ANALYSIS OF RESIDENTIAL HOUSING MARKETS

IN LARGE U.S. METROPOLITAN AREAS

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Yu Lu

Dr. Shawn Ni, Dissertation Supervisor

DECEMBER 2009

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

ANALYSIS OF RESIDENTIAL HOUSING MARKETS IN LARGE U.S. METROPOLITAN AREAS

presented by Yu Lu,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Saku Aura

Professor Douglas J. Miller

Professor Shawn Ni

Professor Neil Raymon

Professor Xuemin (Sterling) Yan

ACKNOWLEDGEMENTS

I would like to gratefully and sincerely thank Dr. Shawn Ni for his advices, guidance, encouragement and patience during my graduate study at the Department of Economics of the University of Missouri. I am grateful to Dr. Saku Aura, Dr. Douglas J. Miller, Dr. Neil Raymon and Dr. Xuemin (Sterling) Yan for being on my committee.

This dissertation is my best gift to my parents who have been supporting my study for the past few years. This dissertation would not be possible without their supports and encouragements.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ABSTRACT

This dissertation consists of three essays focusing on the geography of U.S. residential housing markets, the role of housing attributes in the residential housing markets, as well as the housing tenure choices in large metropolitan areas. The first essay, "The Geography Of Residential Housing Prices In Large U.S. Metropolitan Housing Markets", conducts the clustering analysis to identify the geographic pattern of U.S. residential housing markets using American Housing Survey (AHS) data. Based on the housing prices and other relevant dwelling features, the K-means clustering analysis classifies the residential housing markets into three groups: a Coast group, a Central U.S. group and an in-between group. The clustering analysis rejects the hypothesis that the housing market associations between cities are random, and finds strong evidences of regional differences in housing price variations. The discriminant analysis is used to test the grouping analysis through the comparison with the region-based grouping method. The discriminant analysis provides strong support that the clustering analysis in this paper outperforms the region-based grouping analysis.

In the second essay, "Hedonic Analysis Of U.S. Residential Housing Markets", we apply the hedonic analysis to examine the relationship between prices of owner-occupied dwellings and housing attributes among the three clusters defined in Chapter One. The results show that both the housing attributes and the regional factors play an important role in determining the housing prices. The F-test verifies the disparity of housing attributes effects across clusters.

The third essay, "Analysis Of Tenure Choice And Timing Of Tenure Transition In U.S. Residential Housing Markets", incorporates the effect of household demographics in the

discussion of house tenure choice and the timing of tenure transition. It employs the logistic model to identify the relationship between household features and the tenure status. The Cox proportional hazard model is used to investigate the determinants of the timing of tenure change. The empirical results suggest that household demographics, household composition and economic status features considerably affect household tenure choice, as well as the decision of the transition from housing rental to ownership.

Chapter One: The Geography Of Residential Housing Prices In Large U.S. Metropolitan Housing Markets

I: Introduction

Residential housing prices have long been of great academic and business interest. In the last three decades, residential housing prices in large U.S. metropolitan areas exhibit considerable fluctuations over time and across cities. Exhibit 1 presents the average owner-occupied housing prices in twenty-six major metropolitan areas in the United States from year 1985 to 2004. The average house prices vary from a high value of $393,530 in San Francisco to a low value of $92,146 in Buffalo. The price growth is especially high in Columbus with the percentage change of 301% over time and surprisingly low in Providence with the percentage change of only 9%. The wide variety of house prices over time is plotted in Exhibit 2. The average house price levels vary substantially from a low of $66,380 in Buffalo in 1988 to a high of $561,112 in San Francisco in 1998. Among the twenty-six cities, housing prices in San Francisco are significantly higher than those in other cities and the housing prices increase in San Francisco is stable. The high price levels in San Francisco undoubtedly reflect the rapid economic development, low unemployment and high-quality of living environments in that area. Washington D.C. has high house prices levels in the first two periods, but the prices increase is slowing down in recent years. Some cities such as New Orleans and Buffalo have consistently low house prices, which we will make a further exploration in later part of the chapter. In the early time periods, Buffalo has the lowest house prices levels of $66,381 in 1988 and $87,514 in 1994. It is partly due to its slow economic development and growth in this area. The housing prices in some cities, such as Columbus, Oklahoma City and Seattle have extraordinary increments over the 1985-2002 periods. For instance, the average house price in Columbus jumps from $95,580 in 1987 to $383,283 in 2002.

