An Evaluation of the Effects of the Housing Bubble on Consumer Preferences in the Housing Market: A Hedonic Pricing Model

by

Robert Adam Blair

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 13,2010

Keywords: housing, consumer, preferences,
hedonic, economics

Approved by

Diane Hite, Chair, Professor of Agricultural Economics
Richard Beil, Associate Professor of Economics
Randolph Beard, Professor of Economics

Abstract

The purpose of this thesis is to examine the effects of the recent housing bubble on consumer preferences across the entire United States along with individual census regions. A hedonic pricing model is used to analyze how consumers valued certain aspects of a house then regressions are run on the model to test a theory developed to explain potential changes in consumer preferences in the housing market. Using data from 1997 and 2005, the results of the tests are examined to see how, if at all, consumers' preferences changed during the housing bubble and if these changes were universal across the whole United States or contained in specific regions. The regression results are then compared and any significant changes and variables are discussed and explored.

Acknowledgments

The author would like to thank his friends and family for their constant support and understanding. The author would also like to thank committee chair Dr. Diane Hite, along with committee members Dr. Richard Beil and Dr. Randolph Beard for their assistance, patience, and advice. Thanks must also be extended to Auburn University for providing an excellent learning environment in which to grow and succeed both academically and personally.

Table of Contents

List of Tables

List of Figures

I. Introduction

A house is typically the most expensive asset an individual will own during his or her life. But how exactly is a house valued, is it based on the aspects of the house itself, size, number of rooms, or is it based on the aspects of the area around it, neighborhoods, schools? Perhaps it is a combination of all of these things.

During the 1998-2006 housing bubble in the United States housing prices rose dramatically creating market values that were extremely inflated. This was brought about by many things but was made possible by the federal government's encouragement to grant sub-prime mortgages by deregulating the mortgage market. The question this study attempts to answer is: during the housing bubble did consumers' preferences for different housing characteristics change? Consumer preferences, as defined in this study, are the average consumer preferences of buyer in the United States. The hypothesis of this study is that preferences did in fact change over time. The data used in this study is 1997 and 2005 survey data collected by the United States Department of Commerce: Bureau of the Census, as part of the American Housing Survey. These time periods are directly before the housing market bubble and at its peak. Using data from these two years allows for an excellent comparison between consumer preferences before and during the housing bubble. To gain an insight as to how, if at all, consumer preferences changed during the bubble an analysis on a hedonic pricing model for a house's market value is run. The hedonic model is used to determine the implicit values of different characteristics of a house because houses are heterogeneous goods thus the values must be assessed on the characteristics of the houses. An Ordinary Least Squares regression will be

run on the model and then tested for specification errors. The results of these tests will then be compared and discussed.

There are seven sections in this study, the first being the introduction. This section simply describes the question to be answered in this study and the means by which the answer will be found. The second section is an examination into the causes of the housing bubble. A deeper look into what may have caused this economic changed may offer a better insight into and understanding of the possible effects the bubble may have had or will have still. In section three, literature previously written on subjects pertinent to this study will be reviewed and discussed. These subjects include consumer preferences and hedonic pricing models. By reviewing these works, a better understanding of these subjects and the origins of the theories used in this study can be obtained. The fourth section of this study explains the theoretical model and the variables used. This section will discuss why the variables used were chosen and how they pertain to the question this study aims to answer. Section five discusses the data and methodology of testing this data. Explanations of the data, its origins, and usefulness to this study are given and descriptions of the methods used to test the data along with reasons these methods were chosen are also given in this section. The sixth section contains the results of the multiple tests run on the model developed in section four using the data described in section five. The seventh and final section is a simple overview of the work done in this study and offers conclusions based on its findings. The evidence as to whether consumer preferences changed during the housing bubble is discussed in this final section.

II. Evaluation of the Housing Bubble

From 1998 to 2006 the United States housing market experienced an extremely uncharacteristic rise in housing prices. As home prices rose past their predicted levels, most believed this was not a bubble but in fact a boom. It was not until the bubble "popped" that the problem of a housing bubble became obviously apparent. Typically when bubbles are formed in market economies they are not detected until they burst, however, there were some economists that made claims that the US housing market was experiencing a bubble as early as 2003. Most of these claims were disputed and the majority opinion on the subject remained that there was no housing bubble. When it became apparent that there was in fact a bubble economists started searching for answers to explain what caused this problem.

To begin to understand the causes of the housing bubble the government sponsored enterprises[1] of Fannie Mae and Freddie Mac must first be explained. The U.S. government established the Federal National Mortgage Association (Fannie Mae) in 1938 and Congress later chartered it in 1968 as a private shareholder-owned company. Fannie Mae operates in the U.S. secondary mortgage market by working with mortgage market partners to insure they have funds to lend to homebuyers at affordable rates. Fannie Mae is able to fund its mortgage investments by issuing debt securities in domestic and international capital markets. The Federal Home Loan Mortgage Company (Freddie Mac) was chartered by Congress in 1970 to compete with Fannie Mae. Freddie Mac was created to purchase mortgages on the U.S. secondary mortgage market. Freddie Mac was then to pool these mortgages and sell them as mortgage-backed securities on

[1] Government sponsored enterprises are financial services companies created by the U.S. Congress to increase the flow of credit to specific sectors, agriculture, home finance, and education.

the open market. The goal of Freddie Mac and Fannie Mae is to expand the secondary mortgage market by increasing the supply of money available for mortgage lending and for new home purchases. These two companies would come to play a large role in the creation of the housing bubble.

In 1977 the Community Reinvestment Act (CRA) was passed in order to encourage depository institutions to meet the credit needs of the entire community in which they operate. The CRA mainly focused on preventing these depository institutions from partaking in the act of redlining or, refusing, increasing the cost of, or limiting services such as loans, mortgages, and insurance within specific geographic areas, especially inner-city neighborhoods. In 1995 Fannie Mae started to receive affordable housing credit for buying subprime securities. The Taxpayer Relief Act of 1997 encouraged consumers to purchase second homes and investment properties by reducing taxes on income gained by selling a house. Also during this year Fannie Mae helped launch the first CRA securities available to the public issuing $384.6 million of such securities[2], all of which carried a guarantee to timely interest and principal given by Fannie Mae. The mortgage denial rate was 29% in 1997[3].

In September of 1999 Fannie Mae decreased the credit requirements in an attempt to encourage banks to offer home mortgage loans to individuals whose credit was not good enough to qualify for conventional loans. Due to this ease in requirements banks began to increase the rate at which they issued subprime mortgages. Also in 1999 the Gramm-Leach-Bliley Act deregulated banking, insurance, and securities and as a result allowed financial institutions to become very large. By 2000 Fannie Mae had committed to purchase and securitize $2 billion of CRA-eligible loans and announced that the Department of Housing and Urban Development

[2] American Bankers Association.
[3] Statistic found at www.ffiec.gov

would soon require that 50% of its business be dedicated to low and moderate-income families. Fannie Mae also announced that its goal was to finance over $500 billion in CRA-related business by 2010. The U.S. Federal Reserve lowered the Federal Funds Rate 11 times from 6.5% to 1.75% in 2001 then down to 1% in 2003, the lowest it had been in 45 years[4]. Also in 2003 Fannie Mae and Freddie Mac bought $81 billion in subprime securities[5]. From 2002 to 2003 the mortgage denial rate fell to 14%, half of the 1997 figure[6].

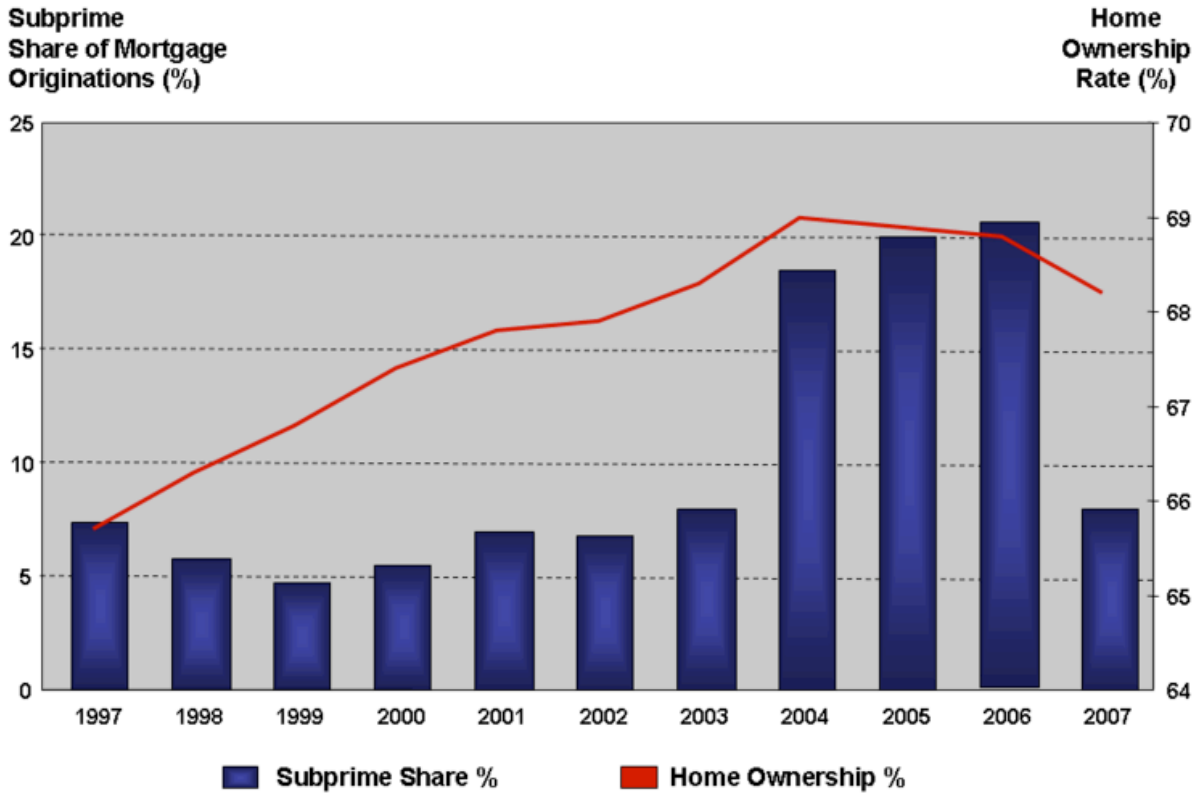
Due to Fannie Mae and Freddie Mac's willingness to purchase subprime mortgages from the banks and institutions lending them and the reduced regulations on these lending institutions the requirements for approval on such loans fell dramatically causing the decrease in the denial rate for mortgage applications. This led to, in 2004, the U.S. having the highest rate of homeownership in its history at 69.2%[7]. Encouraged by the Department of Housing and Unban Development, both Fannie Mae and Freddie Mac together purchased $434 billion in securities backed by these subprime loans. When the Securities and Exchange Commission loosened the rules for five of the major lending firms[8] and allowed them to ignore the government-imposed limits on how much debt they can assume, they quickly increased their debt by making subprime loans. This added additional pressure on Fannie Mae and Freddie Mac, which increased their risky lending. By the end of the year 2004 the U.S. housing market was set up perfectly to expand the housing bubble until it bursts.

