heterogeneity in mortgage products. I restrict my analysis to a subset of mortgages with identical default and prepayment risk. After controlling for the macroeconomic environment and eliminating any bias due to omitted underwriting variables, I am able to examine the differences in prices paid between online and retail consumers. I use a switching regression to control for consumers' endogenous market choice, and I find that consumers searching for mortgages in online markets do not pay a lower price compared with consumers searching in a retail markets. Additionally, I do not find any supporting evidence that consumers select into market types as a result of unobserved characteristics.

I find that online markets provide the firm with the ability to price discriminate by loan characteristics. Contrary to the theory, some consumers may end up paying higher rates by applying for a mortgage in the online market. I conclude that this is evidence of tacit collusion behavior in online markets. Online markets make information accessible and provide the member firms with market power and the ability to monitor their competitors. Although the retail consumer has a higher search cost, they pay equal or lower pric-

es compared with the online consumer. Because of the uncertainty about the retail consumers' information, firms are forced to consider the possibility that the consumers may have already observed a low price. To the extent that this occurs, the prices in retail markets will be lower than prices in online markets. When consumers apply through an ICS site, the firm has information about its immediate competitors and the expected price distribution the consumer will observe; consequently, it can charge higher price for its product.

To examine firm profits, I control for the marginal cost variation across market types. The gatekeeper fees increase the marginal cost in online markets, and the variation in prices may be due to variation in the cost structure. When controlling for the higher marginal cost imposed by the gatekeeper, I find that for a segment of the market the firm receives higher profits in the online market. This provides further support that online markets create an anti-competitive market. For some market segments online profits are lower when compared to the retail profits, and in this case the presence of the gatekeeper increases the competitive environment. Nonetheless, on average, profits in online markets are equal to the profits in retail markets.

The threat of competition in retail markets is more severe than the actual competition in the online market. Empirical literature examining the effects of the internet on pricing should carefully consider the market structure, especially in the presence of a gatekeeper. The anticompetitive effect can be equal or larger than the search cost effect. This may help explain why empirical research focusing on the benefits of the internet has been inconclusive.

Calculating APR for the consumer (Price paid)

$$LA - C = P\left[\frac{(1+apr)^m - 1}{r(1+apr)^m}\right]$$

Where:

LA= Loan Amount at origination

C=Commission paid to the Loan Officer

P= Monthly payment

m= the total duration of the loan in months (30 year loan will have 360 monthly payments)

APR= the annual percentage rate and the variable to solve for. Solving for the APR provides the true price of the loan (in percentage points) to the consumer after controlling for the commission.

Calculating the APR for the firm (Profits received)

$$LA - (C - Fee) = P\left[\frac{(1+apr)^m - 1}{r(1+apr)^m}\right]$$

Where:

Fee= the fees paid to the gatekeeper. Solving for APR will provide the profits (in percentage points) the firm receives for each loan after paying the cost to the gatekeeper.

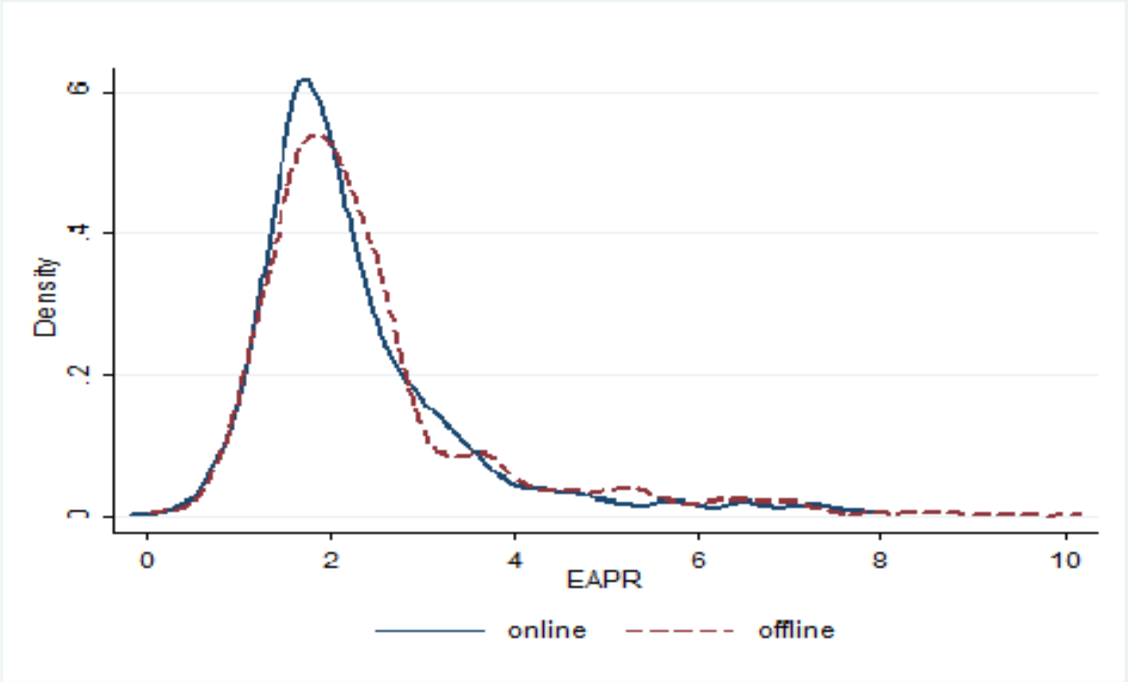

Figure 4.1: Price Paid by Online and Offline Consumers. 30 Year Fixed Mortgages Only.
Kernal densities of the price consumers pay for a mortgage. Source: Firm data.

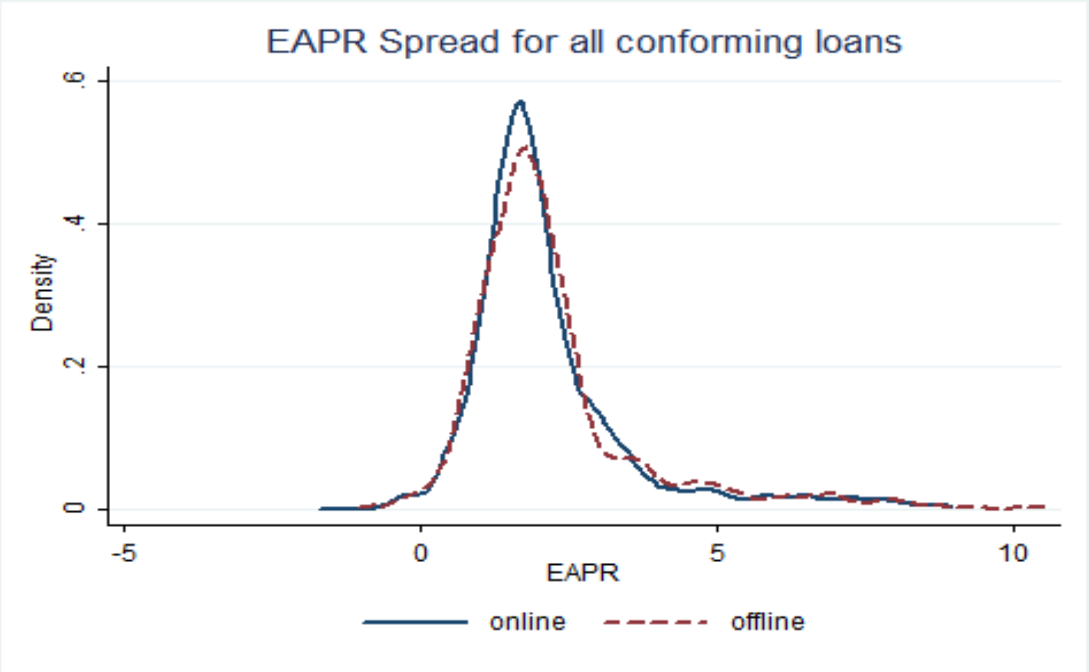

Figure 4.2: Price Paid by Online and Offline Consumers. All Loans
Source: Firm data

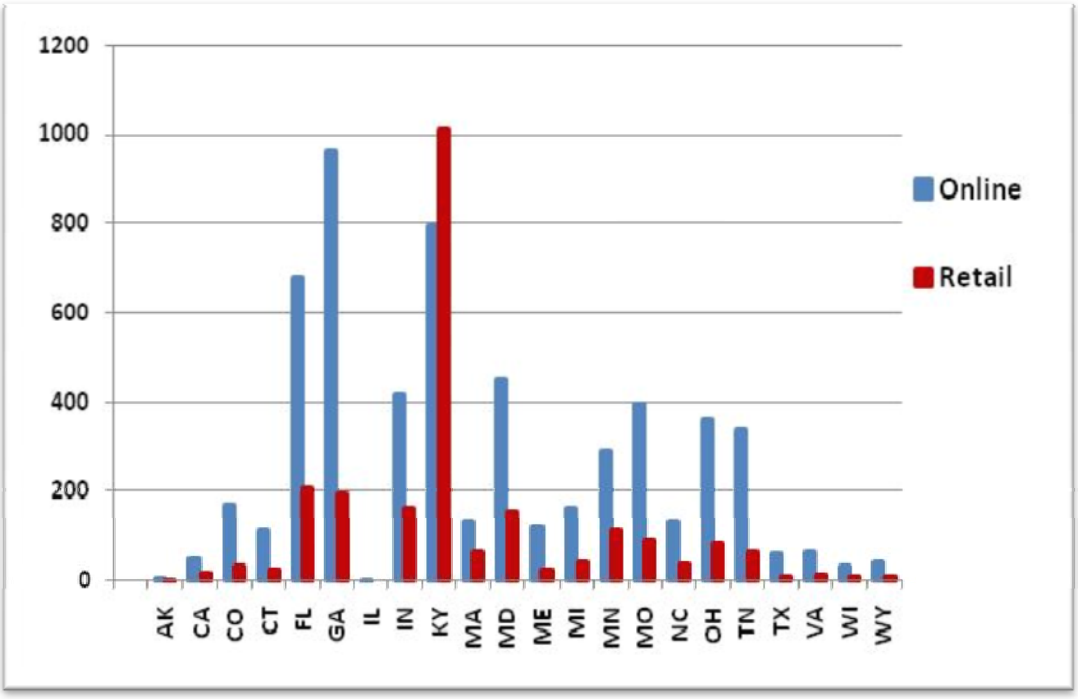

Figure 4.3: Number of Loans Originated by State and Market Type.
Source: Firm data

Table 4.1: Default and Prepayment Variables

Variable	Explanation
Expense to Income	Housing payment/Gross Income
Debt to Income	Total Debt Payments/Gross Income
Net Worth	Savings & stock Accounts, Retirement Accounts, Home Equity (Second Homes, Investment properties)
Employment History	Stability of Employment
Loan to Value	Loan Amount/House value (Equity invested in the house)
Fixed Loan	Fixed interest rate
Adjustable Rate Mortgage (ARM)	Loans with variable interest rate
Term of Loan	Loan Duration10,15,20, 25, 30,or 40 years
Loan Amount	Loan Amount
Credit Score	A measure of consumer credit worthiness
Lien Position	In the case of foreclosure or sale of the house, the lien position determines who receives funds first. It is important in the where sale or foreclosure proceeds are less than total mortgage debt outstanding.
Refinance Type	Rate and Term, or Cashout
Purchase	Loans made to finance a purchase of a home
Liquid Assets	Savings and stock accounts
Bankruptcy	A recent bankruptcy would impact the creditworthiness
Foreclosure	Foreclosure on a property would impact the creditworthiness of the consumer
Prepayment Penalty	Most states allow the lender to impose a prepayment penalty on mortgage for a specific time frame. Usually it ranges from 1-5 years and the penalty can range from 1-9% of the loan amount. They are designed to discourage early prepayment. Usually associated with subprime loans. Agency loans do not have pre-payment penalties.
APR	Annual percentage rate. Measures the cost of a loan including all transaction fees.
EAPR	The APR minus the 10 year Treasury. Measures the real risk of a loan.
Conforming	A loan that confirms to Fannie Mae and Freddie Mac guide-lines.