The study on the geographical differences of housing attributes effects across cities helps us understand the interrelationship of housing prices across cities. This chapter focuses on the geographical pattern of the residential housing prices with the impacts of housing attributes. In particular, this chapter incorporates the impact of housing attributes on forming the national structure of metropolitan housing markets. The clustering method is adopted to examine the cross-sectional characteristics of housing prices in twenty-six U.S. metropolitan residential markets. The residential housing markets are classified into three homogenous groups based on not only the housing prices, but also some relevant elements such as unemployment rate, household income, dwelling size, housing unit quality and neighborhood quality. It provides supports to the hypothesis that the residential markets exhibit distinct structural patterns, and the national structural trends of housing prices are driven by macroeconomic factors.

Second, the discriminant analysis technique is used in this chapter to assess the robustness of the clustering analysis based on the approach introduced by Hoesli *et al.* (1997). Particularly the paper compares this grouping method with the geographical patterns defined by Kasarda (1995) using the discriminant analysis. The results show that the clustering analysis of this paper outperforms the region-based grouping method, so that economics dominates geography on the respect of differentiating housing markets.

The rest part of the chapter is structured as follows. Section II briefly reviews the literature on the application of clustering analysis on real estate studies. The third section illustrates the dataset, the methodologies used and the testing results. Section IV conducts the discriminant analysis to assess the validity of the clustering analysis. The last section is the conclusion.

II: Literature Review

A large body of research literature has studied the variations in the housing prices across the country. They demonstrate that house prices fluctuate over time and price levels vary between cities in the U.S., and propose some explanations to the fluctuations of housing prices over time and across cities.

Gyourko and Voith (1992) use a large panel dataset of median housing prices in 56 metropolitan areas over 18 years. They observe wide divergence in price levels across cities and show persistent appreciation trends in housing prices over the sample period. They focus on the impact of local and national circumstances on the volatility of house prices in these cities, and argue that the clear geographical pattern of the house prices is predominantly induced by economic growths in various cities. Economic growth is associated with employment growth, population growth and wealth increase in an area, so that city-specific income growths offset the increases of implicit trait prices. As a result, continued equal appreciations of housing prices and high divergences of price levels are observed.

Based on the study of the price dynamics in 130 metropolitan areas in the United States, Jud and Winkler (2002) draw the conclusion that major determinants of the level and development of house prices are population growths, real changes in incomes, construction costs and interest rates. Variations in house prices are mainly attributed to the differences in real income growths and population growths across cities and housing prices are also influenced by location-specific fixed effects. In particular, the residuals of housing price appreciation are subject to growth management policies and limitations on land availability in particular cities.

Studies by Abraham and Hendershott (1996) investigate the house price indices in 30 metropolitan areas over the period of 1977-1992. They find a strong localized effect of house price trends, with high price premium in the coastal, northeast and western areas. Large fluctuations in house prices are especially observed in these areas. They also show that owner-occupied house prices are autocorrelated over time and across areas. The real housing price appreciation is positively related to the real construction costs, employment and real income, and negatively related to real interest rates, which explains the changes in the equilibrium price.

Another research on the fluctuation of house prices represented by Kennickell, Starr-McCluer and Surette (2000) emphasizes the interplay between residential turnover rates and the levels of house prices, because both are partially related to the same set of environmental elements.

By recognizing the importance of the cross-sectional differences in real estate market, grouping methods are required for a further investigation on housing market segmentation. Some studies define the housing submarkets based on the geographical regions or the physical characteristics of the dwellings. For instance, housing markets are partitioned in terms of pre-existing geographic boundaries in Adair *et al.* (1996) and Schnare and Struyk (1976), or by socio-economic conditions in Galster (1987), Harsman and Quigley (1995) papers. Another way to classify homogeneous groups of housing market is the clustering analysis, and the literature on the clustering analysis of housing market is not extensive.