Over the next few years the high availability of subprime loans, caused by relaxed regulations on lending firms and the willingness to purchase these loans by Fannie Mae and

[4] Statistic from the U.S. Federal Reserve Website
[5] Statistic from "How HUD Mortgage Policy Fed the Crisis"
[6] Statistic from the Federal Financial Institutions Examination Council
[7] Statistic from the U.S. Census Bureau
[8] Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch, and Morgan Stanley

Freddie Mac, encouraged homeowners to refinance their homes and/or purchase additional houses with low interest rates.

Figure 1: U.S. Subprime Lending Expansion 2004-2006[9]

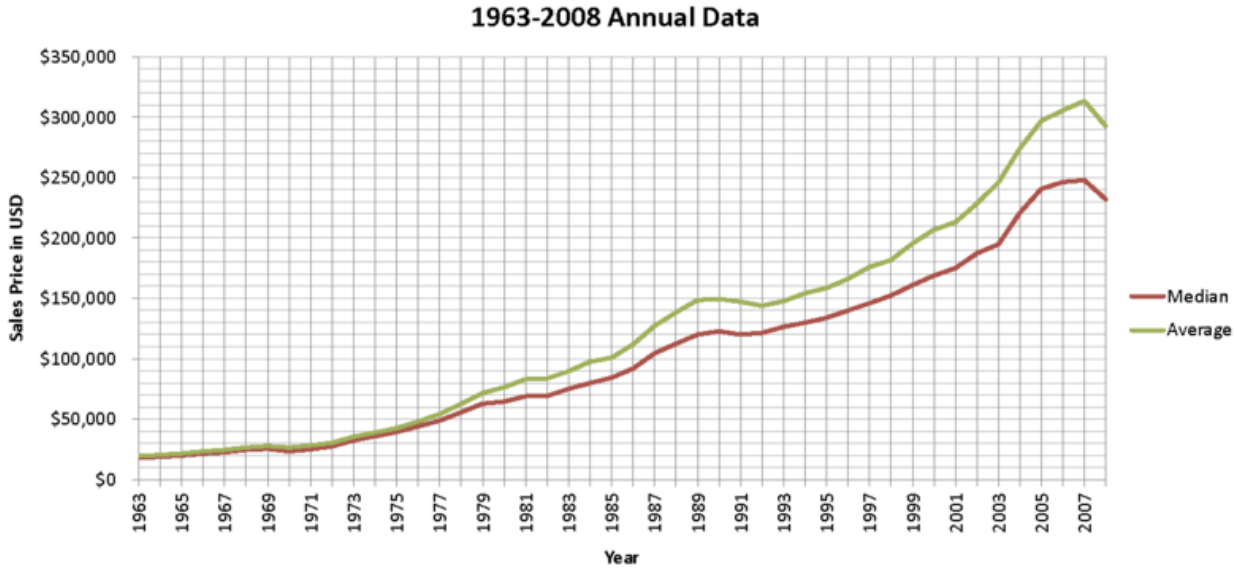


Many of the subprime mortgages in the US were adjustable rate mortgages, these loans have a fixed interest rate for the first few years but then the rate is adjust once or twice a year based on the market interest rate. These subprime mortgages with low adjustable rates also allowed individuals, who would ordinarily not be able to acquire a home loan, purchase their own house. However, when the Federal Reserve began in 2004 to raise the Federal Funds Rate, from 1.25% to 5.25% in 2006[10], the adjustable rates on the subprime mortgages began to adjust. This made it impossible for many individuals to make the payments on the mortgages they took

[9] Data Source: U.S. Census Bureau and Harvard Report- State of the Nation's Housing 2008 Report
[10] Statistic from U.S Federal reserve Website

out. The fact that from 1997 to 2005 mortgage fraud in the U.S. rose by 1,411%[11], and an increase in the number of individuals that defaulted on these subprime mortgages were the final straws that caused the housing bubble to burst. It is important to note that the majority of individuals defaulting on mortgages were on the investment side of the market, trying to "flip" houses.

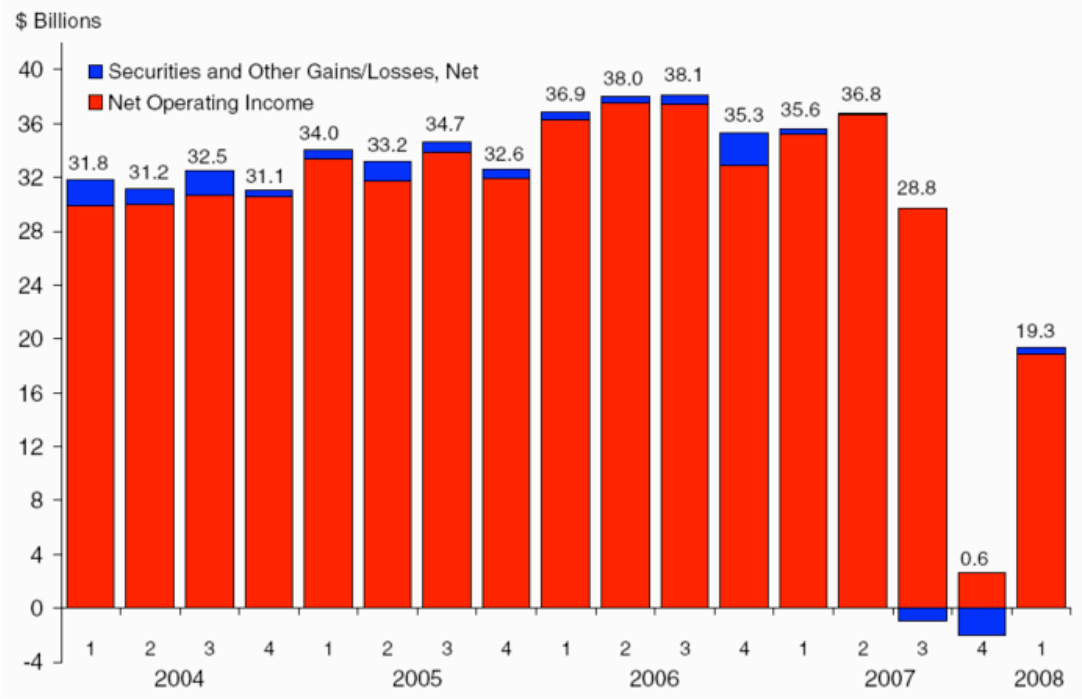
Figure 2: Median and Average Sales Prices of New Homes Sold in the U.S.[12]



In 2008 the national median price of new homes fell nearly 6.4% to $237,100 from a peak of $247,900 in 2007 and by 2009 had fallen over 12.5% to $216,700.By the end of 2007 there were 2,203,295 foreclosures in the U.S. up from 885,468 in 2005[13]. Finally in 2007, more than 25 subprime lenders including New Century Financial, the largest subprime lender in the U.S. declared bankruptcy, reported substantial losses, or were put up for sale because they were unable to cover the debt they had incurred from subprime lending.

[11] Statistic from the U.S. Department of the Treasury, Financial Crime Enforcement Network
[12] Data Source: U.S. Census Bureau
[13] Statistics from RealtyTrac Year End Report

Figure 3: Quarterly U.S. Bank Earnings 2004-2008[14]



The declared bankruptcy of major banks, along with the continued decrease in housing prices lead to the financial crisis and recession in the U.S. economy that began in 2007. Although legislation and regulatory changes put in place by the federal government did not cause the U.S. housing bubble, they did in fact open the door and make it possible for the banks and lending institutions to participate in the actions of writing subprime mortgages, bundling these mortgages into securities, and then selling off these securities and the risk associated with them.

[14] Data Source: FDIC Quarterly Banking Profile

III. Literature Review

Hedonic price models have been used to analyze housing markets for years. Hedonic models came about when the need to find the demand for heterogeneous goods arose. There are two primary applications of the hedonic price model, the first stage model, which consists of estimating marginal implicit prices of house characteristics, and two-stage models, which are used to identify demand curves for the characteristics. Both will be discussed here, but the application used in this study relies on first stage modeling, which avoids some of the potential pitfalls of two-stage models. The demand for heterogeneous goods cannot be explained by the price of the good due to the fact that these goods are, for the most part, "one of a kind". The idea of demand for characteristics was set forth in Rosen's (1974) paper that presented an analysis on hedonic prices in a perfectly competitive market and a model that displayed how hedonic prices influences buyer and seller choices in implicit markets. This was the first introduction of the two-stage model. Rosen described these prices as follows:

> "Hedonic prices are defined as the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them."[15]

Although Rosen's work is very famous and an important starting point when reviewing hedonic pricing models, his was an extension of the literature on first stage models by Griliches (1961) and Griliches (1971). The introduction of dealing with heterogeneous goods through hedonic analysis in Griliches' works allowed the ideas to reach a wide range of economists. Though his research was important to the expansion of hedonic analysis, Griliches cannot lay

[15] "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" pg. 34

claim to introducing the idea of hedonic prices to the world of economics. Waugh (1929) wrote about how the "qualities" of goods affect the price of these goods. By looking at the quality of vegetables and attributing this to a range of observable characteristics Waugh was able to estimate the implicit prices of these characteristics. This was the first example of hedonic analysis, although Waugh did not use this term. Ten years later, Court (1939), was the first to use the term "hedonic" in describing his work on implicit prices and the demand of the attributes that make up goods characterized as heterogeneous. The combined works of these economists show how important the development and introduction of hedonic analysis was, and still is, to the advancement of economics and econometrics as a whole. They lead the way in moving hedonic price analysis from a new, advanced practice to the standard in examining heterogeneous goods and markets. These studies constitute so-called first stage models, which determine the implicit prices for characteristics, and which are in widespread use in the current literature and employed by many urban and real estate economists.

Goodman (1978) performed a hedonic analysis on house price indices using a short-run equilibrium model. Goodman constructed 15 submarkets from data collected in one metropolitan area and compared the hedonic price indices from each submarket. Through this experiment Goodman found that after hedonic price coefficients were aggregated into standardized units it was evident that prices were higher in the central city than those in the suburbs. Applying hedonic price analysis to the separate submarkets a difference in the value of structural improvements, which were valued higher in the suburbs, was measured. Using this same technique it became apparent that the value of neighborhood improvements was constant though out the metropolitan area. Goodman concluded that hedonic price analysis could be

applied to separate submarkets of a larger metropolitan area to measure the differences in prices among these submarkets.