Table 4.2 : Summary Statistics for Large and Reduced Samples

	Large Sample			Reduced Sample		
Dependent Variable	(1) Online	(2) Retail	(3) Average	(4) Online	(5) Retail	(6) Average
Rate	6.98	6.84	6.94	6.33	6.31	6.32
	(1.67)	(1.73)	(1.69)	(1.36)	(1.48)	(1.40)
Commission	1375.79	1091.03	1293.03	1330.89	951.77	1216.28
	(1302.60)	(1446.87)	(1352.02)	(1157.16)	(1181.36)	(1177.11)
APR	7.17	7.01	7.13	6.51	6.46	6.49
	(1.72)	(1.80)	(1.75)	(1.40)	(1.53)	(1.44)
EAPR	2.77	2.70	2.75	2.14	2.21	2.16
	(1.72)	(1.75)	(1.73)	(1.41)	(1.52)	(1.44)
ICS Fee	600.47	0.00	426.47	601.27	0.00	420.32
	(286.98)	0.00	(364.28)	(282.46)	0.00	(363.11)
Credit score	663.16	672.79	665.97	672.62	681.53	675.34
	(62.39)	(67.50)	(64.06)	(58.51)	(63.66)	(60.24)
Loan To Value	76.16	72.38	75.06	74.88	71.89	73.96
	(24.92)	(27.65)	(25.80)	(24.30)	(25.97)	(24.86)
Loan amount	135.78	130.28	134.20	142.10	134.59	139.87
(1,000)	(91.50)	(92.59)	(91.85)	(80.17)	(74.88)	(78.71)
Fixed	0.63	0.62	0.62	0.94	0.92	0.93
	(0.48)	(0.49)	(0.48)	(0.25)	(0.27)	(0.26)
Conforming	0.42	0.44	0.42	1.00	1.00	1.00
	(0.49)	(0.50)	(0.49)	0.00	0.00	0.00
Cash Out	0.58	0.39	0.53	0.61	0.39	0.54
	(0.49)	(0.49)	(0.50)	(0.49)	(0.49)	(0.50)
Purchase	0.14	0.35	0.20	0.11	0.32	0.17
	(0.34)	(0.48)	(0.40)	(0.31)	(0.47)	(0.38)
Second	0.16	0.22	0.18	0.13	0.17	0.14
	(0.36)	(0.42)	(0.38)	(0.34)	(0.38)	(0.35)
term10	0.03	0.04	0.03	0.02	0.03	0.02
	(0.17)	(0.20)	(0.18)	(0.15)	(0.16)	(0.15)
term15	0.09	0.08	0.09	0.15	0.14	0.15
	(0.28)	(0.28)	(0.28)	(0.35)	(0.35)	(0.35)
term20	0.05	0.05	0.05	0.08	0.08	0.08
	(0.22)	(0.22)	(0.22)	(0.28)	(0.28)	(0.28)

Table 4.2 Continued

term25	0.02	0.02	0.02	0.01	0.01	0.01
	(0.13)	(0.14)	(0.13)	(0.09)	(0.09)	(0.09)
term30	0.81	0.80	0.81	0.74	0.74	0.74
	(0.39)	(0.40)	(0.39)	(0.44)	(0.44)	(0.44)
term40	0.00	0.00	0.00	0.00	0.00	0.00
	(0.04)	(0.02)	(0.03)	0.00	0.00	0.00
YR2002	0.00	0.08	0.02	0.00	0.11	0.03
	0.00	(0.28)	(0.15)	0.00	(0.31)	(0.18)
YR2003	0.09	0.20	0.12	0.11	0.25	0.15
	(0.28)	(0.40)	(0.32)	(0.31)	(0.43)	(0.36)
YR2004	0.27	0.16	0.24	0.30	0.16	0.26
	(0.45)	(0.37)	(0.43)	(0.46)	(0.37)	(0.44)
YR2005	0.40	0.34	0.38	0.40	0.33	0.38
	(0.49)	(0.47)	(0.49)	(0.49)	(0.47)	(0.49)
YR2006	0.24	0.22	0.23	0.19	0.15	0.18
	(0.43)	(0.42)	(0.42)	(0.39)	(0.36)	(0.38)
Education	10.10	10.22	10.13	10.09	10.24	10.13
	(2.19)	(2.35)	(2.24)	(2.24)	(2.39)	(2.29)
Income	58935.55	59370.39	59070.88	58628.16	59180.93	58816.78
	(18651.99)	(18657.30)	(18666.83)	(18572.90)	(18669.57)	(18633.24)
Distance	422.39	294.78	385.36	420.98	265.75	374.14
	(321.52)	(327.35)	(328.33)	(338.39)	(318.80)	(340.06)
Obs	5664	2311	7977	2353	1013	3368

Standard deviations in parentheses

Table 4.3: Online Consumer Summary Statistics by Year (Conforming Loans only)

Variable	2002	2003	2004	2005	2006	Total Online
Credit score	.	668.4572	670.5787	676.4926	670.3322	672.6741
	(.)	(59.7745)	(56.2547)	(59.2812)	(59.4917)	(58.5267)
Rate	.	6.1945	6.3518	6.3427	6.3399	6.3288
	(.)	(1.0133)	(1.4466)	(1.3924)	(1.3294)	(1.3615)
APR	.	6.358	6.5337	6.5153	6.5442	6.5091
	(.)	(1.0478)	(1.4837)	(1.4224)	(1.3861)	(1.3992)
EAPR	.	2.2022	2.2644	2.2298	1.7296	2.1435
	(.)	(1.0327)	(1.475)	(1.4219)	(1.3716)	(1.4055)
Loan To Value	.	73.3411	71.2297	75.2852	80.6216	74.8475
	(.)	(21.5715)	(24.5767)	(26.0623)	(20.0747)	(24.3185)
Loan amount	.	141.1411	137.3732	145.9184	142.3318	142.1419
($1000)	(.)	(76.2235)	(72.0773)	(85.959)	(81.9793)	(80.206)
Upfront Fee	.	1407.655	1318.878	1303.765	1359.193	1330.051
	(.)	(1246.57)	(1088.211)	(1188.352)	(1144.159)	(1157.067)
Fixed	.	.9572	.9298	.9365	.9252	.9346
	(.)	(.2028)	(.2557)	(.244)	(.2634)	(.2473)
Cash out	.	.6537	.6362	.5894	.5918	.611
	(.)	(.4767)	(.4814)	(.4922)	(.4921)	(.4876)
Purchase	.	.0467	.1124	.127	.093	.1074
	(.)	(.2114)	(.316)	(.3331)	(.2907)	(.3097)
Second	.	.0389	.1067	.1524	.1927	.1338
	(.)	(.1938)	(.309)	(.3596)	(.3949)	(.3405)
ICS Fee	.	560.625	615.9494	611.6772	578.8889	601.2686
	(.)	(279.2546)	(291.8635)	(279.3189)	(272.7866)	(282.4647)
Distance	.	408.1278	425.851	426.2261	408.24	420.7695
	(.)	(314.3184)	(329.1149)	(367.064)	(300.8934)	(338.3507)
term10	.	.0272	.0225	.0212	.0227	.0225
	(.)	(.1631)	(.1483)	(.144)	(.149)	(.1484)
term15	.	.1907	.1545	.1333	.1406	.1473
	(.)	(.3936)	(.3617)	(.3401)	(.348)	(.3545)
term20	.	.0817	.0955	.0825	.0703	.0841
	(.)	(.2745)	(.2941)	(.2753)	(.2559)	(.2776)
term25	.	.0078	.0042	.0127	.0045	.0081
	(.)	(.088)	(.0648)	(.112)	(.0673)	(.0895)
term30	.	.6926	.7233	.7503	.7619	.738
	(.)	(.4623)	(.4477)	(.4331)	(.4264)	(.4398)
term40	.	0	0	0	0	0
	(.)	(0)	(0)	(0)	(0)	(0)
AHI	.	57286.02	58738.48	59352.16	57851.2	58660.01
	(.)	(17195.32)	(18146.97)	(19085.89)	(19135.45)	(18619.35)
Education	.	10.0185	10.0428	10.1671	10.0377	10.0891
	(.)	(2.1805)	(2.2095)	(2.2887)	(2.2277)	(2.2414)
Observations	0	257	712	945	441	2355

Standard deviations in parentheses

Table 4.4 : Offline Consumer Summary Statistics by Year (Conforming Loans only)