Cluster analysis techniques were first developed and used in the biological and ecological sciences, and have been applied to the research in many disciplines.

The application of cluster analysis in economic studies is first introduced by Elton and Gruber (1971). They emphasize the importance of homogenous grouping to the comparative analysis. They use the distance measuring procedure to assess differences and similarities between objects. They also explore alternative approaches to disaggregate economic data into homogeneous groups for hypothesis testing and forecasting.

Goetzmann (1993) compares the risk and expected return of investment portfolios with and without single family homes. He concludes that spreading investment in residential assets across regions substantially reduces the risk.

Abraham, Goetzmann and Wachter (1994) explore interrelationship of housing market returns using the 1977-1992 returns to housing price indices data in 30 metropolitan areas. They particularly emphasize the role that the interrelationship of housing market returns play on the purposes of equity investment, portfolio diversification and risk hedging. They apply the K-means clustering algorithm, and several grouping outcomes are identified. In addition, the bootstrapping testing is conducted to examine the robustness of the clustering algorithm, and the test outcome supports the results of the clustering analysis. The study verifies that the structural differences in housing markets exist between cities, so that housing market partition is not an effect of random association. The structural features of housing returns play an important role in diversifying debt and equity portfolio as well as hedging the housing market risks.

Goetzmann and Wachter (1995a) examine the clustering approach for the purpose of portfolio diversification with two datasets, effective rents data for twenty-one metropolitan areas and vacancies data for twenty-two metropolitan areas. The results support for the hypothesis of the existence of major families in real estate markets, but the groupings are not consistent with geographical patterns specified by pervious studies. In particular, they specify an oil and gas group and an industrial Northeast group and they observe a strong bicoastal effect. The bootstrapping approach is used to test the validity of clustering analysis and it reveals the significant association among cities.

Besides the study on U.S housing markets, the cluster analytical techniques are widely used in other countries. For instance, Goetzmann and Wachter (1995b) generate the K-means cluster analysis to investigate the real estate returns in the office markets across countries. They find that the global market can be disaggregated into three groups, European, Scandinavian, Iberian and Asian markets. The fluctuation of U.S. real estate market is part of the global market trend, so that there exits a strong cross-sectional relationship in the world office market.

Hoesli *et al.* (1997) conducts a cluster analysis to U.K. commercial property markets with a dataset containing the property returns for 156 property markets, and the dataset include three types of properties: retailing locations, office locations and industrial locations. For the result, the study does not identify a distinct regional clustering, and instead it claims that the property type plays a critical role in differentiating housing market behavior. The paper also uses the discriminant analysis and the test of the stability of the cluster structures to examine the study, and the results are supportive to the findings.

Past research has well recognized the substantial geographical variations in housing price fluctuations across cities. However, economists identified the clustering of residential housing prices in a limited manner, and very little empirical work has addressed the role of regional differences in housing price variations with the cluster analysis. This chapter develops the past works and extends the understanding of the volatility nature of the housing prices with the consideration of housing market characteristics. It focuses on the interaction between the house prices and the structure of metropolitan housing areas through the analysis to the marginal values of housing attributes. The research in this chapter will extend the understanding of the nature of volatility of the house prices with the certain characteristics of various housing markets. This chapter investigates the geographical patterns of residential housing markets and analyzes whether the regional characteristics can help to explain the price trends in owner-occupied housing markets. Moreover, it compares the grouping outcome with the previous definition of residential housing market segmentation, and discusses the influence of macroeconomic factors on owner-occupied housing prices.

III: Analysis Of Geographic Pattern Of Metropolitan Housing Markets

A. Data description

To investigate the geography of the house prices, we try to find groups of cities with a comparable pattern in the house prices trends over time. The data for our analysis were abstracted from the data files of Annual Housing Surveys (AHS) for the metropolitan statistical areas (MSAs) data from 1985 through 2004. The MSA American Housing Survey (AHS) is a household survey designed and sponsored by the Department of Housing and Urban Development, and conducted by the Bureau of the Census. AHS conducts a national survey for the U.S as a whole and a metropolitan area survey for individual MSAs, and the later is used by this analysis.