Hedonic pricing models can occasionally suffer from identification problems. However, Palmquist (1984) found that using data from multiple cities could reduce these problems especially when they are used to estimate demand curves. Rosen stated that the problem with hedonic estimation was the interaction of demand and supply but Palmquist suggested a model that ignored the supply side of the market correctly results in prices that are endogenous when a non-linear hedonic equation was used. Also refuting Rosen, Bartik (1987) found that the problem with estimating demand parameters hedonically is caused by the endogeneity of both prices and quantities when households face a non-linear budget constraint. Bartik offered the solution of using instrumental variables that would exogenously shift the budget constraint and thus the hedonic pricing function.

Now that some of the troubles that arise when using hedonic price functions to predict demand parameters began to come to light, many economists started investigating which functional form in hedonic prediction eliminated or reduced these problems. Cropper, Deck, and McConnell (1988) tried to determine the best choice when it comes to the functional form of hedonic price functions and found that in general the simple forms such as, linear, semi-log, double-log, and linear Box-Cox, worked the best for reducing bias in a hedonic prediction in which there are omitted variables. The linear and Box-Cox preformed the best when misspecification was present, the linear functional form produced the smallest maximum bias while the Box-Cox form had the lowest average bias. From these findings, the linear form or linear Box-Cox form were suggested when estimating hedonic price functions.

Goodman and Thibodeau (1995) discovered another problem with hedonic estimation on house prices. Possibly due to depreciation and vintage effects, along with the effects of demand for construction and renovation, house age, as a parameter in hedonic pricing models, actually has a heteroskedastic effect. This is important because most hedonic house price models include age as a parameter. After comparing derived submarkets in Dallas to imposed submarkets like zip codes and census tracts, Goodman and Thibodeau (2003) found that smaller submarkets improved hedonic price prediction accuracy and due to this discovery, developed a method of deriving submarkets to improve prediction accuracy.

Hedonic pricing models have helped advance demand estimation techniques along with developing new theories of demand based on the demand for heterogeneous goods. A significant aspect of demand theory is that of consumer preferences. Demands are set based on these preferences, which are set by individuals' utility functions. Utility is the "happiness" an individual receives from something, utility theory suggests that an individual will make a decision that maximizes his or her utility based on the current circumstances surrounding that decision. Utility cannot be directly measured and is revealed in the choices that an individual makes. If there are two cars that are exactly the same except for color, one blue and one red, and someone chooses the blue car then it is said that this "reveals" their preference showing that a blue car brings them more utility than a red car. The concept of revealed preference was first introduced by Samuelson (1938). Marshal (1920) described how an individual's utility is revealed in prices,

> "Utility is taken to be correlative to Desire or Want. It has been already argued that desires cannot be measured directly, but only indirectly, by the outward phenomena to which they give rise: and that in those cases with which economics is chiefly concerned the measure is found in the price which a person is willing to pay for the fulfillment or satisfaction of his desire."[16]

[16] Principles of Economics: An Introductory Volume pg. 78

From this it is easy to see how consumer preferences bridge the gap between utility and demand. Due to the direct immeasurability of utility, consumer preferences act as a measurable observation and allow demands to be quantitatively measured. Further proof of this lies in a quote from Basmann (1956) in which is described the definition of a change in consumer preference by two earlier works of Ichimmra (1950) and Tintner (1952).

> "Ichimura and Tintner defined a change in preferences by a change in the form of the ordinal utility function or indifference map and derived, for shifts in demand, algebraic expressions which are linear combinations of the Slutsky-Hicks substitution terms which play a central role in existing consumer demand theory."[17]

Current research into consumer preferences has started to use hedonic pricing models as is such in a work by Lancaster (1966) which states,

> "The crucial assumption in making this application has been the assumption that goods possess, or give rise to, multiple characteristics in fixed proportions and that it is these characteristics, not goods themselves, on which the consumer's preferences are exercised."[18]

Goods having multiple characteristics on which consumers exercise their preferences mirrors the ideas behind hedonic pricing models used to estimate demands for heterogeneous goods. Lancaster seems to suggest there is a possibility that homogeneous goods do not exist, that consumers do not choose between two different pairs of shoes based on the price of the shoes but, instead choose based on the "prices" or perceived values of the shoes components, characteristics, and attributes.

[17] "A Theory of Demand with Variable Consumer Preferences" pg. 48
[18] "A new Approach to Consumer Theory" pg. 154

IV. Theoretical Model

The model used in this study is a hedonic pricing model developed by Rosen as a method of determining the implicit or hedonic prices of the characteristics that make up a commodity. This model is being used in this study based on the heterogeneity of houses. Rosen's model beings with the assumption that a good has *n* number of characteristics and that this good, good *z,* is composed of different amounts of each characteristic, $z = (z_1, z_2, \dots, z_n)$ where z_i is the amount of the *i*th characteristic in each good. The price of the good is dependent on z_i, $p(z) = p(z_1, z_2, \dots, z_n)$. In this model the assumption is made that both consumers and producers base their decisions on maximizing behavior and that the market clearing price $p(z)$, is determined by consumer preferences and producer costs. Also an assumption of indivisibility is made, in other words saying that two four bedroom house are not the same as an eight bedroom house nor is living in an eight bedroom house half the year and a four bedroom house the other half the same as living in a six bedroom house all year. This also includes that assumption that sellers do not repackage goods, either for lack of ability or because it is not profitable to do. The utility function for this model can be written as $U(x, z_1, z_2, \dots, z_n)$ and is assumed to be strictly concave along with the other usual properties, where *x* is all other goods consumed. Income, *y*, is measured in terms of *x*: $y = x + p(z)$. To maximize utility over *x* and $(z_1, \dots, z_n)$ must be chosen to satisfy the budget constraint, $y \geq x + p(z)$ where *x* is a numeraire good with a price of one, and the first-order necessary conditions

$\partial p/\partial z_i = p_i = U_{z_i}/U_x, i = 1, \dots, n$. The value function is then defined as

$\theta(z_1, \dots, z_n;\ u, y)$ according to

$$U(y-\theta, z_1, \ldots, z_n) = u \tag{1}$$

$\theta(z; u, y)$ represents the amount a consumer is willing to pay for alternative values of $(z_1, \ldots, z_n)$. This allows for z_i to be related to "money", and has been used often in urban economics. By differentiating (1), the following can be obtained

$$\theta_{z_i} = U_{z_i}/U_x > 0, \theta_u = -1/U_x < 0 \qquad \theta_y = 1 \tag{2}$$

$$\theta_{z_i z_i} = (U_x^2 U_{z_i z_i} - 2U_x U_{z_i} U_{x z_i} + U_{z_i}^2 U_{xx})/U_x^3 < 0 \tag{3}$$

where the inequality in (3) is due to the assumption about the bordered Hessian matrix of *U*. Also the assumption of strict concavity for U implies that θ is concave in z. Equations (2) and (3) show that in fact, θ is increasing in z_i at a decreasing rate. In other words, θ is the marginal rate of substitution between z_i at a given utility index and income.

It has been shown that $\theta(z; u, y)$ is the amount a consumer is willing to pay and $p(z)$ is the minimum price the consumer must pay in the market. Based on this, utility is maximized when $\theta(z; u, y) = p(z)$ *and* $\theta_{z_i}(z^*; u^*, y) = p_i(z^*), i = 1, \ldots, n$, where z^* and u^* are optimum quantities. After differentiating θ_{z_i} with respect to u, it can be seen that $\theta_{z_i u} = (U_x U_{u z_i} - U_{z_i} U_{xx})/U_x^2$. The numerator of this equation determines the sign of the income elasticity for the characteristic or "good" z_i, ceteris paribus.[19] If all of these derivatives are positive, then z_i is normal and additional income always increases attainable utility. Thus, if $p(z)$ is convex it might be expected that consumers with higher incomes would purchase larger amounts of all of the characteristics. This reveals a consequence of this model in that it displays natural tendencies toward market segmentation. However, this is a common result in spatial equilibrium model, as shown in Tiebout's (1956) analysis of the implicit market for neighborhoods, using the local public goods as "characteristics" in this case. The results of his

[19] Latin phrase commonly used in economics meaning "holding other things constant"

analysis showed that neighborhoods tend to be segmented by distinct income and preference groups. This result has shown to hold true for other differentiated products as well.

A parameter for consumer preferences can be taken into account in this model which results in a utility which can be written as $U(x_1, z_1, \ldots, z_n;\ \alpha)$, where α differs from person to person. The equilibrium value functions are dependent on income and consumer preferences and the market hedonic price model is the envelope of the group of value functions that characterize the equilibrium of all consumers, represented by the joint distribution function, $F(y, \alpha)$, as given in the whole population.

V. Variables

The first stage empirical model is employed in this study, including variables that reflect house quality and that of the neighborhood in which they are located. These variables are commonly used in hedonic pricing models used to analyze the housing market. The dependent variable in this study is the log of the value[20] of the house (LOGVALUE) and is assumed to be a function of the remaining independent variables used in this study. The reason for using the log of the market value rather than the level value is that prices are generally considered to be log normally distributed and the log function can help reduce heteroskedasticity. There is also a chance that using the linear value may result in negative predicted values, which obviously would be incorrect. The variable of the market value in this study is the homeowner's estimation of the house's value. This has been shown to have a large bias in individual studies but when the sample size is large enough then this bias is greatly reduced, Kain and Quigley (1972). Goodman and Ittner (1992), by using data from 1985 and 1987 that homeowners tend to overestimate the value of their house by about 6%. Since the actual price of each house was not included in the data, and because the homeowner estimated market value is fairly accurate in large samples, the data provided for the market value of the house by the Census Bureau was used in this study. The variables that were chosen for this study were selected based on the fact that they are, in large part, what most real estate companies used to describe the properties they are selling and their appearance in prior analyses of the housing market.

[20] Current Market Value

The first independent variable used is the annual real estate tax payment divided by one thousand (AMTXK). This variable was set in thousands to reduce a problem with scaling. The amount of real estate taxes a consumer will have to pay annually could be expected to show a negative relationship with LOGVALUE under the assumption that paying taxes is undesirable and as the taxes on a house increase, the price of the house should fall so that consumers can compensate the amount they pay in taxes with a lower price on the house. However, in this study a positive relationship is expected between AMTXK and LOGVALUE due to the idea the taxes could be based on amenities. This would make the desirable part of higher taxes, the amenities for which the taxes are paid.

The second variable included in this study is the number of bedrooms the house has (BEDRMS). This variable has shown to be a significant characteristic of a house in previous studies. A bedroom is defined as a room in the house not specifically designated as any kind of living room, dining room, hall, foyer, or bathroom. Typically this variable is used by realtors in the description of a house when advertizing on the market. This, along with the evidence from other studies, shows that this variable is extremely important to consumers when purchasing a home and leads to the expectation of a significant positive relationship with the dependent variable.