Variable	2002	2003	2004	2005	2006	Total Offline
Credit score	694.32	687.9602	683.5	676.5714	671.3799	681.6429
	(64.80)	(62.7149)	(63.5518)	(63.456)	(63.0511)	(63.6622)
Rate	6.05	6.1326	6.3911	6.4817	6.2832	6.3042
	(1.02)	(1.2798)	(1.6868)	(1.702)	(1.2535)	(1.481)
APR	6.18	6.2789	6.5549	6.631	6.4675	6.4581
	(1.10)	(1.3183)	(1.7382)	(1.7638)	(1.2791)	(1.532)
EAPR	2.12	2.3187	2.2763	2.3737	1.6698	2.2103
	(1.06)	(1.3174)	(1.7115)	(1.7171)	(1.2994)	(1.5186)
LTV	70.16	70.0321	68.7876	71.2974	80.2708	71.8217
	(20.30)	(22.0907)	(28.8535)	(29.3977)	(22.7995)	(26.0018)
Loan amount	135.48	131.0411	135.8228	134.6435	139.0619	134.7046
($1,000)	(67.24)	(71.4039)	(85.8955)	(74.9377)	(74.3307)	(74.9965)
Commission	671.47	806.5239	876.1493	1040.21	1268.448	950.5802
	(1157.21)	(1316.279)	(1137.503)	(1085.387)	(1131.018)	(1180.603)
Fixed	0.91	.9402	.8902	.9107	.9286	.9172
	(0.29)	(.2375)	(.3135)	(.2856)	(.2584)	(.2757)
Cash out	0.35	.3705	.3963	.3929	.4091	.3852
	(0.48)	(.4839)	(.4906)	(.4891)	(.4933)	(.4869)
Purchase	0.06	.1315	.4268	.4286	.4481	.3182
	(0.25)	(.3386)	(.4961)	(.4956)	(.4989)	(.466)
Second	0.05	.0558	.1951	.2262	.2922	.1695
	(0.21)	(.2299)	(.3975)	(.419)	(.4563)	(.3753)
Distance	140.69	173.1625	284.7453	338.6992	325.4575	265.5781
	(234.41)	(268.8194)	(294.1475)	(370.8595)	(284.8075)	(318.5766)
term10	0.03	.0438	.0183	.0208	.0195	.0266
	(0.16)	(.2051)	(.1344)	(.143)	(.1387)	(.161)
term15	0.11	.1514	.1524	.125	.1753	.1419
	(0.31)	(.3591)	(.3605)	(.3312)	(.3815)	(.3491)
term20	0.13	.1155	.0854	.0595	.0455	.0828
	(0.33)	(.3203)	(.2803)	(.237)	(.209)	(.2757)
term25	0.04	.004	0	.0119	0	.0089
	(0.19)	(.0631)	(0)	(.1086)	(0)	(.0938)
term30	0.70	.6853	.7439	.7827	.7597	.7399
	(0.46)	(.4653)	(.4378)	(.413)	(.4286)	(.4389)
term40	0.00	0	0	0	0	0
	0.00	(0)	(0)	(0)	(0)	(0)
AHI	60411.15	59692.48	58553.74	58811.79	59421.17	59253.67
	(19848.82)	(18925.8)	(18489.75)	(18478.73)	(18865.3)	(18774.16)
education	10.42	10.2691	10.4027	10.1396	10.1507	10.2462
	(2.57)	(2.4977)	(2.3268)	(2.3044)	(2.3539)	(2.3915)
Obs	694.32	251	164	336	154	1015

Standard deviations in parentheses

Table 4.5: Switching Regression

Variable	(1) EAPR Online	(2) EAPR Offline	(3) Probit Online
Credit Score	-0.0055*	-.0055*	-0.000795
	12.57	7.98	1.93
Loan Amount	-0.0051*	-.0061*	0.0038
	15.66	10.38	1.09
Fixed	1.1661*	1.286*	0.134349
	11.18	8.13	1.38
30 yr term	.8038*	.628*	-0.0457
	4.68	2.55	0.28
25 yr term	1.774*	0.0498	0.1297
	5.42	0.1	0.42
20 yr term	.7848*	0.4961	0.046
	4.15	1.69	0.25
15 yr term	0.2566	-0.011	-0.00778
	1.42	0.04	0.04
Loan To Value	-0.0000916	0.0025	.00295*
	0.08	1.26	2.6
Cash out	0.0739	0.0192	.1589*
	1.27	0.17	2.68
Purchase	.5294*	-0.0054	-0.897*
	5.22	0.03	12.04
Second	0.1117	0.19	-0.126
	1.28	1.39	1.57
Coverage Law	-0.0079	-0.0339	-0.0617*
	0.49	1.03	3.62
Enforcement Law	-0.0139	-0.011	-0.1131*
	0.8	0.29	6.39
2006	-.5548*	-.6757*	.9193*
	5.19	3.53	11.04
2005	-0.0413	0.0657	.9441*
	0.42	0.38	13.56
2004	-0.638	0.0801	1.199*
	0.62	0.37	15.54
Distance			0.0045*
			5.04
Education			-0.016
			1.36
Constant	4.922*	5.093*	0.5066
	12.65	7.62	1.32
/lns1 and /lns2	0.2005*	0.2818592*	
	13.12	11.89	
/r1 and /r2	-0.139	0.079	
	-1.89	0.49	
Sigma	1.22*	1.33*	
	0.0187ˣ	0.0314ˣ	
Rho1 and Rho2	-0.138	0.0787	

LR test of Indep Eqns chi2(1)=7.35 Prob>chi2=0.0067

ˣstandard error

*significant at the 5% level

Table 4.6: OLS Price Estimation

	(1) Online Price Equation	(2) Off Price Equation
Credit score	-0.006	-0.005
	(12.78)**	(8.01)**
Loan amount	-0.005	-0.006
	(15.55)**	(10.37)**
Fixed	1.177	1.274
	(11.27)**	(8.10)**
term30	0.803	0.672
	(4.67)**	(2.53)*
term25♠	1.790	0.041
	(5.47)**	(0.08)
term20♠	0.789	0.493
	(4.17)**	(1.66)
term15♠	0.257	-0.013
	(1.42)	(0.05)
LTV	0.000	0.002
	(0.11)	(1.19)
Cash out	0.086	0.005
	(1.48)	(0.04)
Purchase ♠	0.450	0.054
	(4.86)**	(0.45)
Second	0.101	0.200
	(1.16)	(1.47)
Coverage law	-0.012	-0.029
	(0.76)	(0.92)
Enforcement law	-0.023	0.001
	(1.42)	(0.02)
YR2006♠	-0.474	-0.738
	(4.82)**	(5.10)**
YR2005	0.041	0.000
	(0.47)	(0.00)
YR2004	0.034	-0.003
	(0.38)	(0.02)
Constant	4.832	4.949
	(12.51)**	(8.16)**
Observations	2351	1010
R-squared	0.25	0.24

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%
♠ indicates coefficients are not equal across equations

Table 4.7: Results Using Heckman Selection and Exclusion Restriction Tests

	(1) All exclusion Online Price Eq	(2) All exclusion Online Probit	(3) switching exclusion Online Price	(4) switching exclu-sion Online Probit	(5) All exclusion Offline Price Eq.	(6) All exclusion Offline Probit	(7) switching exclusion Offline Price Eq	(8) switching exclusion Offline Probit
Credit Score	-0.005	-0.001	-0.005	-0.001	-0.005	0.001	-0.006	
	(9.74)**	(1.67)	(9.32)**	(1.92)	(7.50)**	(1.62)	(7.59)**	(1.88)
Loan Amount	-0.005	0.000	-0.005	0.000	-0.006	-0.000	-0.006	-0.000
	(14.24)**	(0.94)	(13.86)**	(1.24)	(10.19)**	(0.95)	(9.81)**	(1.28)
Fixed	1.099	0.138	1.085	0.135	1.271	-0.128	1.305	-0.125
	(9.23)**	(1.41)	(8.88)**	(1.39)	(7.86)**	(1.31)	(7.98)**	(1.29)
term30	0.814	-0.050	0.807	-0.049	0.671	0.047	0.674	0.048
	(4.19)**	(0.30)	(4.07)**	(0.30)	(2.55)*	(0.29)	(2.54)*	(0.29)
term25	1.677	0.112	1.658	0.139	0.037	-0.112	0.069	-0.139
	(4.52)**	(0.36)	(4.37)**	(0.44)	(0.07)	(0.36)	(0.13)	(0.45)
term20	0.758	0.039	0.749	0.038	0.491	-0.040	0.501	-0.038
	(3.53)**	(0.21)	(3.42)**	(0.21)	(1.67)	(0.22)	(1.69)	(0.21)
term15	0.249	-0.011	0.242	-0.011	-0.014	0.009	-0.006	0.010
	(1.22)	(0.06)	(1.16)	(0.06)	(0.05)	(0.05)	(0.02)	(0.06)
LTV	-0.002	0.003	-0.002	0.003	0.002	-0.003	0.003	-0.003
	(1.14)	(2.52)*	(1.24)	(2.62)**	(1.13)	(2.45)*	(1.34)	(2.54)*
Cash out	-0.012	0.153	-0.021	0.158	0.000	-0.152	0.044	-0.156
	(0.18)	(2.58)**	(0.29)	(2.67)**	(0.00)	(2.55)*	(0.36)	(2.64)**
Purchase	1.083	-0.897	1.136	-0.896	0.074	0.896	-0.113	0.895
	(6.11)**	(12.03)**	(5.61)**	(12.04)**	(0.30)	(12.02)**	(0.41)	(12.04)**
Second	0.187	-0.135	0.195	-0.125	0.203	0.139	0.171	0.129
	(1.88)	(1.68)	(1.91)	(1.55)	(1.45)	(1.73)	(1.20)	(1.61)
Coverage	0.022	-0.058	0.026	-0.061	-0.027	0.058	-0.043	0.062
Law	(1.09)	(3.39)**	(1.22)	(3.60)**	(0.75)	(3.42)**	(1.14)	(3.63)**
Enforcement	0.053	-0.086	0.062	-0.113	0.005	0.086	-0.031	0.113
Law	(1.99)*	(4.55)**	(2.09)*	(6.43)**	(0.09)	(4.57)**	(0.56)	(6.45)**
YR2006	-1.115	0.900	-1.181	0.920	-0.759	-0.898	-0.562	-0.917
	(6.02)**	(10.77)**	(5.55)**	(11.05)**	(2.83)**	(10.75)**	(1.88)	(11.02)**
YR2005	-0.613	0.939	-0.677	0.944	-0.022	-0.935	0.184	-0.941
	(3.40)**	(13.44)**	(3.24)**	(13.56)**	(0.08)	(13.40)**	(0.62)	(13.53)**

Table 4.7 Continued

YR2004	-0.748	1.180	-0.824	1.195	-0.030	-1.178	0.229	-1.193
	(3.59)**	(15.26)**	(3.39)**	(15.51)**	(0.09)	(15.25)**	(0.62)	(15.50)**
Distance		0.001		0.000		-0.001		-0.000
		(5.99)**		(4.91)**		(5.96)**		(4.87)**
distance2		-0.000				0.000		
		(4.16)**				(4.14)**		
AHI		0.000				-0.000		
		(1.31)				(1.41)		
Education		-0.033		-0.016		0.033		0.016
		(1.84)		(1.41)		(1.88)		(1.36)
Constant	5.545	0.265	5.625	0.508	4.902	-0.268	5.350	-0.512
	(11.91)**	(0.68)	(11.56)**	(1.33)	(6.20)**	(0.69)	(6.30)**	(1.33)
Observa-tions	3358	3358	3359	3359	3360	3360	3361	3361

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

Table 4.8: OLS Profit Estimation

	(1) Online Price Equation	(2) Off Price Equation
Credit score	-0.004	-0.001
	(9.61)**	(8.01)**
Loan amount	-0.01	-0.01
	(13.97)	(10.37)**
Fixed ♠	.911	1.27
	(8.27)**	(8.10)**
term30♠	.98	0.67
	(5.42)**	(2.53)*
term25♠	2.27	0.04
	(6.53)**	(0.08)
term20♠	1.895	0.49
	(9.43)**	(1.66)
term15♠	2.24	-0.01
	(11.72)**	(0.05)
LTV	0	0
	(0.61)	(1.19)
Cash out	0.12	0.01
	(2.00)**	(0.04)
Purchase ♠	.237	0.05
	(2.42)	(0.45)
Second	.032	0.2
	(0.36)	(1.47)
Coverage law	-0.029	-0.03
	(1.70)	(0.92)
Enforcement law	-0.052	0
	(2.98)**	(0.02)
YR2006♠	0.075	-0.74
	(0.79)**	(5.10)**
YR2005	0.062	0
	(0.67)	(0.0)
YR2004♠	-0.425	0
	(4.08)**	(0.02)
Constant	4.80	4.95
	(11.74)**	(8.16)**
Observations	2351	1010
R-squared	0.29	.24

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

♠ indicates coefficients are not equal across equations

5. CHAPTER 5. Asymmetric Information and the Lemons Problem in Mortgage Banking

5.1 Introduction

The literature investigating asymmetric information in the mortgage market has mainly been concerned with principal-agent problems that arise during the process of securitization. Lenders supplying mortgage loans to the secondary market are relatively more informed about the quality of the underlying collateral. Specifically, the lender is more informed about the probability of default or prepayment associated with each loan than are potential buyers of the loan in the secondary market. Since the value of an asset is the discounted value of the future stream of income, information about the probability of default or prepayment is valuable to the secondary market and in order to accurately price the asset.