In information collection, Census Bureau interviewers visit or telephone the dwelling's occupants and ask questions about the quality of housing. The survey includes questions about household characteristics such as family size, race and income; dwelling characteristics such as number of rooms; neighborhood condition such as noise, trash and other related information on housing cost, location and so on.

The metropolitan survey consists of 47 metropolitan areas, which include both central cities and suburbs. The metropolitan areas are widely distributed geographically. However, the metropolitan area survey data are not necessarily representative of the whole housing market. Some metropolitan areas are excluded from this test due to inadequate time periods of observations, so there are totally 26 metropolitan areas used and Exhibit 2 contains the information of 26 metropolitan statistical areas (MSAs). The time dimension of the data used in this analysis is from 1998 to 2004.

B. Cluster analysis of residential housing prices

All variables used for this study are explained in Exhibit 3. The primary explanatory factor is the property value, which is assumed to represent the trends and changes in residential housing markets. The analysis also includes some descriptive variables of household characteristics and the dwelling qualities, including the unemployment rate, tax payment, mortgage rate, household income, unit size, rooms, crowding, neighborhood quality rating and unit quality rating. Specifically, the unemployment rate data is extracted from Bureau of Labor Statistics. It provides the unemployment rate in each metropolitan area in every survey year. Totally 2,385 observations are included and all variables are averaged by cities. Therefore, the analysis uses a 10×26 data matrix.

The clustering of data employs the hierarchical algorithm, to which each of the objects stands out as its own cluster initially and they are combined into a hierarchy or treelike structure based on the similarity of objects. Euclidean distance approach is the representative distance measurement to quantify the inter-object similarity in the hierarchical algorithm, and it is defined as the straight-line distance between objects in n-dimensional space. It focuses on the magnitude of the distances, and group objects that are close to each other.

This study involves 26 objects (metropolitan areas) and each object has a profile of 10 variables. The distance measure is used to distinguish the similarity of objects in the 10-dimensional space, and objects with the least distances are combined together. The distance approach is sensitive to the differences in scales of the variables. Since clustering is performed on the squared error, variables with a larger absolute value will contribute more to the error. Therefore variables with larger dispersion will have more impacts on the similarity measures and undue influences on clusters. As the ten variables in this analysis have different ranges, standardization is done by computing mean and variance of each variable and then converting variables to their standardized values with the mean of zero and the standard deviation of one.

Ward's method is used to assess the similarity between clusters that have multiple components in this analysis. The distance of two clusters is defined as the sum of squared errors within the clusters summed over all variables. The distance between cluster C_K and C_L is calculated as $D_{KL} = \frac{\| \bar{x}_K - \bar{y}_L \|^2}{\frac{1}{N_K} + \frac{1}{N_L}}$, where $\bar{x}_K$ and $\bar{y}_L$ are the mean of object values in cluster C_K and C_L. Therefore, the Ward's method minimizes the sum of squared errors across all variables in all clusters and the similarity of two clusters is based on the increase in squared error when two clusters are merged.

The empirical results of Ward's method are presented by Exhibits 4-6. Exhibit 4 displays the cluster history and the relevant statistics of Ward's Minimum Variance Cluster Analysis. In particular, the first column lists the number of clusters and second column displays the names joined. The *Frequency* column indicates the number of objects in the cluster. *SPRSQ* column exhibits the semipartial R-square, which is the proportion of variance reduced by joining the two clusters. For instance, while the "Cluster Joined" to Cluster 3 is Cluster 8 and Cluster 9, *SPRSQ* (the semipartial R-square) of Cluster 3 is 0.1021. It illustrates that the proportion of variance decreases by 10.21% by joining Cluster 9 and Cluster 8 to form Cluster 3. The column *RSQ* is R-square, the squared multiple correlation, which is the proportion of variance accounted for by the clusters. For example, when the "Number of Clusters" is 3, RSQ has a value of 0.331. It means that, if the 26 cities are grouped into three clusters, the proportion of variance accounted for is 33.1%. The last column *PSF* contains the pseudo F statistic, and it is useful in determining the number of clusters in the data. Local peak of the PSF value stands for a stopping point. The value of PSF reached a local maximum at 3 clusters, indicating that the dataset should be divided into three groups.