Another variable that has proved to be significant in many previous studies is the age of the house (AGE). This variable was created by subtracting the year the house was built from the year in which the data was collected, providing that age of the house in years. Although they have shown it to be significant, previous studies have also shown age to cause heteroskedasticity[21]. As stated earlier Goodman and Thibodeau (1995) discovered that age did in

[21] OLS makes the assumption that the variance of the error term is constant (Homoscedasticity). If the error terms do not have constant variance, they are said to be heteroscedastic.

fact cause heteroskedasticity in hedonic pricing models. This variable may also be somewhat skewed in the second year of data due to the housing bubble causing a large influx of new homes in the market. Ideas on how to adjust for these potential problems are discussed in later sections. Once again, as Goodman and Thibodeau (1995) pointed out, age in some cases may be valued due to the vintage effect, however one would expect a negative relationship with LOGVALUE due to the effect of depreciation seeming more common in the market.

Another variable chosen on the fact that it is included in real estate listings is a dummy variable describing whether or not the house has a garage. Although the size of the garage (one car, two car, etc.) is not included, it is believed that this variable will have a significantly positive relationship with the log of value of a home.

The fifth variable used in this study is the first to attempt to describe the neighborhood in which the house is located (HOWN). This variable is composed of the survey results for how the homeowner rated their neighborhood as a place to live on a ten point scale, ten being the highest rating and one being the lowest. This variable is also expected to have a significant positive relationship with the dependent variable. This expectation is based on the idea that houses in highly rated neighborhoods should have a greater demand than houses in neighborhoods with lower ratings, ceteris paribus. And due to the law of demand this should drive the prices of the houses in neighborhoods with lower ratings down, all else held constant.

Another variable used in this study that may cause the specification error of heteroskedasticity is that of lot size (LOT10K). The lot size is the measure of the area of land on which the house is built and is sold along with the house such as a lawn. In this study the square footage of the lot was divided by 10,000 in order to, once again, correct a scaling problem. The reason this could result in heteroskedasticity is that fact that the lot size of a house indirectly

determines the size or square footage of the house, another variable used in this study.. This is because if a house has a lot size of 10,000 ft^2, then the house built on this lot would have to have a square footage measurement of 10,000 ft^2 or less. The effect of this variable on the value of a house cannot be predicted due to the fact that some consumers may prefer a larger lawn thus a large lot size while others, who do not want to deal with the upkeep of a large lawn, may prefer a smaller lot size.

The second variable that deals with the quality of the neighborhood the house is in is (SATPOL) which is a dummy variable stating whether or not the homeowner is satisfied with the police protection and presence in the area. This variable is expected to have a positive relationship with the dependent variable. The reasoning behind this that if a home owner is dissatisfied with the police in the area, then the area may possibly have a high crime rate and the homeowner sees this as a direct effect of the police not doing a good enough job to prevent the crime. Previous studies have shown crime rate to have a negative effect on the value of a house, so if SATPOL is 0 when the crime rate is high and 1 when it is low then the opposite effect of crime rate is expected.

With the bedrooms and bathrooms being the only rooms accounted for in this study, the need for another variable arose. The number of square feet of living space in the house (UNITSFH) was used to account for the other rooms in the house not designated bedrooms or bathrooms. This variable is the square footage of the house divided by one hundred to correct for scaling. This variable will help in assessing how consumers value the other rooms, as a whole, in a house. This variable is also widely used by realtors and in previous studies. Although a positive effect on value is expected, a problem of heteroskedasticity could arise with this variable as it is linked to multiple variables included in this study.

The next variable analyzed in this study is that of the household income of the owners of the house (ZINC2K). This variable has also been divided by 1000 to adjust for scaling. Although the income of the family buying a house may not change the market value of a house directly, but it does determine the consumer's budget constraint and is also examined when applying for a home mortgage loan. Both of which directly influence the price range a consumer can actually purchase a house. Due to this influence, a positive relationship with the dependent variable is expected. This is based on the fact that the model used in this study leans towards market segmentation, that consumers with higher income should purchase more expensive houses. Another reason for using the household income is that, as discussed in Ball (1973), the use of aggregate data has become less popular in hedonic models due in part to both aggregate bias, and the increasing availability of household level data.

The final variable used in this study is the total number of bathrooms in the house (TOTBATH). This variable has shown, like BEDRMS, to be a significant aspect of a house to consumers. This variable was created by combining two separate variables included in the data source. These two variables were the number of full bathrooms and the number of half bathrooms. Full bathrooms are defined as bathrooms that include a sink, toilet, and bathtub or shower whereas half bathrooms are defined as bathrooms not having the bathtub or shower aspect. Based on results from previous studies the expected relationship to LOGVALUE is positive. This should make sense once again based on the fact that this variable is a key piece of information included by realtors on the listing of a house.

The variables describing the quality of the neighborhood combined with those that describe the characteristics of the house itself should provide a strong model to explain the value of house. This model may help in analyzing whether or not average consumer preferences for

houses changed during the housing bubble. The descriptive statistics for all of these variables could provide an idea as to the type of results that will be produced from the analysis of this model. These descriptive statistics for both 1997 and 2005 are listed in the following tables.

TABLE 1.1: 1997 Variable Descriptive Statistics

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	1.553 (-1.575)	2.675 (1.764)	1.5023 (1.310)	1.129 (1.549)	1.423 (1.299)
BEDRMS	3.112 (-0.853)	3.195 (0.938)	3.110 (0.871)	3.056 (0.771)	3.147 (0.888)
AGE	36.877 (-22.762)	44.020 (23.763)	42.090 (23.729)	31.827 (20.674)	32.576 (20.804)
GARAGE	0.771 (0.420)	0.724 (0.447)	0.846 (0.361)	0.681 (0.466)	0.873 (0.333)
HOWN	8.208 (1.798)	8.365 (1.632)	8.232 (1.740)	8.191 (1.883)	8.066 (1.847)
LOT10K	7.415 (19.071)	7.220 (17.862)	8.481 (21.076)	8.081 (19.776)	4.849 (15.169)
SATPOL	0.918 (0.274)	0.939 (0.238)	0.932 (0.252)	0.901 (0.298)	0.913 (0.282)
UNITSFH	20.069 (9.386)	23.094 (10.194)	20.949 (9.546)	18.648 (8.953)	18.797 (8.375)
ZINC2K	52.542 (39.938)	59.096 (44.157)	51.738 (37.179)	47.499 (37.665)	57.240 (42.456)
TOTBATH	1.989 (0.856)	1.941 (0.901)	1.894 (0.851)	2.010 (0.851)	2.125 (0.813)
VALUE	109981.56 (70893.20)	129770.62 (74004.84)	99092.91 (63801)	92435.46 (62023.05)	148380.11 (91648.36)
Number of Observations	19589	3354	5375	7050	3810

Note: The mean value of the variable for each region is listed with the standard deviation in parentheses below.

TABLE 1.2: 2005 Variable Descriptive Statistics

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	2.333 (2.928)	3.915 (3.679)	2.193 (2.469)	1.651 (2.621)	2.341 (2.733)
BEDRMS	3.160 (0.840)	3.218 (0.879)	3.129 (0.827)	3.123 (0.786)	3.218 (0.908)
AGE	43.732 (23.4190	53.139 (24.043)	48.433 (24.237)	37.627 (21.286)	40.116 (21.326)
GARAGE	0.790 (0.407)	0.717 (0.450)	0.875 (0.331)	0.706 (0.456)	0.890 (0.313)
HOWN	8.223 (1.673)	8.327 (1.598)	8.223 (1.651)	8.197 (1.762)	8.176 (1.600)
LOT10K	8.108 (20.219)	7.820 (19.464)	9.978 (23.318)	86.954 (204.551)	48.659 (150.063)
SATPOL	0.921 (0.270)	0.930 (0.255)	0.930 (0.254)	0.907 (0.290)	0.923 (0.266)
UNITSFH	21.681 (17.392)	24.422 (20.994)	21.598 (16.828)	21.026 (16.735)	20.555 (15.438)
ZINC2K	71.304 (72.718)	82.879 (84.069)	66.328 (62.662)	64.917 (69.711)	79.011 (77.478)
TOTBATH	2.101 (0.870)	2.066 (0.898)	2.008 (0.850)	2.131 (0.887)	2.198 (0.829)
VALUE	168437.48 (249349.21)	206266.56 (223622.77)	125234.5 (137845.76)	125492.96 (198472.23)	304304.01 (419692.51)
Number of Observations	18648	3287	4949	6646	3766

Note: The mean value of the variable for each region is listed with the standard deviation in parentheses below.

VI. Empirical Data and Methodology

The U.S. Census Bureau collected the data used in this study as a part of the American Housing Survey. The Census Bureau conducts this survey every year but only the years 1997 and 2005 were used in this study. These years were chosen based upon their relation, in time, to the housing bubble. Due to the fact that the housing bubble started in 1998 and continued until 2006 the data from 1997 are used as the base year data to which the data from 2005 are compared to determine if consumer preferences for houses changed during the housing bubble.

To avoid any specification error or bias resulting from the fact that the market values in the 2005 data set will be inflated, due to the normal inflation experienced in the U.S. monetary market, a transformation of this data is necessary. This transformation was conducted by converting the 2005 market values of the houses in the data set to values set in 1997 dollars accounting for the normal inflation observed in the U.S. over the eight years between data sets. This deflating of the 2005 values results in real prices, which are comparable across time periods. The hedonic price model is a reduced form model, meaning it involves the demand and supply side of the economy. Deflating the 2005 values helps in washing out the effects from the supply side of the model, such as inflation of building costs.

Although the housing bubble affected the entire United States, there were specific regions that were hit harder than others. To observe if preferences in these hard hit regions changed significantly more than preferences in other regions, the data were divided into four separate regions. These regions were determined by the Census Bureau. The classifications of the

regions are as follows: Region 1 is the Northeast[22] section of U.S. and Region 2 is the Midwest[23]. Region 3 encompasses the South[24], and Region 4 contains data from the West[25]. There is the possibility that consumer preferences did not change significantly in some regions while showing specific changes in other regions. If this occurs then it is possible that the price changes, in the regions in which preferences shifted, are so significant that if the data was analyzed at a national level without the information about the individual regions, the data could provide evidence supporting the assumption that preferences changed across the entire U.S. However, regionalized data would prove this assumption to be false.

The first test run on this model is an Ordinary Least Squares (OLS) regression. OLS is a method for linear regression that estimates the values in a statistical model by minimizing the sum of the residuals squared. The residuals are defined as the difference between the predicted and observed values. The OLS approach satisfies the Gauss-Markov theorem, which states that least squares estimators are BLUE (Best Linear Unbiased Estimators). Each estimator is assumed to be linearly related to the dependent variable. The estimators are unbiased, defined as their average or expected values being equal to the true parameter value. Although there may be more linear, unbiased estimators, estimators that are BLUE have the lowest variance of any such estimators. This is true under the assumptions that the error terms are expected to be zero, have a constant variance, and are uncorrelated with one another (Gujarati, 2003). OLS regression starts with the assumption that the true form of the data follows the equation $y_i = x_i ß_i + e_i$.