The described securitization process would predict a lemons market outcome, as suggested by Akerloff (1970). This occurs because lenders possess more accurate information about the true quality of each loan, while demanders of loans only identify the distribution of overall quality. As a result of the information asymmetries, the lemons problem would predict that lenders will sell the low quality loans to the secondary market and retain the least risky (high quality) loans on their own balance sheets. The outcome of the lemons problem would suggest that the behavior of agents could lead to a collapse of the Mortgage Backed Securities market.

Demanders of mortgage loans would expect that the mortgages supplied to the secondary loan market to be of low quality. However, in 2004, 65% of all originated

loans were repackaged as Mortgage Backed Securities, creating a total of $1.85 trillion of bonds.[46] The volume of activity in the Mortgage Backed Securities market would suggest the presence of a vibrant market despite the existence of asymmetric information. This leads to a postulation that other market factors exist which encourage lenders to supply mortgages to the secondary market, or that market mechanisms have been developed to mitigate the lemons problem.

In this chapter, I test for the presence of asymmetric information in the mortgage market. Using the Oaxaca-Blinder decomposition, I examine pricing variations between the broker and lender segments of the mortgage market. The decomposition allows the separation of any difference in price into explained and unexplained differences. The unexplained difference is a measure of asymmetric information. I do not find any evidence of asymmetric information or sorting behavior in the mortgage market.

5.2 Literature

5.2.1 *Lemons Market*

The lemons problem, as described in Akerloff (1970), exists in markets where the quality of a good varies and some market participants have imperfect information. Specifically, when sellers know the true quality of their product but buyers only know the distribution of overall quality in the market. Buyers entering the market must infer the quality of the product being sold. Any buyer's best estimate of quality would be the market's average quality for that product. Since price does not reflect the true quality of the

[46] Downing, C., D. Jaffee, et al. (2005). "Information Asymmetries in the Mortgage Backed Securities Market." University of California at Berkeley **May**.

seller's product, sellers with high-quality goods would not have an incentive to enter the market. As high-quality sellers exit the market, the average market quality and price fall. This forces more sellers to exit the market. Eventually if enough sellers exit, the market will collapse.

The severity of the lemons problem will depend on how much private information buyers can gain about the seller's product. Sellers' behaviors can serve as a signal and provide buyers with valuable information about the quality of their product. Genesove (1993) explores how the market responds when buyers gain information about a seller's type. The example provided is of a consumer in search of an apple. Only the seller of the apple knows the true quality of the product. Buyers only know that there are two types of sellers; farmers that love apples and farmers that hate apples. A Farmer's preference for apples can send valuable information to the buyer and save the market from its collapse. It is expected that a farmer that loves apples will retain the good apples for their own consumption, while the farmer who hates apples will take all their apples to the market. Incorporating this information into the market will provide a price premium for apples sold by farmers that hate apples.

Genesove test this hypothesis by examining used car auctions and finds some support. In the study, sellers of used cars are either new-car dealers or used-car dealers. New-car dealers are expected to sell all their used cars at the auction, while used-car dealers would retain the good-quality and sell the low-quality cars. Therefore it is expected that new-car dealers should receive a premium for their cars. In a similar study, Chezum and Wimmer (1997) explore price differentials in the thoroughbred yearlings market. In the thoroughbred industry two types of sellers exist. There are "breeders," who

only breed horses to be sold, and "racers," who both bread and race horses. The authors find evidence that "breeders" earn a premium for their yearlings relative to "racers". It is expected that "racers" would only bring their low-quality horses to the market, and retain their good-quality horses.

5.2.2 *Asymmetric Information in the Mortgage Market*

Glaeser et al (1997) provide a model that actually credits the presence of a liquid securities market to the existence of ex-ante information asymmetries in the securitization process. If the secondary market gains access to accurate information about the loans grouped into the Mortgage Backed Securities, they would only demand the high quality loans. Downing et al (2005) further explain that the government sponsored enterprises (GSEs), Fannie Mae and Freddie Mac, purposefully restrict information about the mortgages which back their securities. The GSEs insure loans which meet their guidelines and securitize them to increase liquidity for mortgages and facilitate mortgage origination. The GSE groups loans with common duration but allows all other quality measures to vary as long as they meet the minimum standards. Purchasers of the Mortgage Backed Securities do not have any information about the quality of the pool. They just know the expected duration of the income stream. Investors in mortgage securities invest in the portfolio of loans and are concerned with the overall quality of the pool. The pooling of mortgage loans would suggest distribution of quality within the pool of securitized loans, and consequently, there must be an incentive for lenders to also sell high quality loans.

When examining the securitization decision for banks, a subset of mortgage lenders, Ambrose et al (2005) contend that banks should securitize all loans. The securitiza-

tion process would minimize exposure to credit risk and interest rate risk, would increase liquidity, and would reduce the required reserve requirements. Specifically, the authors propose that the capital requirements established by Basel capital rules[47] create an incentive for banks to sell low risk loans and retain high risk loans. This occurs because the Basel I capital requirements do not vary with the quality of mortgages, and thus the capital requirements would be too low for high risk loans. In an empirical test the authors find that securitized loans experience lower default rates compared to loans retained as part of the banks' portfolios. They attribute their findings to two reasons: first, banks are responding to the incentives provided by Basel I; and second, they are responding to the reputation effect. The reputation effect exists due to the repeated interaction between sellers (lenders) and buyers of loans (secondary market). The repeated interaction compels banks to supply the secondary market with loans which have lower probabilities of default and prepayment, therefore maintaining demand for the mortgages they supply.

Third party originators, like mortgage brokers, do not face the same incentive structure as banks. Third party originators are classified as middlemen between the lender and the consumer and are usually a good source of new business to the lender. Mortgage brokers often carry lower overhead cost, typically use incentive pay structure, and consequently attract more productive loan officers (LaCour-Little and Chun 1999).

TPO operations are a large fraction of the overall mortgage market as indicated by their 52% share of all mortgages from 1997 to 2004. These operations employ 418,700

[47] Basel accords are banking laws established by Basel Committee on Banking and Supervision (BCBS) at the Bank of International Settlements. In 2004, Basel II was recommended and it allowed capital requirements to adjust with the risk of each loan. Basel I had a fixed 50% capital requirement for all residential mortgages. http://www.cml.org.uk/cml/policy/issues/748

people at 53,000 companies.[48] However, their presence further magnifies the principal-agent problem so inherent in the securitization process. LaCour-Little et al (1999) and Alexander et al. (2002) provide evidence that loans originated through TPOs are more likely to prepay or default compared to loans originated through the lender. TPOs are compensated for originating loans but are not directly accountable for the loan performance of the originated loans. Therefore, it is expected that TPOs would not be meticulously screening borrowers for default.

Alexander et al. (2002) find that loans originated by TPOs are more likely to default after controlling for observable default characteristics. LaCour-Little et al. (1999) find that TPO loans are more likely to be refinanced and prepaid in comparison to retail loans. Mortgage brokers have an incentive to "churn" the consumer by refinancing the mortgage when market conditions are favorable because market interest rates are falling. The higher prepayment and default rates reduce the discounted value of the income stream and are therefore costly to investors. According to Alexander et al. (2002) and Spader et al. (2009), the recent market response has been to increase interest rates on TPO originated loans by as much as 50 basis points, or one-half of one percent. Another mechanism to control the principal-agent problem inherent with TPO loans is to impose pre-payment penalties on brokered loans. (LaCour-Little et al. 1999). Brokers can either choose to incur the 50 basis point penalty or impose a prepayment penalty on consumer loans.

The 50 basis point penalty is implemented to reflect the higher risk of TPO loans relative to lender-originated loans. However, the penalty has reduced the competitiveness

[48] National Mortgage Broker Association website http://www.namb.org/namb/Mission.asp?SnID=1585056730

of the TPOs in mortgage origination. The difference in prices between lender originated loans and third-party originators has created an incentive for brokers to vertically integrate into lending operations. The new structure of the TPOs allows them to lend their own funds in addition to maintaining the option of serving as a middleman between other lenders and the consumer. The hybrid lender (or hybrid TPO) originates and lends its own funds; however, it does not service the loans or maintain any loans on portfolio. Financial constraints force the hybrid lender to securitize the loans immediately after origination. Additionally, the firm maintains the option to act as a regular broker and originate loans for other lenders. Since these firms do not maintain any loans on portfolio and, unlike banks, do not face Basel I capital requirements, any sorting behavior can be attributed to asymmetric information. Ambrose et al. (2005) are unable to separate the capital requirement effect and the asymmetric information effect in their results. By examining the hybrid lender, I can attribute any sorting behavior directly to the presence of asymmetric information in the mortgage market. Furthermore, these firms are new in the mortgage industry and their behavior has not been examined. Alexander et al. (2002) postulate that the 50 basis point penalty was introduced during the period of their study, which ranged from 1996 to 1998. This would suggest the incentive for a TPO to vertically integrate into lending occurred between 1996 and 1998.

5.3 Hybrid Lender

During the origination of a loan, a hybrid lender has two options: it can choose to operate as a TPO or as a lender. As a TPO, the firm serves as a middleman between the consumer and downstream lender. In this method, the hybrid lender receives commission

from the consumer and may earn yield spread premium (YSP) from the lender. The firm does not experience any cost of default or repayment for loans originated as a TPO.

Alternatively, the hybrid lender can choose to act as a lender and incur the responsibility of underwriting and funding the loan. In this instance, the firm is responsible for insuring that the loan meets the secondary market's requirements, and any default or repayment would become the responsibility of the firm. Upon funding, the hybrid lender sells the loan into the secondary market to be pooled into securities. The sale of the loan usually occurs within 90 days of funding (Ambrose, LaCour-Little et al. 2005).

During the sale of the loan to the secondary market, the hybrid lender can use either the cash or swap program. The swap program requires the lender to provide a pool of loans, which are swapped for a Mortgage Backed Securities from either of the GSEs or from a private firm (to remain consistent with the literature, I will refer to the private firms in the securitization market as Private Label). The hybrid lender operations do not generate large enough volume to utilize the swap program. In addition, due to financial constraints, hybrid lenders typically need to sell their loans quickly and cannot wait to create a large pool of loans to sell. Therefore, hybrid lenders usually sell their loans through the cash program.

The cash program allows the firm to sell the rights to the loan to the GSE or private label firm and receive the increased liquidity.[49] The purchaser of the loan is responsible for packaging the loans into multi-lender pools and securitizing them. By liquidat-

[49] Government Sponsored Enterprises insure and purchase conventional or conforming loans. Loans that do not meet Fannie Mae and Freddie Mac guidelines are sold to Private Label firms. Private Label loans are not insured and consequently are packaged into riskier Mortgage Backed Securities.

ing its mortgages, the firm is able to finance future loans and continue its operations. Lending its own funds allows the hybrid lender to reduce its cost. As a lender, the firm bypasses the 50 basis point penalty imposed on TPO loans and thus can remain competitive in the mortgage market. The cost to the firm for operating as a lender is that it bears the risk of prepayment or default. Since the firm has to continuously transact with agents in the secondary market, it must maintain a good reputation. If the firm continuously sells poor quality loans to the secondary market, the secondary market may impose conditions on future purchases. On the other hand, if the loan is brokered, then the downstream lender incurs all credit risk, interest risk, and reputation cost.

5.4 Sorting Hypothesis.

The structure of the origination process suggests some loan qualities that may be used to sort loans and may affect pricing of a mortgage. If information asymmetry does affect the firm's lend-broker decision, then the following relationships should be observed.

5.4.1 *Termination Risk*

It is expected that the firm will choose to lend loans with lower probabilities of termination. If the reputation effect holds, then the firm will attempt to originate loans with better observable qualities. To maintain demand for its loans in the future, the hybrid lender will insure that the loans it provides to the secondary market have a low probability of default or prepayment based on observable information. The firm may sort by specifically selecting mortgages which have low risk or it may sort using prices as a mechan-

ism. To attract good quality customers and obtain only low risk loans, the firm may increase prices on high risk characteristics,therefore creating the incentive for loan officers to lend when risk is low and brokering when risk is high. If pricing is used as a sorting mechanism, then the firm's pricing for non-brokered loans would be lower than the pricing for brokered loans and when controlling for the lower cost of funds.

5.4.2 *Distance*

The relationship between lender and borrower suffers from asymmetric information itself. The lender does not know the true default or prepayment probabilities of the loan. As an approximation of termination risk, the lender uses known credit and consumer characteristics. In addition, the lender has to rely on a third party (appraiser) to provide information on the housing conditions in the consumer's market. Therefore, it is expected that the degree of asymmetric information increases with distance from the location of the firm. Garmaise et al (2004) provide evidence that distance is a factor in real estate transactions, and serves as an indirect measure of the degree of asymmetric information.

The hypothesis is that the hybrid lender will be more likely to broker loans as the distance between the consumer and the firm increases. As suggested by Heuson et al. (2001), banks are more familiar with their local economies, and the strategies they pursue will be based on a combination of their own information, local conditions, and statistical evidence. Thus, distance can be utilized as an indirect measure of asymmetric information.

5.4.3 *Market Interest rates and volatility*

To reduce interest rate risk, it is expected that the firm will broker loans when the market interest rates are volatile. Furthermore, given the evidence provided by LaCour-Little et al (1999) that mortgage brokers are likely to re-solicit consumers to refinance when market interest rates fall, I expect that brokering activity will increase as market interest rates increase. As a broker, the firm can solicit the consumer to refinance when the interest rates decrease without jeopardizing its reputation in the secondary market. The shift in demand is expected to adjust the prices in the broker and lender markets and should be accounted for.

5.4.4 *Liquidity*

The firm's objective may be to act as a lender with all loans, regardless of qualities. However, since the firm faces liquidity constraints, its ability to lend depends on whether previous loans have been sold to the secondary market. Therefore, the firm may choose to broker loans as it approaches its credit constraint.

A limitation of this study and a weakness of the empirical test is that I cannot test for this hypothesis. However, liquidity constraints are expected to be negligible during the time of the study as investment in mortgage-backed securities and real estate was high during the housing boom. Therefore, the firm would have been able to liquidate mortgages easily.

5.5 Data

Firm-level data are used to test for information asymmetries in the presence of a hybrid lender. The dataset contains loan level information for a hybrid lender based in Louisville, Kentucky. The sample of transactions includes 9,140 loans originated between July 2002 and August 2007. During the time period, many mortgage brokers expanded their mortgage business to include both lending and brokering operations. Although it is unclear how many firms simultaneously operate in the broker and lender markets,[50] the firm in this study and five other local competitors expanded into mortgage lending between 2003 and 2006.

Initially the firm operated solely as a third party originator (broker), serving as a middleman between consumers and lenders. In September 2005, the firm expanded its operations so it could act as a lender on its own behalf. After the commencement of lending operations, the firm originated a total of 4,479 loans; 57% of these were originated as a broker. Therefore, for 4,479 transactions, the firm had a choice between acting as a lender or a third party originator. A summary of transactions by period is provided in Table 5.1.

The firm's decision to expand into lending operations may have been an attempt to (1) expand available loan programs to the consumer, (2) specialize in a subset of loan programs, (3) or to bypass the 50 basis point penalty. Table 5.2 compares the quality of loans originated during third party origination (pre-2005) and hybrid lending operations (post-2005). Comparing the quality of loans between the two periods does not indicate

[50] http;//www.mortgageloan.com/mortgage-lenders.

any obvious difference between the samples. Prepayment and default variables such as loan amount, loan to value ratio, and consumer credit scores are similar across samples. The amount of commission the loan officers received during the hybrid lending operations was $1,837, almost identical to average commission of $1,834 received under the broker only system. Loans originated under the hybrid system paid a higher interest rate of 7.4%, compared to 6.5% for loans originated before lending operations began. However, as Figure 5.1 reveals, the variation in interest rates is representative of the macroeconomic environment and increasing interest rates between 2002 and 2007. There is no evidence of moral hazard or adverse selection across the two periods. By expanding the type of operations, the firm originated the same type of loans. The firm's decision to vertically integrate into lending operations seems to be an effort to increase its profitability by eliminating the 50 basis point penalty.

Summary statistics comparing the non-brokered and brokered loans during the hybrid lending operations are presented in Table 5.3. The hypothesis that the firm is more likely to lend higher quality loans and broker lower quality loans is not immediately evident in the data. The samples are similar in default and prepayment variables like loan to value ratio, and credit scores; consequently, they do not indicate any sorting behavior by quality. However, the absence of sorting does not negate the presence of information asymmetries. Default and prepayment variables are observable to all market participants, while the lemons problem (Akerlof 1970) occurs when the quality measures are observable to one party of the market but not the other. The difficulty in testing for asymmetric information is the fact that by assumption the information is unobservable.

The distance variable provides an indirect measure of information asymmetries. Distance between lender and consumer is known to the firm but not observable to the buyers of mortgages in the secondary market. However, the prediction that the firm would be less likely to lend as the distance between consumer and the firm increases is unsubstantiated. As a lender, the average distance between the consumer and firm is 437 miles, while the distance for brokered loans is 397 miles.[51] If information asymmetries increase with the distance between the consumer and firm, then it would be expected that the firm would sort loans by distance, choosing to lend loans that are closer to the home office. There is a slight difference between the samples with respect to the loan amounts. The mean loan amount for lending is $123,000; this is slightly lower than brokered loans, which averaged $142,000.

5.6 Methodology

The central question of this research is whether there is a systematic difference between the loans the firm originated as a lender and the brokered loans. If there is a variation in the type of loans originated, is it due to quality differences and/or profitability? If there is no systematic difference between the loans in the operations, this suggests that the firm is willing to act as a lender for all loans but resorts to brokering loans when it meets its capital constraint.

To test for the presence of asymmetric information and consequently adverse selection, I use the Blinder-Oaxaca decomposition (Blinder 1973; Oaxaca 1973). Using the Lending and Brokered loans as the groups of interest, the decomposition will divide the

[51] As an alternative measure of indirect information, I use local vs. nonlocal classification. During Hybrid operation, 12% of loans originated were in the Louisville area.

difference in price into "explained" differences arising from variation in loan quality and "unexplained" differences that are unaccounted for by the observable information. The systematic presence of unexplained differences would suggest that there are unobserved determinants in mortgage pricing and would imply that asymmetric information creates sorting behavior. Nevertheless, due to the 50 basis point broker penalty, it is expected that unobserved difference may exist between the groups and may be reflected in the price variation. Consequently, the magnitude of the unexplained variation would also be of importance.

The Blinder-Oaxaca decomposition is applied most often in labor market and discrimination studies. It is often utilized to identify explained and unexplained differences in the male-female wages. The explained portion of the wage gap occurs because of variations in productivity between the groups, while any unexplained difference may be considered a measure of discrimination. This approach can be applied in the mortgage market by examining mortgage prices. Since the interest rate of a mortgage is the price of the underlying risk, any differences between prices can be attributed to differences associated with observable and unobservable risk. An efficient market should base pricing on observable information, which accounts for prepayment and default risk. It is expected that, if asymmetric information does not exist in the mortgage market, pricing would be solely based on observable information and any variations in pricing would be accounted for by explained differences only.