[22] Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, and Pennsylvania

[23] Ohio, Indiana, Illinois, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas

[24] Delaware, Maryland, Washington D.C., Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, Texas

[25] Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada, Washington, Oregon, California, Alaska, Hawaii

In this equation y_i is the dependent variable (LOGVALUE in this study) and x_i are the independent variables (the characteristic variables). $ß_i$ represents the coefficient of x_i and shows the effect x_i has on y_i. The error term is represented by e_i, this term accounts for any discrepancy between the actual values of y_i and the predicted outcome of $x_i ß_i$. This regression was used due to the research by Cropper, Deck, and McConnell (1988), which showed that linear models work best with reducing bias in hedonic models so to test the linear model; the linear regression form of OLS was used. The OLS regression is not immune to specification errors and in particular to this study, heteroskedasticity. To check to see if the data used had the problem of heteroskedasticity White's Test was run.

This test was proposed by White (1980) and it tests whether the residual variable in a regression is constant or homoskedastic. This is done by regressing the squared residuals from the regression onto the independent variables, the cross-products of the independent variables, and the squared independent variables. This results in a chi-square statistic with a null hypothesis of homoskedasticity.

OLS estimators are not BLUE in the presence of heteroskedasticity, however they do remain linear and unbiased. When OLS is run allowing for heteroskedasticity the resulting estimators can have overly larger variances causing inaccurate results from any *t* and *F* tests. This may result in coefficients that appear to be statistically insignificant (caused by a *t* value that is smaller than what is appropriate) when it may, in fact, be significant if the correct confidence intervals had been established. White also developed a method of correcting for heteroskedasticity known as White's heteroskedasticity- corrected standard errors. White's standard errors, also known as robust standard errors can be larger or smaller than the OLS standard errors, which should correct for heteroskedasticity and result in more accurately

estimated *t* values than those obtained by OLS. Another solution to heteroskedasticity if the form of it is unknown is to run a regression using generalized methods of moments estimation (GMM). After running the regressions and obtaining the results a conversion must be made so that the results can be analyzed. To transform the regression results from the hedonic pricing model into values upon which a strong analysis and statistical inference can be made the results are converted into the marginal implicit prices (MIP) for each independent variables. These MIPs are found using the equation $MIP_i = \delta p/\delta x_i = \delta(exp(\hat{\beta}`x))/\delta x_i = \hat{p} \cdot bi$ where, $\hat{p} = exp(pred\ LOGVALUE)$. This transformation is the same equation as the first-order necessary conditions shown in Rosen's paper. It should be noted that the MIPs for the dummy variables of GARAGE and SATPOL were not calculated using the previous equation since they do not represent continuous variables. Instead the implicit prices of these variables were calculated by taking the difference between the implicit prices that would result if all of the observations had the value of one and if they all were zero. After the MIPs are calculated and analyzed, the true question of this study needs to be answered; did the housing bubble significantly affect how consumers value these characteristics of a house? An analysis of *t* tests run across time periods for each region and the whole U.S. will provide substantial statistical information that should help provide an answer for this question.

VII. Results

After running the first OLS regression the results were tested for heteroskedasticity using White's test. The results from the White test on each OLS regression are given in Table 1; they are listed by year and region.

Table 2: Results of White's Heteroskedasticity Test

Year-Region	Statistic	DF	Pr > ChiSq
1997-All Regions	1699	63	<.0001
1997-Region 1	467	63	<.0001
1997-Region 2	706.1	63	<.0001
1997-Region3	733.6	63	<.0001
1997-Region4	544.9	63	<.0001
2005-All Regions	998.7	63	<.0001
2005-Region 1	285.2	63	<.0001
2005-Region 2	645.4	63	<.0001
2005-Region 3	533.3	63	<.0001
2005-Region 4	475.9	63	<.0001

Note: X^2 critical value for the .001 significance level with 63 DF = 103.46

With the null hypothesis of White's test being homoskedasticity it is easy to see that the null hypothesis is rejected in all of the regions in both 1997 and 2005. These results show that the specification error of heteroskedasticity is present in every model. As stated before a simple approach to correcting heteroskedasticity is to use White's standard errors. These standard errors were found and using these, new *t* values were calculated. These new, more accurate *t* values have been reported with the OLS results. These results are shown in the following table.

TABLE 3.1: 1997 OLS Results With White's Standard Errors

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
INTERCEPT	9.9358***	10.1547***	9.8919***	9.6925***	10.0783***
	(318.00)	(134.92)	(159.14)	(201.30)	(138.55)
AMTXK	0.1199***	0.1435***	0.1729***	0.0773***	0.1551***
	(39.89)	(22.33)	(26.23)	(17.35)	(16.41)
BEDRMS	0.0713***	0.0527***	0.0631***	0.0921***	0.0730***
	(12.78)	(5.04)	(6.98)	(8.60)	(5.61)
AGE	-0.0008***	-0.0010***	-0.0022***	-0.0005	0.0033***
	(-4.19)	(-2.63)	(-6.44)	(-1.37)	(7.36)
GARAGE	0.2514***	0.1407***	0.2836***	0.1824***	0.3484***
	(23.36)	(6.69)	(11.77)	(11.73)	(10.08)
HOWN	0.0295***	0.0432***	0.0361***	0.0273***	0.0249***
	(12.07)	(7.43)	(8.09)	(7.41)	(4.39)
LOT10K	0.0006**	0.0001	0.0006	0.0017***	0.0013**
	(2.51)	(0.11)	(1.49)	(3.77)	(2.05)
SATPOL	0.0844***	0.0674*	0.0969***	0.1033***	0.0669*
	(5.09)	(1.72)	(2.90)	(4.14)	(1.92)
UNITSFH	0.0097***	0.0060***	0.0081***	0.0174***	0.0119***
	(19.78)	(6.79)	(10.45)	(17.10)	(9.74)
ZINC2K	0.0034***	0.0032***	0.0024***	0.0033***	0.0031***
	(32.02)	(16.55)	(12.00)	(17.72)	(13.70)
TOTBATH	0.1372***	0.0731***	0.1032***	0.1589***	0.0871***
	(22.30)	(5.94)	(10.72)	(14.16)	(5.41)
Number of Observations	19589	3354	5375	7050	3810
Adjusted R^2	0.4498	0.4904	0.5058	0.4541	0.4206

Notes: Dependent Variable is LOG(VALUE). Parameter estimates shown with t-ratio in parentheses below.
*** 0.01 level of statistical significance; ** 0.05 level of statistical significance; * 0.1 level of statistical significance.

TABLE 3.2: 2005 OLS Results With White's Standard Errors

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
INTERCEPT	9.9363***	10.3867***	9.9393***	9.7210***	9.9058***
	(236.44)	(104.20)	(137.43)	(152.71)	(94.11)
AMTXK	0.0900***	0.0958***	0.0901***	0.0631***	0.1045***
	(30.60)	(19.33)	(14.39)	(12.05)	(15.87)
BEDRMS	0.0989***	0.0827***	0.1057***	0.1030***	0.0714***
	(12.58)	(4.97)	(9.42)	(7.42)	(3.98)
AGE	0.0014***	0.0004	-0.0014***	0.0009*	0.0082***
	(5.57)	(0.73)	(-3.56)	(1.95)	(12.52)
GARAGE	0.2863***	0.1050***	0.3252***	0.1594***	0.5988***
	(19.94)	(3.69)	(10.87)	(7.82)	(12.06)
HOWN	0.0434***	0.0478***	0.0379***	0.0420***	0.0554***
	(12.63)	(6.15)	(7.11)	(7.99)	(6.50)
LOT10K	-0.0022***	-0.0023***	0.000023	-0.0001	-0.0003***
	(-7.40)	(-3.64)	(0.05)	(-1.50)	(-2.80)
SATPOL	0.1035***	0.1354***	0.1140***	0.1211***	0.0355
	(5.02)	(2.73)	(3.46)	(4.00)	(0.76)
UNITSFH	0.0013***	0.0007	0.0023***	0.0040***	0.0005
	(3.51)	(1.14)	(3.72)	(6.01)	(0.44)
ZINC2K	0.0023***	0.0017***	0.0017***	0.0023***	0.0023***
	(22.31)	(9.16)	(9.43)	(12.99)	(11.17)
TOTBATH	0.1975***	0.1092***	0.1530***	0.2525***	0.2134***
	(20.86)	(5.48)	(11.66)	(15.00)	(9.17)
Number of Observations	18648	3287	4949	6646	3766
Adjusted R^2	0.3802	0.4012	0.4571	0.3862	0.3885

Notes: Dependent Variable is LOG(VALUE). Parameter estimates shown with t-ratio in parentheses below. *** 0.01 level of statistical significance; ** 0.05 level of statistical significance; * 0.1 level of statistical significance.

It should be noted that the regression technique of GMM can also be used to correct for heteroskedasticity and was done as part of this study. The results however, were basically the same as the results from the OLS regression with White's standard errors so only results from the OLS regression are presented here. The results from the OLS regression run on the 1997 data

show that all of the variables had the correct expected relationship with the dependent variable. The only exception to this is the variable AGE in Region 4. A possible explanation for this could be that in the western region of the U.S. older homes maybe valued higher due to a vintage effect. AGE was also significant at the 0.01 level in the whole U.S. and every region except Region 3 where it was not significant at any level. The variables AMTXK, BEDRMS, GARAGE, HOWN, UNITSFH, ZINC2K, and TOTBATH all were significant in the data from the whole U.S. and across every region at the 0.01 level of significance. LOT10K was significant in Region 3 at the 0.01 significance level and the whole U.S. and Region 4 at the 0.05 level. However, in Regions 1 and 2 LOT10K had no statistical significance. Finally the variable SATPOL was significant in the whole U.S. and Regions 2 and 3 at the 0.01 level and was also significant in the remaining regions but only at the 0.1 level.