The method of the decomposition is to determine how much of the mean mortgage price differential is accounted for by the observable determinants of risk. Using the

model provided by Ambrose et al (2004; 2005), I estimate a linear price equation (equation 5.1) for the brokered loans and another for non-brokered loans by including observable risk and pricing determinants:

$$\boldsymbol{P_i = \beta_0 + \beta_1 X + \beta_2 mrktconditions + \beta_4 law\ index + \varepsilon} \qquad (5.1)$$

The price $\boldsymbol{P_i}$ is defined as the annual percentage rate minus the 10 year treasury. The matrix X includes loan and consumer characteristics which are used to measure the probability of default and prepayment. It includes the loan amount, consumer's credit score, a dummy variable indicating whether the loan meets GSE guidelines, and the loan-to-value ratio. To control for loan characteristics, I include a dummy variable indicating whether the loan is a fixed rate mortgage, a dummy variable for second lien loans, and a dummy variable for 10, 15, 20, 25 and 30 year mortgages. I follow Ambrose and Pennington-Cross (2000) and include the law index to control for spatial variation in regulatory environments.[52] The estimation also controls for the macroeconomic environment by including a measure of market credit risk premium, defined as the difference between the Aaa and Baa bond rates. Additionally, I control for the slope of the yield curve by calculating the difference between the 10 year and the one year Treasury bond rates. The credit risk premium will capture the market's preference for safe assets (Aaa) relative to riskier assets (Baa). An increasing credit risk premium signifies an increase in demand for safe or high quality assets. A steeper slope of the yield curve represents a macroeconomic environment that incentivizes investment in long term assets.

The difference in the mean outcome is given by:

[52] The law index is included to capture any differences in pricing that arise due to state regulatory environments.

$$\boldsymbol{D = E(P_b) - E(P_l)} \tag{5.2}$$

Here, $E(P_b)$ denotes the expected price of brokered loans, and $E(P_l)$ is the expected price of the non-brokered loans. For simplicity, equation (5.1) can be written as:

$$\boldsymbol{P_i = X_i' B_i + \epsilon_i, \qquad E(\epsilon_i) = 0, \quad i \in \{b, l\}} \tag{5.3}$$

Then the difference in expected price can be rewritten as:

$$\boldsymbol{D = E(P_b) - E(P_l) = E(X_b)' B_b - E(X_l)' B_l} \tag{5.4}$$

By assumption $E(B_i) = B_i \; and \; E(\epsilon_i) = 0$. Rearranging equation 5.4, the difference in expected price can be divided into three parts and provides the "three-fold" decomposition.

$$\boldsymbol{D = [E(X_b) - E(X_l)]' B_b + [E(X_b)]'(B_b - B_l) + [E(X_b) - E(X_l)]'(B_b - B_l)} \tag{5.5}$$

$$\boldsymbol{D = E + C + I}$$

where $E = [E(X_b) - E(X_l)]' B_b$ accounts for group difference in the quality of loans. The variation in pricing due to differences in coefficients is given by "C" and is equal to $[E(X_b)]'(B_b - B_l)$. Finally," I" is the interaction of differences of attributes and coefficients between the broker and lender groups.

The explained component (E) is the expected change in the price of brokered loans, if brokered loans attributes were equal to the attributes of non-brokered loans. The "C" portion measures the change in the price of brokered loans, if brokered loans had the coefficients associated with non-brokered loans. The Blinder-Oaxaca decomposition will

assist in identifying if there is a systematic difference in quality of loans between the groups. Additionally, if the unexplained portion of the difference, denoted by 'C,' is less than -50 basis point, it implies that non-brokered loans have positive unobserved qualities, which leads to prices being lower in the lending operations. It is possible that unobserved qualities exist if C is larger or equal to -50 basis points. This will occur if a portion $\alpha \in [0,1)$ of the lower marginal cost associated with lending are not passed on to the consumer. In this instance, unobserved qualities may be present but will be unidentifiable.

The above decomposition can be rearranged to provide an explained and unexplained outcome (Jann 2008). The decomposition can be rewritten as the "two-fold" decomposition if it is assumed that one group pays standard pricing and the other receives "preferential" pricing. That is, pricing adjustments only occur on one side of the market. If it assumed that the broker market has a standard pricing $\beta^* = \beta_b$ and non-brokered loans prices are adjusted, then the difference can be written as in equation (5.6). This assumption is similar to the discrimination literature, which usually assumes that discrimination is directed towards one group only (Women).

$$\boldsymbol{D = [E(X_b) - E(X_l)]'\beta^* + [E(X_b)]'(B_b - \beta^*) + [E(X_b) - E(X_l)]'(\beta^* - B_l)} \qquad (5.6)$$

$$\boldsymbol{D = Q + U}$$

In the two fold decomposition, the explained portion is denoted by Q which is equal to $[E(X_b) - E(X_l)]'\beta^*$. The unexplained difference in expected means is given by:

$$\boldsymbol{U = [E(X_b)]'(B_b - \beta^*) + [E(X_b) - E(X_l)]'(\beta^* - B_l)} \qquad (5.7)$$

The "two fold" decomposition provides an estimate of the effect of unobserved qualities on the prices of mortgages. Similar to the "three fold" decomposition, if unobserved qualities do exist, they will be captured by the coefficient on U. If the estimate exceeds 50 basis points, it will imply that unobserved qualities do influence pricing and sorting by the firm. However, this interpretation assumes that all observed predictors are included in the estimation.

5.7 Results

The results of the OLS estimation for brokered (column 1), non-brokered loans (column 2), and pooled (column 3) samples are reported in Table 5.4. The coefficients of the estimation are consistent with expectation and have signs similar to the findings in Ambrose (2004; 2005). In both the broker and lending specifications, the price of a mortgage increases as observable risk of a loan increases. Although the signs are as expected in both specifications, there is a variation in pricing of the risk characteristics across markets.

The pricing of each observable risk factor is most higher in the broker market relative to the lending market. In the broker market, loans meeting conforming guidelines are priced 51 basis points lower than non-conforming loans. The magnitude of the conforming status in the lending market is only 5 basis points and is statistically insignificant. For an increase of 10 units in consumer's credit scores, broker and lending prices increase by 4 and 1 basis points, respectively. Mortgages used to fund the purchase of a new home are priced equally in both markets; they receive a price of 8 basis points less

than refinance loans. As the loan-to-value increases by one percentage point, lending prices increase by 2 basis points compared to 1 basis point in the broker market.

Prices in the broker market are responsive to changes in the macroeconomic environment. For every unit increase in the difference between the 10 year treasury and one year treasure, mortgage prices increase by 27 basis points. However, an increase of one unit in the yield curve results in a decrease of 21 basis points in lending market, but the coefficient is insignificant. Prices in both the broker and lending markets are not responsive to changes in credit risk. As market interest rate volatility increases, prices in the broker market increase by 91 basis points and by 133 basis points in the lending market. This indicates that both markets pass on the risk of interest rate movements to the consumer.

The main coefficient of interest to test for the role of asymmetric information in mortgage pricing is the distance variable. For every additional mile of distance between the consumer and firm, the price of a mortgage increases by 5 basis points on average. For the broker market this represents an increases of 5.5 basis points, while price in the lending market increases by 4.2 basis points. Although there is evidence of pricing due to informational frictions, it was expected that the effect of distance on price would be more prominent in the lending market. The dummy variable for non-brokered loans in the pooled specification provides an estimate of the reduction of price lending consumers receive. On average, loans originated through the firm's lending division receive a price 18 basis points less than brokered loans.

The variation in prices can arise due to either unobservable information or the lower marginal cost of lending. If the variation is due to lower marginal cost, then the firm is passing on 18 basis points to the consumer because it is bypassing the broker penalty. Alternatively, if it is assumed that none of the cost is transferred to the consumer, then the variation in price is due to unobservable qualities of the loan. If the unobservable qualities of a loan are high, then the firm can price a loan lower than the observable information would predict. Consequently, since non-brokered loans are priced 18 basis points lower than brokered loans and when assuming lower cost are not transferred to the consumer, then it must be that the unobserved qualities of non-brokered loans is higher.

To further test the source of variation in pricing between brokered and non-brokered loans, I use Blinder-Oaxaca decomposition. Table 5.5 reports the results from the "three-fold" Blinder-Oaxaca decomposition and the predicted expected price for each group. The mean expected price for brokered loans is 3.445% (annual percentage rate minus the 10 year treasury rate); this is compared to 3.410% for non-brokered loans. The difference between the mean expected prices across the groups is 0.035%. The difference is statistically insignificant at the 95% level.

The decomposition of the difference in expected prices yields a -0.22% coefficient attributed to the endowment effect. It is statistically significant and implies that the mean price of non-brokered loans would decrease by 22 basis points, if non-brokered loans had the loan qualities associated with the brokered loans. If the price of a mortgage reflects the underlying risk, then on average, the observable qualities of loans for which the firm chooses to act as a lender is lower than brokered loans. The result indicates the

presence of a systematic sorting behavior by the firm. Loans with higher qualities, based on observable information, are brokered, whereas loans with low observed quality are part of the lending portfolio. This result conflicts with the termination risk and reputation hypothesis, which predicts that the firm will lend when termination risk is low in order to develop a reputation in the secondary market.

The portion of the decomposed difference attributed to variations in coefficients accounts for 14 basis points and is statistically significant. If non-brokered loans were priced similar to the brokered loans, non-brokered loans would be priced 14 basis points higher. This implies that either the firm does not price observable risk as aggressively as does the broker market or unobservable information is used to determine the price of non-brokered loans. Alternatively, we can assume that the firm is passing on the gains from lower marginal cost of lending to the consumer. The results do not indicate any explicit evidence that the firm charges lower prices in lending markets due to unobserved loan qualities. Therefore, the hypothesis that the firm brokers the "lemons" and retains the good quality loans is unsupported.

The interaction term in "three-fold" method absorbs some of the effect of the un-explained differences. Consequently, as a robustness test, I use the two-fold decomposition and assume that the broker market has the standard mortgage pricing. In this specification, the preferential pricing occurs towards one side of the market--the lending market. The results from the "two-fold" decomposition are given in Table 5.6. The difference in pricing and predicted prices remains unchanged, as would be expected. However, the explained portion of the decomposition now accounts for -.15%. Consequently, due to the

variations in loan qualities, non-brokered loans would be priced 15 basis points lower if they had the characteristics associated with the brokered loans. The unexplained variations account for 18 basis points of the difference.[53] Therefore, brokered loans are priced 18 basis points higher than non-brokered loans. This result is well within the 50 basis point difference expected due to the broker penalty, and adjusting the specification does not change the results.

5.8 Conclusion

This chapter examines the presence of information asymmetries in the mortgage industry. Specifically, it asks if information available to a mortgage firm but unobserved by the secondary market will affect a firm's decision to lend its own funds. Previous studies, like Ambrose (2005), have examined banks and found that sorting does exist in the mortgage market. Banks are more likely to retain loans in their own portfolio when loans have positive unobserved risk qualities. However, studies have not been able to disaggregate the reputation effect and the capital requirement effect.

In this study, I examine a hybrid lender, which does not confront capital requirements. The firm can either broker or lend its own funds when originating a mortgage. Since the reputation effect and asymmetric information exist in the lending market and not the broker market, I can test for sorting by observed quality. To test for variation in unobserved qualities of loans between market types, I use the Blinder-Oaxaca decomposition.

[53] To ensure that these results are not driven by income variations, which are unobserved in the data but observed by the firm and the secondary market, I estimate the model using the specification in Chapter 4. I reduce the sample to conforming loans to eliminate income as a determinant of price. The results are unchanged.

The results of the decomposition provide evidence of a variation in pricing between the broker and lending markets that is not explained by observable risk factors. Loans that are part of the firm's lending business are priced lower than brokered loans. The firm is charging lower prices because of the higher unobserved qualities of the non-brokered loans. However, these results are inconclusive because the effect cannot be disaggregated from the effect due to the lower marginal cost of lending.

Table 5.1: Number of transactions by type

Period	Transactions	Brokered	Lending
Entire Sample (2002-2007)	9,140	7,209	1,931
TPO Period (2002-2005)	4,661	4,661	0
Hybrid lender Period (2005-2007)	4,479	2,548	1931

Table 5.2: Hybrid vs TPO Samples

	Hybrid		TPO	
Variable	Mean	Std. Dev	Mean	Std. Dev.
Credit score	671.13	67.13	669.84	64.95
Note rate	7.40	1.69	6.51	1.60
Commission ($)	1387.86	1439.42	1383.27	1304.77
Loan amount ($1,000)	134.40	101.37	132.97	86.50
Loan To Value	65.70	28.43	70.29	25.74
Fixed	0.64	0.48	0.62	0.48
Second	0.23	0.42	0.13	0.33
Cashout	0.53	0.50	0.54	0.50
Distance	413.00	321.33	370.86	337.66
Loan term (months)	338.91	58.83	323.03	69.77
Purchase	0.19	0.39	0.17	0.38
Rate and term	0.25	0.43	0.27	0.44
30 year mortgage	0.84	0.37	0.75	0.43
Observations	4479		4661	

Table 5.3: Descriptive Statistics

	Non-brokered Loans		Brokered Loans	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Credit score	668.24	67.43	673.31	66.83
Note rate	7.50	1.56	7.32	1.78
Commission ($)	1363.92	1220.51	1406.00	1585.37
Loan amount ($1,000)	123.52	84.56	142.64	111.76
Loan To Value	67.17	28.23	64.58	28.53
Fixed	0.68	0.47	0.60	0.49
Second	0.22	0.41	0.23	0.42
Cash out	0.53	0.50	0.53	0.50
Distance	437.59	312.34	394.36	326.80
Loan term (months)	340.21	58.69	337.94	58.92
Purchase	0.19	0.39	0.20	0.40
Rate and term	0.25	0.43	0.25	0.43
30 year Mortgage	0.84	0.36	0.83	0.37
Observations	1931		2548	

Table 5.4: OLS Price Regression by Origination Type

	(1) Broker Price Equation	(2) Lender Price Equation	(3) Pooled Price Equation
Conforming	-0.51**	-0.05	-0.36**
	(5.7)	(0.51)	(5.30)
Credit score	-0.005**	-0.001**	-0.0034**
	(9.24)	(3.33)	(9.79)
Fixed	-0.87**	-1.29**	-1.06**
	(10.16)	(17.41)	(18.16)
HELOC	-1.71**	-2.73**	-1.83**
	(7.73)	(7.25)	(10.16)
Loan Amount	-0.003**	-0.005**	-0.004**
	(8.63)	(12.91)	(14.15)
Purchase	-0.087	-0.08	-0.07
	(0.96)	(0.97)	(1.08)
Rate and Term	-0.63**	-0.31**	-0.47**
	(7.49)	(4.18)	(8.20)
Term 10	-1.54**	-0.93*	-1.24**
	(3.31)	(2.15)	(3.90)
Term 15	-0.89*	-0.28	-0.62**
	(2.39)	(1.06)	(2.66)
Term 20	-1.01*	-0.55	-0.708**
	(2.69)	(1.89)	(2.94)
Term 25	-0.37	0.15	-0.28
	(0.81)	(0.35)	(0.94)
Term 30	-1.06**	-0.16	-0.65**
	(3.09)	(0.68)	(3.04)
Second Lien	1.84**	2.35**	2.13**
	(10.21)	(14.9)	(17.23)
Distance	0.06**	0.043*	0.058**
	(2.92)	(2.5)	(4.41)
Distance squared	-0.001	-0.00126	-0.002*
	(1.12)	(1.18)	(2.04)
Law Index	0.02*	-0.00185	0.009*
	(2.64)	(0.37)	(2.16)
Loan To value	0.011**	0.025**	0.017**
	(4.11)	(11.07)	(9.66)
Yield Curve	0.27*	-0.21	0.09
	(2.5)	(1.62)	(1.13)
Credit Risk	-0.18	0.85	0.20
	(0.24)	(1.3)	(0.39)
Interest Rate Volatility	.91*	1.33**	0.95**
	(2.57)	(3.86)	(3.85)
Constant	7.64**	3.07**	5.78**
	(8.67)	(4.11)	(9.75)

Table 5.4: Continued

non-brokered loans	-	-	-0.186**
			(3.49)
observations	2539	1925	4464
R-squared	.33	.44	.36

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table 5.5: Three Fold Oaxaca-Blinder Decomposition

Prediction (Mean Expected price)	Coefficient
Broker	3.445**
	(0.04)
Lending	3.409**
	(0.03)
Difference (D)	0.035
	0.06
Decomposition	
Endowments	-0.23**
	(0.056)
Coefficients	0.14**
	(0.057)
Interaction	0.11**
	(0.59)

Table 5.6: Two Fold Oaxaca-Blinder Decomposition

Prediction	Coefficient
Broker	3.445**
	(85.54)
Lending	3.409**
	(89.70)
Difference (D)	0.035
	0.52
Decomposition	
Explained	-0.15**
	(3.44)
Unexplained	0.185**
	(3.77)

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

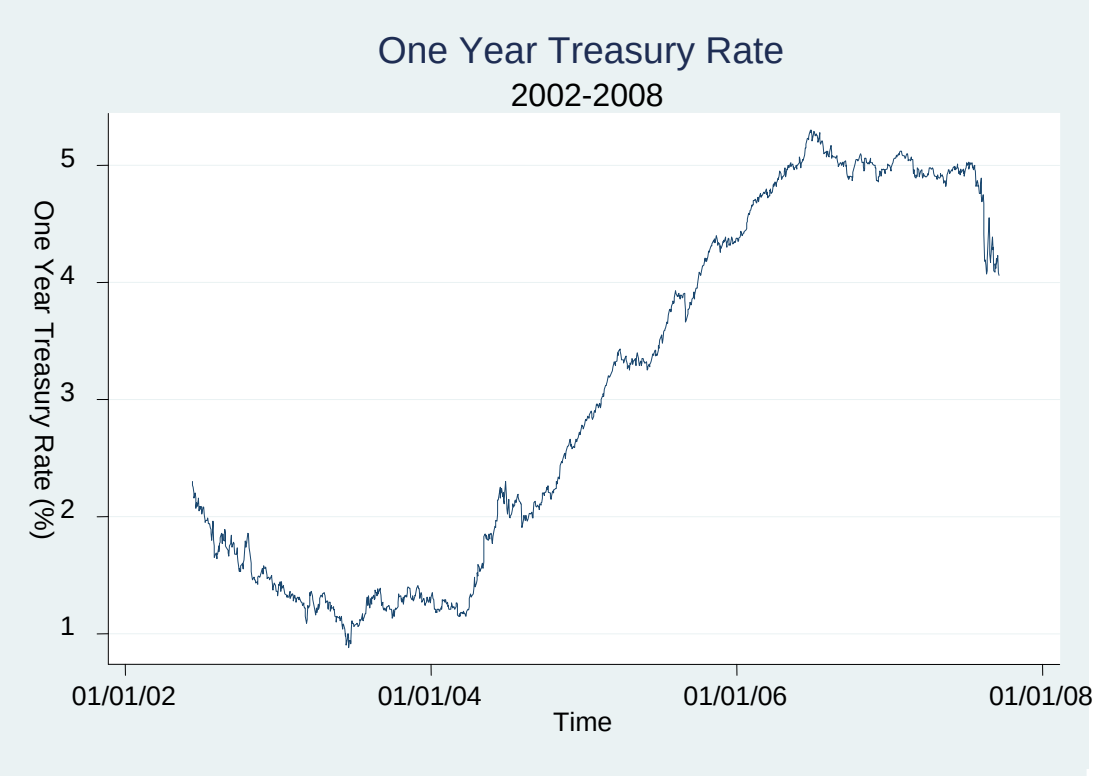

Figure 5.1 One Year Treasury Rate from 2002-2008

6. Conclusion

In this dissertation, I provide a detailed look at the mortgage industry and the role of informational frictions in determining the structure and performance of the mortgage market. In Chapter 2, I provide an analysis of the market structure of the mortgage industry. The institutional details and the complex structure of the mortgage market are discussed. I find that the configuration of the mortgage market at its present state is vastly different than its historical structure due to the reduction of information costs. The innovation of the internet and local area networks have reduced the cost of transmitting information and have increased the ability for firms to collaborate. As a result, the mortgage market has disintegrated into specialized firms, where each firm provides a single component and collaborates with the components of the supply chain to produce the final product. However, the vertical disintegration of the supply chain has introduced principal-agent problems. Asymmetric information concerns are mainly present in the broker segment of the mortgage market. The incentives and objectives of the broker are incompatible with the overall objective of the secondary market. Consequently, I find that more complete contracts are needed to realign the broker and market objectives.

Chapter 3, "A Model of Electronic Commerce in the Presence of a Gatekeeper," provides a theoretical explanation of the effect of the internet on market outcomes. Empirical studies examining the effect of the internet have provided inconsistent results. Search models assume that the reduction in search frictions would lead to more competitive markets. However, I argue that gatekeepers operating in online markets may create

an anticompetitive effect, in addition to reducing the consumers' search cost. Therefore, the conduct of the gatekeeper can cause prices in online markets to be higher than in retail markets and provide online firms with larger profits. When testing for the effect of the internet on market outcomes, it is important to consider the conduct of the gatekeeper in the market and test for both prices and profits.

In Chapter 4, "Joint Determination of Consumer Search Behavior and Mortgage Pricing," I examine the role of the Internet and Internet Comparison Search sites in reducing consumer search costs and their effects on the prices consumers pay for mortgages. Additionally, given the results from the third chapter, I expand the study to test for firm profits. Using a unique data set, I examine a mortgage firm's pricing strategies in online and retail markets, and find evidence of market power in online markets that does not exist in retail markets. To control for consumers' selection into market type, I use a switching regression and find that selection into online commerce is random and thus the estimation method is reduced to a pooled OLS. I find that online and retail consumers pay the same price on average for a mortgage. After controlling for variation in marginal cost, I find that the firm earns higher profits in online markets relative to the retail market. In addition to reducing the cost of search to the consumer, the availability of information in online markets allows firms to observe their competitors' pricing, and therefore reduces the firms' cost of monitoring. Both consumers and firms benefit from the availability of information, and consequently the presumed benefits to the consumer from the reduction of search costs are offset by the anticompetitive environment created by Internet Comparison Search sites in online markets. For some mortgage types, the anticompetitive effect dominates the search cost effect, which leads to higher prices in online markets.