A quick analysis of the results from the OLS regression run on the 2005 data shows there are some interesting changes in the parameter estimates from those obtained from 1997. Although there are some obvious differences, it should first be pointed out that the variables AMTXK, BEDRMS, GARAGE, HOWN, ZINC2K, and TOTBATH are all still significant for the whole U.S. and in every region at the 0.01 level, which was the same case for these variables in the 1997 results. However, unlike in 1997 where in every region except Region 4 AGE was had a negative effect on the dependent variable, in the 2005 results AGE is actually positive except in Region 2. This change in the sign of the estimate could be contributed to the fact that the housing bubble attracted "flippers" to the market and with the plan of purchasing a house, renovating it, and then selling it for profit this type of action could have driven the demand for older homes up due to the fact that older home would be the "fixer-upper" type of house a person intending to flip would want. AGE is not significant in Region 1 while Region 3 shows a 0.1

level of significance and in the whole U.S., Region 2 and Region 4 AGE is at the 0.01 level. Another variable that changed signs from 1997 to 2005 is LOT10K. In 1997 LOT10K is positive in every region however, like AGE the signs have changed for every region but Region 2. This could be due to, like the sign change in the variable AGE, by flippers. The fact that flipping a house can take months to accomplish perhaps influenced the flippers to buy homes with smaller lots to avoid as much lawn renovation and care as possible in order to cut back on the time and money needed to flip a house. This also could have been caused by an increase in the cost of lawn care. This trend would have caused the demand for houses with larger lots to fall during the housing bubble. The observation can be made that neither AGE nor LOT10K changed signs in Region 2 as they did in every other region and could possibly be explained by the fact that the Midwest region of the United States was not affected as much as the other regions by the housing bubble. It is well known that major metropolitan areas in the West, South and Northeast were hit very hard by the collapse of the housing market. This could lead one to conclude that these areas were highly concentrated with housing bubble activities such as flipping and subprime lending, more so than the Midwest region. LOT10K, however, is not significant in Region 2, or Region 3 but does display a significance level of 0.01 in the whole U.S. and the remaining regions. SATPOL is not significant in the 2005 data across the whole U.S. and in every region as it is in the 1997 data because it shows no significance in Region 4 but it does have a significance level of 0.01 in all other regions and the whole U.S. In the results from the 1997 data UNITSFH is one of the seven variables that shows a 0.01 significance level across the whole U.S. and all the regions. However, the 2005 data shows a change in regions 1 and 4, where UNITSFH becomes non-significant. By looking at the results from the OLS regressions, a picture can begin to be drawn about how the values of the characteristics included

in this study have changed over the time period. An analysis of the marginal implicit prices and the resulting comparative *t* tests should help complete this picture. The MIPs in the following table were calculated by the formula presented earlier in this study and using the results from the OLS regressions. The MIP for each variable was calculated for every observation in the data set and then the mean and standard deviation were also calculated in order to run the necessary *t* tests analysis. The mean MIPs and their standard deviations are shown in Tables 3.1 and 3.2.

TABLE 4.1: 1997 Marginal Implicit Prices

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	13190.09 (8502.22)	18616.89 (10616.73)	17132.37 (11030.68)	7141.29 (4791.71)	23008.86 (14211.64)
BEDRMS	7838.72 (5052.77)	6842.55 (3902.13)	6253.36 (4026.23)	8509.24 (5709.59)	10824.33 (6685.75)
AGE .	-89.09 (57.42)	-132.37 (75.48)	-216.02 (139.09)	-46.22 (31.01)	482.38 (297.95)
GARAGE	17404.83 (23303.28)	10528.34 (16539.46)	20259.89 (20896.77)	8530.50 (16481.73)	37339.66 (35189.31)
HOWN	3245.12 (2091.77)	5602.59 (3195.01)	3578.05 (2303.73)	2520.81 (1691.43)	3691.25 (2279.94)
LOT10K	70.94 (45.73)	7.53 (4.29)	62.13 (40.00)	156.86 (105.25)	195.12 (120.52)
SATPOL	7767.80 (7203.01)	7673.76 (5991.83)	8208.66 (7150.02)	7648.84 (7795.31)	8165.66 (7789.45)
UNITSFH	1063.41 (685.47)	773.30 (440.99)	803.84 (517.55)	1607.18 (203.75)	1766.02 (279.53)
ZINC2K	371.96 (239.76)	410.46 (234.08)	237.03 (152.61)	303.65 (203.75)	452.56 (279.53)
TOTBATH	15090.57 (9727.26)	9484.42 (5408.72)	10223.12 (6582.16)	14684.48 (9853.11)	12931.18 (7987.06)
Number of Observations	19589	3354	5375	7050	3810

Note: Mean marginal implicit prices followed by the standard deviation in parentheses below.

TABLE 4.2: 2005 Marginal Implicit Prices

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	15165.77 (22450.90)	19757.45 (21419.93)	11284.25 (12420.59)	7919.74 (12525.38)	31786.08 (43838.98)
BEDRMS	16666.22 (24672.11)	17052.88 (18487.79)	13242.05 (14575.54)	12928.66 (20447.21)	21721.22 (29957.65)
AGE	243.39 (360.31)	84.98 (92.13)	-172.82 (190.23)	115.96 (183.39)	2491.34 (3436.02)
GARAGE	32113.54 (67675.34)	12902.69 (27446.27)	30302.05 (41886.94)	11387.96 (32675.12)	124980.09 (197317.34)
HOWN	7318.78 (10834.47)	9858.92 (10688.5)	4741.63 (5219.12)	5265.94 (8328.29)	16867.27 (23263.14)
LOT10K .	-375.62 (556.05)	-468.23 (507.62)	2.88 (3.17)	-94.12 (148.85)	-1046.81 (1443.74)
SATPOL	14740.45 (25647.58)	23541.63 (30491.78)	12201.08 (15920.14)	12459.86 (23696.38)	9408.24 (15451.81)
UNITSFH	221.83 (328.39)	151.61 (164.36)	288.54 (317.6)	496.7 (785.55)	154.59 (213.20)
ZINC2K	390.77 (578.49)	355.81 (385.75)	211.9 (233.24)	294.53 (465.81)	697.77 (962.35)
TOTBATH	33271.46 (49253.95)	22529.05 (24424.75)	19160.25 (21089.71)	31691.74 (50121.78)	64943.95 (89569.99)
Number of Observations	18648	3287	4949	6646	3766

Note: Mean marginal implicit prices listed followed by the standard deviation in parentheses below.

These results allow for a better understanding of the actual value consumers place upon these individual characteristics of a house. MIPs are easily defined as the marginal willingness to pay and now that results from both years have been obtained a comparison can be made across these years to determine if the willingness to pay by consumers changed significantly for any of the characteristics. Perhaps consumers preferred smaller homes, or more bedrooms, or even smaller lot sizes during the housing bubble than compared to the time period before it. Before testing for any significant changes in the MIPs it may be helpful to see the magnitude of these changes. This can be done by calculating the percentage change in the mean MIPs for each variable from 1997 to 2005. This percentage can be calculated by formula $[(\bar{X}_1 - \bar{X}_2)/\bar{X}_2)]100$.

Table 5 contains the percentage change in each variable in the whole U.S. and each individual region.

TABLE 5: Percentage Change in Mean MIPs from 1997 to 2005

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	14.98%	6.13%	-34.13%	10.90%	38.15%
BEDRMS	112.61%	149.22%	111.76%	51.94%	100.67%
AGE	373.20%	164.20%	19.99%	350.89%	416.47%
GARAGE	84.51%	22.55%	49.57%	33.50%	234.71%
HOWN	125.53%	75.97%	32.52%	108.90%	356.95%
LOT10K	-629.49%	-6318.19%	-95.36%	-160.00%	-636.50%
SATPOL	89.76%	206.78%	48.64%	62.90%	15.22%
UNITSFH	-79.14%	-80.39%	-64.10%	-69.09%	-91.25%
ZINC2K	5.06%	-13.31%	-10.60%	-3.00%	54.18%
TOTBATH	120.48%	137.54%	87.42%	115.82%	402.23%

This table shows that there was a positive change in the MIPs of the variables BEDRMS, AGE, GARAGE, HOWN, SATPOL, and TOTBATH in 2005 from 1997 across every region including the whole U.S. The variables LOT10K and UNITSFH show a decrease in MIPs in the Whole U.S. and each individual region in 2005. The two variables that do not show consistent change in every region are AMTXK and ZINC2K. AMTXK shows an increase everywhere except for Region 2 while, ZINC2K only increased in the whole U.S. and Region 4.

To find out if any of these changes are statistically significant a statistical test needs to be run on the MIPs for each variable. The test used in this study to determine the significance of changes in the MIPs is Welch's *t* test. Welch's *t* test is used because this study uses independent samples. This means that when testing significance across the samples the *t* test must deal with

the fact that the two statistics being analyzed have unequal variances and unequal sample sizes. The Welch *t* test accommodates these inequalities and is defined by the following equation.

$$\text{t stat} = (\bar{X}_1 - \bar{X}_2)/s_{\bar{X}_1-\bar{X}_2}$$

Where:

$$s_{\bar{X}_1-\bar{X}_2} = \sqrt{s_1^2/n_1 + s_2^2/n_2}$$

With:

$$\text{DF} = \frac{(s_1^2/n_1+s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1)+(s_2^2/n_2)^2/(n_2-1)}$$

X1 is the mean value of the MIPs for 1997 and X2 is the mean value for 2005. The standard deviations of the MIPs for 1997 and 2005 are s1 and s2 respectfully, while the sample sizes for each year are represented by n1 and n2. By using the 1997 MIPs as the base line estimates this *t* test will show if there is a statistically significant difference between the mean MIPs from 1997 and those in 2005. If the *t*-values resulting from this test are significant at the given level of degrees of freedom then it can be said that the mean values of MIPs in 2005 are significantly different from those in 1997. The results from the tests comparing the MIPs from both time periods are included in the following table.

TABLE 6: Results of Welch *t* test for Independent Samples MIPs

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTX	11.27*** (23666)	2.74*** (4787)	-25.21*** (9925)	4.75*** (8453)	11.69*** (4539)
BEDRMS	47.91*** (20135)	30.99*** (3573)	32.61*** (5642)	17.01*** (7619)	21.79*** (4135)
AGE	124.51*** (19549)	105.05*** (6342)	13.08*** (9008)	71.14*** (7003)	35.75*** (3821)
GARAGE	28.13*** (22817)	4.26*** (5374)	15.21*** (7131)	6.40*** (9689)	26.84*** (4002)
HOWN	50.46*** (19970)	21.89*** (3858)	14.44*** (6684)	26.36*** (7162)	34.59*** (3836)
LOT10K	-109.32*** (18887)	-53.73*** (3286)	-108.23*** (5447)	-113.32*** (11894)	-52.61*** (3817)
SATPOL	35.81*** (21437)	29.29*** (3534)	16.20*** (6742)	15.77*** (7992)	4.41*** (5548)
UNITSFH	-154.24*** (28435)	-76.41*** (4283)	-61.50*** (9029)	-111.75*** (7486)	-282.32*** (7118)
ZINC2K	4.12*** (24617)	-6.96*** (5397)	-6.42*** (8414)	-1.47 (8983)	15.02*** (4389)
TOTBATH	49.50*** (20031)	29.91*** (3601)	28.56*** (5832)	27.17*** (7129)	35.50*** (3824)

Note: The *t*-statistic for the variable and region is listed with the degrees of freedom for that test in parentheses below. *** 0.01 level of statistical significance; ** 0.05 level of statistical significance; * 0.1 level of statistical significance using the critical values for DF = ∞[26]. The null hypothesis for this test is, H0: $\bar{X}_{05} - \bar{X}_{97} = 0$

As the results of the *t* test show, every variable experienced a statistically significant change in its mean MIP from 1997 to 2005 in the whole U.S. and in every individual region. The only exception is the variable ZINC2K in Region 3. According to the results shown in the above table the mean MIP for the household income in the southern region of the U.S. did not significantly change from 1997 to 2005.