In the final chapter, I analyze a mortgage firm's portfolio choice to test for the lemons problem in the mortgage market. Using the unique data set from the third chapter, I investigate the loan characteristics that affect the firm's decision to retain mortgages as part of its own portfolio. Mortgage lenders can use their own funds to finance a mortgage or alternatively can broker the loan using other lenders' financing. I estimate the quality differences between lender originated loans and brokered loans using an Oaxaca decomposition. I am able to decompose differences in prices between the broker and lender market into portions explained by the observed data and unexplained pricing which I attribute to asymmetric information. Since the Interest rate of a mortgage is the price of its underlying risk, then any variation in price is due to risk characteristics.

I find that non-brokered loans are on average priced 3 basis points lower than brokered loans. However the difference is statistically insignificant. By decomposing the difference, it is possible to disaggregate the difference in pricing into differences due to observable risk characteristics and difference due to unobserved characteristics. Non-brokered loans are priced on average 18 points lower than brokered loans because of unobserved qualities. Consequently, it is possible to assume that the lower unexplained price for non-brokered loans is in response to better unobserved characteristics of lender originated loans. Although this supports the presence of asymmetric information in the mortgage market, the unobserved difference cannot be attributed to asymmetric information with certainty. It is possible that the difference arises because of the lower cost of funds when the firm operates as a lender. Therefore, the lower prices in the lending market may be due to the firm passing on the reduction in cost to the consumer.

7. References

Akerlof, G. A. (1970). "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." The Quarterly Journal of Economics **84**(3): 488-500.

Alexander, W., S. Grimshaw, et al. (2002). "Some Loans Are More Equal than Others: Third Party Originations and Defaults in the Subprime Mortgage industry." Real Estate Economics **30**(4): 667-698.

Ambrose, B., M. LaCour-Little, et al. (2004). "The Effect of Conforming Loan Status on Mortgage Yield spreads: A Loan Level Analysis.`." Real Estate Economics **32**(4): 541-570.

Ambrose, B., M. LaCour-Little, et al. (2005). "Does Regulatory Capital Arbitrage, Reputation or Asymmetric Information Drive Securitization?" Journal of Financial Services Research **28**(1): 113-133.

Bailey, J. P. (1998). Intermediation and electronic markets : aggregation and pricing in Internet commerce, Massachusetts Institute of Technology.

Baye, M. R. and J. Morgan (2002). "Information gatekeepers and price discrimination on the internet." Economics Letters **76**(1): 47-51.

Belsky, E. and J. Prakken (2004). Housing Wealth Effects: Housing's Impact on Wealth Accumulation, Wealth Distribution and Consumer Spending, Harvard University. **W04**.

Blinder, A. S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates." The Journal of Human Resources **8**(4): 436-455.

Brown, J. R. and A. Goolsbee (2002). "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." The Journal of Political Economy **110**(3): 481-507.

Campbell, C., G. Ray, et al. (2005). "Search and Collusion in Electronic Markets." MANAGEMENT SCIENCE **51**(3): 497-507.

Campbell, J. Y. and J. F. Cocco (2003). "Household Risk Management and Optimal Mortgage Choice*." Quarterly Journal of Economics **118**(4): 1449-1494.

Carlin, B. I. (2009). "Strategic price complexity in retail financial markets." Journal of Financial Economics **91**(3): 278-287.

Carrillo, J. J. (2008). "Dangerous Loans: Consumer Challenges to Adjustable Rate Mortgages." Berkeley Business Law Journal, Vol. 5, p. 1, 2008.

Chen, Y., G. Iyer, et al. (2002). "Referral Infomediaries and Retail Competition." Review of Marketing Science **1**(2).

Chezum, B. and B. Wimmer (1997). "Roses or Lemons"Adverse Selection in the Market for Thoroughbred Yearlings." Review of Economics and Statistics **79**(3): 521-526.

Crawford, G. W. and E. Rosenblatt (1999). "Differences in the Cost of Mortgage Credit Implications for Discrimination." The Journal of Real Estate Finance and Economics **19**(2): 147-159.

Dell'Ariccia, G., D. Igan, et al. (2009). "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market " European Banking Center Discussion Paper No. 2009-14s.

Diamond, P. (1971). "A Model of Price Adjustment." Journal of Economic Theory **3**(2): 156-168.

Dietrich, J. (2005). "Under-specified Models and Detection of Discrimination: A Case Study of Mortgage Lending." The Journal of Real Estate Finance and Economics **31**(1): 83-105.

DiPasquale, D. and E. L. Glaeser (1999). "Incentives and Social Capital: Are Homeowners Better Citizens?" Journal of Urban Economics **45**(2): 354-384.

Downing, C., D. Jaffee, et al. (2005). "Information Asymmetries in the Mortgage Backed Securities Market." University of California at Berkeley **May**.

Ellison, G. and S. F. Ellison (2009). "Search, Obfuscation, and Price Elasticities on the Internet." Econometrica **77**(2): 427-452.

Follain, J. (1990). "Mortgage Choice." Journal of the American Real Estate and Urban Economics Association **18**(2): 125-144.

Follain, J. and P. M. Zorn (1990). " Unbundling of residential mortgage finance. ." Journal of Housing Research. **1**: 117-137.

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

Genesove, D. (1993). "Adverse Selection in the Wholesale Used Car Market." Journal of Political Economy **101**(Aug): 644-665.

Glaeser, E. L. and H. D. Kallal (1997). "Thin Markets, Asymmetric Information, and Mortgage-Backed Securities." Journal of Financial Intermediation **6**(1): 64-86.

Green, R. K. and M. J. White (1997). "Measuring the Benefits of Homeowning: Effects on Children." Journal of Urban Economics **41**(3): 441-461.

Greene, W. (2003). Econometric Analysis, Prentice Hall.

Heuson, A., W. Passmore, et al. (2001). "Credit Scoring and Mortgage Securitization: Implications for Mortgage Rates and Credit Availability." The Journal of Real Estate Finance and Economics **23**(3): 337-363.

Hitt, L. M. and F. X. Frei (2002). "Do Better Customers Utilize Electronic Distribution Channels? The Case of PC Banking." MANAGEMENT SCIENCE **48**(6): 732-748.

Hong, H. and M. Shum (2006). "Using Price Distributions to Estimate Search Costs." The RAND Journal of Economics **37**(2): 257-275.

Jacobides, M. G. (2003). How Do Markets Emerge? Orginizational Unbundling and Vertical Dis-integration in Mortgage Banking. Working Paper.

Jaffee, D. (2010). The Role of the GSEs and Housing Policy in the Financial Crisis. Working Paper

Jann, B. (2008). A Stata Implementation of the Blinder-Oaxaca Decomposition. Zurich, ETH, Chair of Sociology.

LaCour-Little, M. (2000). "The Evolving Role of Technology in Mortgage Finance." Journal of Housing Research **11**(2): 173-206.

LaCour-Little, M. and G. Chun (1999). "Third party originators and mortgage prepayment risk: An agency problem?" The Journal of Real Estate Research **17**(1): 55-71.

Lamalfa, T. (1998). Speech given at the annual convention of the National Association of Mortgage Brokers, as reported in Inside Mortgage Finance, June 26: 7-8. Mas-Colwell, A., M. D.

Lee, H.-G. (1998). "Do Electronic Marketplaces Lower the Price of Goods?" Communication **41**(1): 73-80.

Lee, L.-F. (1978). "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." International Economic Review **19**(2): 415-433.

McDonald, D. and D. Thornton (2008). "A Primer on the Mortgage Market and Mortgage Finance." Federal Reserve Bank of St. Louis Review **January/February 2008, 90**(1): 31-45.

Morgan, P. and R. Manning (1985). "Optimal Search." Econometrica **53**(4): 923-944.

Morton, F. S., F. Zettelmeyer, et al. (2005). "Cowboys or Cowards: Why Are Internet Car Prices Lower?" Discussion Paper, Haas School of Business, University of California, Berkeley.

Neumark, D. (1988). "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination." The Journal of Human Resources **23**(3): 279-295.

Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labor Markets." International Economic Review **14**(3): 693-709.

Order, R. V. (2000). "The U.S Mortgage Market: A Model of Dueling Charters." Journal of Housing Research **11**(2): 233-256.

Passmore, W., R. Sparks, et al. (2002). "GSEs, Mortgage Rates and the Long-Run Effects of Mortgage Securitization." Journal of Real Estate Finance and Economics **25**(2/3): 215-242.

Poole, W. (2007). "The GSEs: Where Do We Stand?" Federal Reserve Bank of St. Louis Review **May/June, 89**(3): 143-151.

Scitovsky, T. (1950). "Ignorance as a Source of Oligopoly Power." The American Economic Review **40**(2): 48-53.

Sorensen, A. T. (2000). "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs." The Journal of Political Economy **108**(4): 833-850.

Spader, J. and R. Quercia (2009). "Mortgage Brokers and the Refinancing Transaction: Evidence from CRA Borrowers." The Journal of Real Estate Finance and Economics.

Stahl, D. O. (1989). "Oligopolistic Pricing with Sequential Consumer Search." The American Economic Review **79**(4): 700-712.

Stigler, G. J. (1957). "Perfect Competition, Historically Contemplated." Journal of Political Economy **65**(1): 1.

Stigler, G. J. (1961). "The Economics of Information." Journal of Political Economy **69**(3): 213.

Stiglitz, J. E. (1979). "Equilibrium in Product Markets with Imperfect Information." The American Economic Review **69**(2): 339-345.

Viswanathan, S., J. Kuruzovich, et al. (2007). "Online Infomediaries and Price Discrimination: Evidence from the Automotive Retailing Sector." Journal of Marketing **71**(July): 89-107.

Williamson, O. E. (1981). "The Economics of Organization: The Transaction Cost Approach." The American Journal of Sociology **87**(3): 548-577.

8. Vita

ABDULLAH A. AL-BAHRANI

Date and Place of Birth
January 31, 1981, Muscat, Oman

Education
M.S. Economics, University of Kentucky, December 2007
M.A. Economics, American University, August 2003
B.S Business Economics, University of Louisville, 2002

Professional Experience
Teaching Assistant
University of Kentucky, Fall 2006-Spring 2009

Licensing Specialist
New Equity Financial Corporation, March 2005-November 2006

Mortgage Professional
New Equity Financial Corporation, December 2003-March 2005

Honors and Award
Kentucky Opportunity Fellow, 2009
Outstanding Teaching Assistant, 2010

Professional Affiliations
American Economic Association (AEA)
American Real Estate and Urban Economics Association (REUEA)
Middle East Economic Association (MEEA)
Southern Economic Association (SEA)

Abdullah A Al-Bahrani

Date