[26] This critical value can be found in the *t* table in Appendix D: Statistical Tables, Basic Econometrics 4th ed. by Gujarati. Pg. 961

Other interesting statistics to look at in this study are the marginal implicit expenditures as a proportion of house prices (MIEs). These MIEs are calculated using the formula for MIPs and the multiplying this by x_i/p. The MIEs are basically the MIPs adjusted by quantity consumed as a proportion of price. If the MIE formula is written out, $(\partial p/\partial x_i) \cdot (x_i/p)$, it can be seen that the MIEs are actually price elasticities. By analyzing the MIEs the changes in preferences can be observed without problems arising from inflated house prices during the bubble because they are calculated as a proportion of price. This will reduce the chances of an insignificant change appearing significant. The calculated MIEs are shown in the following two tables.

TABLE 7.1: 1997 Marginal Implicit Expenditures as a Proportion of House Price

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	0.1863023	0.3837405	0.259843	0.0872051	0.2206937
	0.1888738	0.253104	0.226524	0.1196882	0.2015367
BEDRMS	0.2218235	0.1684812	0.1962357	0.2813388	0.2295532
	0.060814	0.0494334	0.0549378	0.0709544	0.0647733
AGE	-0.02987	-0.0449001	-0.0917571	-0.0159135	0.1059058
	0.0184369	0.0242383	0.0517294	0.0103372	0.0676338
GARAGE	0.1713307	0.0950009	0.2088204	0.1135308	0.2567386
	0.0934319	0.0586778	0.0892453	0.0777094	0.0980671
HOWN	0.24218	0.3611523	0.2972545	0.2233746	0.2006549
	0.0530655	0.0704786	0.0628249	0.0513446	0.0459591
LOT10K	0.0047824	0.00041877	0.0053173	0.0137134	0.0637642
	0.0123011	0.001036	0.0132145	0.0335606	0.1994695
SATPOL	0.0743583	0.061226	0.0860765	0.0884229	0.0590673
	0.0221541	0.0155426	0.0232696	0.0292669	0.0182519
UNITSFH	0.194051	0.1376143	0.1699349	0.3242364	0.223718
	0.0907501	0.0607447	0.0774362	0.1556604	0.0996828
ZINC2K	0.1776979	0.1869208	0.1237563	0.1560336	0.1745832
	0.1350688	0.1396697	0.0889323	0.1237285	0.1294915
TOTBATH	0.272909	0.1418574	0.1954127	0.3193915	0.1852088
	0.1175049	0.0658546	0.0877602	0.1352489	0.0708191
Number of Observations	19589	3354	5375	7050	3810

Note: Mean marginal implicit expenditure as a proportion of house price listed with the standard error in parentheses below.

TABLE 7.2: 2005 Marginal Implicit Expenditures as a Proportion of House Price

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	0.2100754	0.3749598	0.1976293	0.1041821	0.2445221
	0.2635927	0.3523918	0.2224769	0.1654072	0.2854573
BEDRMS	0.3127188	0.2660558	0.3308666	0.3217337	0.2296821
	0.0831566	0.0726375	0.0874507	0.0810087	0.0648062
AGE	0.0631923	0.0218933	-0.0668378	0.0347672	0.32843
	0.0338399	0.0099059	0.0334468	0.0196685	0.1745994
GARAGE	0.1966226	0.0908351	0.2427565	0.1040106	0.4010199
	0.1014252	0.0570238	0.0919638	0.0671507	0.1409522
HOWN	0.3564578	0.3980224	0.3113344	0.3439672	0.4531755
	0.0725327	0.0763734	0.0625146	0.0739396	0.0887013
LOT10K	-0.018081	-0.0177509	0.000229496	-0.000608675	-0.0016544
	0.0450877	0.0441828	0.00053631	0.0014319	0.0051021
SATPOL	0.0905214	0.0914588	0.1002123	0.1034626	0.090822
	0.0265494	0.0250317	0.0273924	0.0330414	0.0261876
UNITSFH	0.0285544	0.0179504	0.0497617	0.0832208	0.0104422
	0.0229051	0.0154309	0.0387729	0.0662375	0.0078423
ZINC2K	0.165425	0.1429659	0.1122263	0.1523606	0.1761402
	0.1687054	0.1450193	0.1060241	0.1636108	0.1727227
TOTBATH	0.4149316	0.2256234	0.3072884	0.5381727	0.4691682
	0.1719202	0.0980445	0.129986	0.2241034	0.1768335
Number of Observations	18648	3287	4949	6646	3766

Note: Mean marginal implicit expenditure as a proportion of house price listed with the standard error in parentheses below.

The variables BEDRMS, HOWN, SATPOL, and TOTBATH all had positive MIEs in 1997 and show an increase in 2005. AMTXK also had positive MIEs in 1997 however, in 2005 only the whole U.S. and regions 3 and 4 show increases in MIEs while regions 1 and 2 show a decrease in the MIEs from 1997 to 2005. In 1997 AGE had positive MIEs except in region 4 and shows an increase across every region in 2005. LOT10K and UNITSFH were also positive in every region in 1997 but showed consistent decreases in all regions in 2005. ZINC2K followed the trend of LOT10K and UNITSFH but instead of a decrease across all regions, it showed an increase in Region 4 with a decrease everywhere else. Finally GARAGE had positive MIEs in

1997 but showed slight decreases in regions 1 and 3 from 1997 to 2005. Two variables also showed changes in the signs of the MIEs over time. AGE changed from negative to positive in regions 1 and 3 and in the whole U.S. where as LOT10K changed from positive MIEs in 1997 to negative in 2005 expect in Region 2. Since the MIEs are also price elasticities and they all lie in the range between zero and one, in both 1997 and 2005, when the absolute value is taken it can be seen that these fall in the inelastic category of elasticities.

As with the MIPs, Welch's *t* test for independent samples is run and the results are shown in table 8.

TABLE 8: Results of Welch *t* test for Independent Samples MIEs

Variable	All Regions	Region 1	Region 2	Region 3	Region 4
AMTXK	10.09*** (33672)	-1.16 (59562)	-14.07*** (102792)	6.85*** (120552)	4.19*** (6766)
BEDRMS	121.51*** (34063)	63.87*** (5779)	92.75*** (8200)	30.97*** (13216)	0.09 (7573)
AGE	331.60*** (28507)	147.52*** (4461)	29.29*** (9283)	187.11*** (9922)	72.99*** (4859)
GARAGE	25.33*** (37590)	-2.93*** (6639)	19.00*** (10193)	-7.68*** (13592)	51.66*** (6711)
HOWN	175.11*** (34070)	20.43*** (6573)	11.41*** (10260)	110.24*** (11767)	155.31*** (5636)
LOT10K	-66.92*** (21282)	-23.57*** (3290)	-28.20*** (5393)	-35.80*** (7076)	-20.24*** (3814)
SATPOL	64.47*** (36355)	58.99*** (5473)	28.14*** (9746)	28.13*** (13268)	61.16*** (6718)
UNITSFH	-247.11*** (22193)	-110.51*** (3792)	-100.87*** (8051)	-119.07*** (9640)	-131.65*** (3857)
ZINC2K	-7.83*** (35695)	-12.58*** (6617)	-5.96*** (9691)	-1.48 (12356)	0.44 (6981)
TOTBATH	93.86*** (32753)	40.79*** (5735)	50.82*** (8581)	68.67*** (10792)	91.55*** (4931)

Note: The *t*-statistic for the variable and region is listed with the degrees of freedom for that test in parentheses below. *** 0.01 level of statistical significance; ** 0.05 level of statistical significance; * 0.1 level of statistical significance using the critical values for DF = ∞[27]. The null hypothesis for this test is, H0: $\bar{X}_{05} - \bar{X}_{97} = 0$.

[27] This critical value can be found in the *t* table in Appendix D: Statistical Tables, Basic Econometrics 4th ed. by Gujarati. Pg. 961

There appear to be some differences between the results of this test and the test run on the MIP values. Like the MIP results most of the variables in every region including the whole U.S. experienced significant changes from 1997 to 2005. However, these results show that AMTXK no longer has a significant change from 1997 to 2005 in Region 1 and that BEDRMS also shows a non-significant change in Region 4. Like the results from the MIP *t* test ZINC2K still shows no significant change in Region 3 but, the MIE *t* test shows that ZINC2K experienced no significant change in Region 4 from 1997 to 2005.

VIII. Summary and Conclusion

The question was proposed; did the housing bubble in the United States from 1998 to 2006 have an effect on consumer preferences as they pertain to the housing market? Due to houses being heterogeneous goods a hedonic pricing model was constructed to help determine the implicit prices of certain house and neighborhood characteristics. This study used data from 1997 and 2005 in order to capture data from before the housing bubble and at its height. The data was also separated into 5 distinct samples, one from the whole U.S. and four from predetermined census regions. After running OLS regressions using the log of the market value as the dependent variable, the marginal implicit prices for each independent variable were calculated. Results showing that the changes in the mean MIPs were statistically significant except for the household income in the southern region of the U.S. were found by running a *t* test that is suited for independent samples.

The conclusions that can be drawn from the results of the regressions and the *t* tests are that during the housing bubble consumer preferences in the housing market did in fact change. From the percentage change in the MIPs (Table 5) it can be seen that Region 4, the western region of the U.S., typically had the highest, or one of the highest, percentage changes of all the regions including the whole U.S. percentage. These large changes in almost every variable should be expected due to the fact that the western region of the U.S. was largely affected by the collapse of the housing bubble. The large affects this region felt after the bubble popped can

only be assumed to have been caused by a large affect in the housing market in this region during the boom. The results found in this study offer evidence to support this assumption.

From tables 5 and 6 it can be seen that the variables describing the number of bedrooms in the house, the age of the house, whether or not the house has a garage, how the neighborhood rates as a place to live, whether or not the police protection in the neighborhood is satisfactory, and the total number of bathrooms in the house were all more valuable to consumers in 2005 as compared to 1997. All of these variables also had a positive significant effect on the dependent variable in the 2005 OLS regression with the exception of Age in Region 3. Therefore, it can be concluded that with the exception of AGE in Region 3, a one unit increase in any of these variables would have resulted in a significantly larger increase in the market value for the house in 2005 than the effect on the value the same one unit change would have had in 1997. In other words consumers preferred more of each of these variables.

Tables 5 and 6 also show that the variables describing the lot size and square footage of the house saw a significant decrease in the value consumers placed on them during the housing bubble. While the variable LOT10K went from having a positive effect on market value in 1997 to a negative effect during the bubble, the square footage of the house stayed positive during both time periods. There was however, a decrease in the MIP of UNITSFH from 1997 to 2005. This leads to the conclusion that although UNITSFH still had a positive effect on the market value of a house (the larger the house, the more its worth), it was not as important of a characteristic in 2005 as it was in 1997.

The final two variables had significant regional differences in the percentage changes of their MIPs. The amount of real estate taxes were discussed earlier as possibly having a positive effect on the dependent variable due to a reactive effect as in, as the market value of a house

increased, the taxes paid on that house would also increase. This variable is shown in Table 5 to have had a positive increase in MIPs for every region except Region 2. This leads to the conclusion that except in the Midwestern region, every where in the U.S. and in the U.S. as a whole the amount of taxes rose significantly in 2005 as opposed to 1997. This would make sense following the belief that AMTXK is a reactive variable and mean house prices rose across all regions. In the Midwest region AMTXK still had a positive significant effect on the dependent variable in 2005, it was just not as great of an effect as it had in1997. This could easily be caused by a possible reduction in the real estate tax rate in Region 2.

The household income is the second variable that was inconsistent from region to region. ZINC2K is positive and significant in the OLS regression from 1997 and 2005, however, the percentage change in the mean MIPs from 1997 to 2005 was negative in Regions 1, 2, and 3. This leads to the conclusion, like UNITSFH and AMTXK, that in these three regions the effect of ZINC2K was still positive but not as high in 2005 as it was in 1997. These could have been caused by a possible increase in the household income levels in these three regions during the housing bubble. However, to achieve a negative percentage change in MIP the percentage of the increase in household income would have to be more than the percentage increase of the housing prices for those regions. It should also be noted that ZINC2K in Region 2 was the only Change in MIP that was not statistically significant.

Tables 7.1, 7.2, and 8 show that when the MIPs are adjusted by the amount consumer and then calculated as a proportion of house price that a few of the variables did not in fact, change significantly over the specified time period. The amount of taxes paid did not experience a significant change in preference in Region 1 and the same can be said for the number of

bedrooms in Region 4. The household income remained insignificant in Region 3 but was found to be insignificant in Region 4 when the MIEs were calculated and tested.

Although this study can answer the question of whether or not consumer preferences significantly changed during the housing bubble, it cannot accurately say what underlying factors caused these significant changes. They may have been caused by increases in expected income or wealth, or by the expectation of continuing increasing housing prices, which would not actually change the consumer preferences but instead just reveal them in a new light with new information. These changes could also be attributed to a possible true change in consumer preferences brought about by the attitudes present during the housing bubble. The scope of this study cannot provide anything but speculation as to the causes of the significant changes however, a more in depth study into the underlying causes, using more than the data provided in this study, could possibly be able to make an inference into what truly caused the significant changes in consumer preferences shown in this study.

Another interesting aspect of this topic that cannot be covered in this study due to the lack of data is if consumer preferences only changed during the housing bubble or if the changes were simply natural changes that occur over time. A study that encompasses data from 1997, 2005, and 2013 could give and insight into whether or not consumer preferences will begin to revert back to the 1997 levels, will be statistically the same as those found in 2005, or will they be different from both years showing that preferences just change over time.

Finally, although this study can not make an inference about either the underlying causes of the changes in consumer preferences nor if these changes are last or will revert back to pre-housing bubble levels, it can answer the question it originally proposed; did consumer preferences in the housing market change significantly during the housing bubble? The answer

is, according to the results found in this study, yes; consumer preferences in the housing market during the recent housing bubble did significantly change when compared to preferences before the bubble began to inflate. Whether or not these changes are permanent or will revert back to pre-bubble levels still remains to be seen.

References

Books and Journals

Ball, Michael J. "Recent Empirical Work on the Determinants of Relative House Prices," *Urban Studies*, 1973, pp. 213-233.

Bartik, Timothy J. "The Estimation of Demand Parameters in Hedonic Price Models," *The Journal of Political Economy*, 1987, Vol. 95, No. 1, pp. 81-88.

Basmann, R. L. "A Theory of Demand with Variable Consumer Preferences," *Econometrica*, 1956, Vol. 24, No. 1, pp. 47-58.

Cropper, Maureen L., and Leland B. Deck, and Kenenth E. McConnell. "On the Choice of Functional Form for Hedonic Price Functions," *The Review of Economics and Statistics*, 1988, Vol. 70, No. 4, pp.668-675.

Court, A. T. "Hedonic Price Indexes with Automotive Examples," *The Dynamics of Automobile Demand*, General Motors, New York, 1939.

DiMartino, Danielle and John V. Duca. "The Rise and Fall of Subprime Mortgages," *Economic Letter, Federal Reserve Bank of Dallas*, 2007, Vol. 2, No. 11.

Goodman, Allen C. "Hedonic Prices, Price Indices and Housing Markets," *Journal of Urban Economics*, 1978, Vol. 5, pp. 471-484.

Goodman, Allen C. and Thomas G. Thibodeau. "Age-Related Heteroskedasticity in Hedonic House Price Equations," *Journal of Housing Research*, 1995, Vol. 6, No. 1, pp. 25-42.

Goodman, John L. Jr. and John B. Ittner. "The Accuracy of Homeowner's Estimates of House Value." *Journal of Housing Economics*, 1992, Vol. 2, No. 4, pp. 339-357.

----"Housing market Segmentation and Hedonic Prediction Accuracy," *Journal of Housing Economics*, 2003.

Greene, William. Econometric Analysis. Fifth Edition, Singapore: Pearson Education, Inc., 2003.

Griliches, Z. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change," *The Price Statistics of the Federal Government*, 1961, No. 73.

---"Price Indexes and Quality Change: Studies in New Methods of Measurement," Harvard University Press, Cambridge, MA, 1971.

Gujarati, Damodar. Basic Econometrics. Forth Edition. New York: Tata mcgraw-Hill, 2003.

Ichimura, S. "A Critical Note on the Definition of Related Goods," *Review of Economic Studies*, 1950-1951, Vol. 18, pp. 179-183.

Kain, John F. and John M. Quigley. "Note on Owner's Estimate of Housing Value," *Journal of the American Statistical Association*, 1972, Vol. 67, No. 343, pp. 803-806.

Lancaster, Kelvin J. "A New Approach to Consumer Theory," *The Journal of Political Economy*, 1966, Vol. 74, No. 2, pp. 132-157.

Marshall, Alfred. Principles of Economics: An Introductory Volume. Eighth Edition. London: Macmillan, 1920.

Palmquist, Raymond B. "Estimating the Demand for the Characteristics of Housing," *The Review of Economics and Statistics*, 1984, Vol. 66, No. 3, pp. 394-404.

Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, 1974, Vol. 82, No.1, pp. 34-55.

Samuelson, Paul A. "A Note on the Pure Theory of Consumer Behavior," *Economica*, 1938, Vol. 1, pp 61-71.

Tiebout, Charles M. "A Pure Theory of Local Expenditure," *The Journal of Political Economy*, 1956, Vol. 64, pp. 416-424.

Tinten, G. "Complementarity and Shifts in Demand," *Metroeconomica*, 1952, Vol. 4, pp.1-4.

Waugh, F.V. "Quality as a Determinant of Vegetable Prices," Columbia University Press, New York, 1929.

White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 1980, Vol. 48, No. 4, pp. 817-838.

Websites

Business Editors. "Fannie Mae's Targeted Community Reinvestment Act Loan Volume Passes $10 Billion Mark; Expanded Purchasing Efforts Help Lenders Meet Both Market Needs and CRA Goals." Business Wire May 7, 2001. Available online at: http://findarticles.com/p/articles/mi_m0EIN/is_2001_May_7/ai_74223918/

Fannie Mae. "About Fannie Mae." Available online at: http://www.fanniemae.com/kb/index?page=home&c=aboutus

Federal Financial Institution Examination Council. "Community Reinvestment Act." Available online at: http://www.ffiec.gov/cra/default/.htm

Federal Reserve. "Open Market Operations: Intended federal funds rate, Change and Level." Available online at: http://www.federalreserve.gov/monetarypolicy/openmarket.htm

Federal Trade Commission. "The Gramm-Leach Bliley Act." Available online at: http://www.ftc.gov/pricacy/privacyinitiatives/glbact.html

Freddie Mac. "Company Profile." Available online at: http://www.freddiemac.com/corporate/company_profile/

Holmes, Steven, A. "Fannie Mae Eases Credit To Aid Mortgage Lending." New York Times. September 30, 1999. Available online at: http://query.nytimes.com/gst/fullpage.html?res=9C0DE7DB153EF933A0575AC0A96F958260&sec=&spon=&pagewanted=1

Joint Center for Housing Studies of Harvard University. "The State of the Nation's Housing". 2008. Available online at: http://www.jchs.harvard.edu/publications/markets/son2008/son2008.pdf

Labaton, Stephen. "The Reckoning – Agency's '04 Rule Let Banks Pile Up New Debt." New York Times. October 2, 2008. Available online at: http://www.nytimes.com/2008/10/03/business/03sec.html?pagewanted=1&_r=1

Leonnig, Carol D. "How HUD Mortgage Policy Fed the Crisis: Subprime Loans Labeled Affordable." Washington Post. June 10, 2008. Available online at: http://www.washingtonpost.com/wp-dyn/content/article/2008/06/09/AR2008060902626.html

Middle Tennessee State University. "*Welch t* Test for Independent Samples." Available online at: http://frank.mtsu.edu/~dkfuller/notes302/welcht.pdf

RealtyTrac. "National Real Estate Trends." Available online at: http://www.realtytrac.com/trendcenter/

St. Louis Federal Reserve. "Lending Standards in Mortgage Markets." 2009. Available online at: http://research.stlouisfed.org/publications/es/09/ES0923.pdfu

U.S. Bureau of Labor Statistics. "CPI Inflation Calculator." Available online at: http://data.bls.gov/cgi-bin/cpicalc.pl

U.S. Census Bureau. "Median and Average Sales Prices of New Homes Sold in United States." 2010. Available online at: http://www.census.gov/const/uspriceann.pdf

U.S. Government Printing Office. "Taxpayer Relief Act of 1997." Available online at: http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=105_cong_public_laws&docid=f:publ34.105

Data

U.S. Dept. of Commerce, Bureau of the Census. AMERICAN HOUSING SURVEY, 2005: NATIONAL MICRODATA [Computer file]. ICPSR04593-v1. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 2006. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-06-13. doi:10.3886/ICPSR04593

U.S. Dept. of Commerce, Bureau of the Census. AMERICAN HOUSING SURVEY, 1997: NATIONAL MICRODATA [Computer file]. ICPSR02912-v2. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 1999. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-05-30. doi:10.3886/ICPSR02912