On the cluster numbers versus semipartial R-square curve in Exhibit 5, there is constantly large drop in semipartial R-square values from 0 to 3 clusters, suggesting that values between 0 and 3 are good numbers of clusters. When there are more than 3 clusters, the curve becomes flatter so that the reduction in semipartial R-square values is insignificant. The results indicate that Ward's method is sufficient in grouping objects in this analysis and the feasible number of clusters is 3.

If we put the cluster identification data back into the original city data and list them out, we get the result shown in Exhibit 6. The three groups are (1) Baltimore, Birmingham, Buffalo, Cincinnati, Cleveland, Indianapolis, Kansas City, Milwaukee, New Orleans, Pittsburgh, St. Louis; (2) Atlanta, Hartford, Minneapolis, Columbus, Memphis, Oklahoma City, Portland, Providence, Rochester, San Antonio and (3) Boston, Miami, San Francisco, Seattle, Washington D.C.

The three-cluster partition of the cities reveals a clear geographical pattern of housing markets shown by Exhibit 7. Cluster 1 is composed of eleven cities, which are located in the central area of the country. These cities have the lowest house prices levels, the highest unemployment rate and the lowest household income over the period of 1998-2004. They are the least developed areas and denoted as “cold” housing markets. On the other hand, cities in Cluster 1 generally have the largest dwellings, the best unit qualities and neighborhood qualities. The observed effects can be explained by the economic characteristics of these areas. The Central U.S. area contains the majority of the old-styled manufacturing industries, which require less-educated people and provide low incomes. The developments in these areas are relatively slow and the pressure on land use is small.

Ten cities are grouped in Cluster 2. The cities in this group have relatively higher house prices, lower unemployment rate, and higher household incomes compared to cities in the first category. Cluster-2 cities are closer to the coasts and are economically more developed than cities in Group 1. As an effect, the dwelling sizes are smaller and quality levels are lower than those of Cluster 1.

The rest five cities belong to Cluster 3. All these cities are large coastal cities and the house prices levels are distinctively higher than the other two groups, so that they are specified as “hot” housing markets. The rapid postwar growths of the aerospace, defense, electronics, service industries spurred the economies of the coastal areas. Due to the rapid economic development and high industrialization, high-educated and high-skilled people dominate the city residential bases; so that unemployment rates are low and household incomes are high in these areas. On one hand, good job opportunities and high earnings attract immigrants from other parts of the country to move in. On the other hand, extensive transportation conditions, relatively well-educated labor force, ready access to investment capital and strong local markets attract more and more firms to enter the markets. As a result, the competition on land use becomes tense, so that in the residential housing markets of these areas, the dwelling qualities and neighborhood qualities are not as high as those in other areas, and the units are typically small.

There may be some exceptions in housing markets grouping, which can be caused by the accuracy of the data source or other reasons. Assessment on misallocation is conducted in the later part of this chapter.

C. Robustness test by discriminant analysis

The discriminant analysis is conducted to test the robustness of the cluster analysis. The intuition of the discriminant analysis is that it categorizes a set of observations into groups based on certain discriminant criteria. The classification criterion is developed upon some measures of generalized squared distance. This analysis assumes a linear discriminant function for clusters, so that the generalized squared distance is determined by pooled covariance matrix.

In this analysis, the geography of the house prices for the 26 cities displays a contrast between the central and the coastal areas. It reflects the general pattern of economic and demographic changes over the past three decades. The assessment of the clustering analysis is established on the comparison with the region-based grouping by Kasarda (1995). The well-known housing grouping by Kasarda (1995) divides residential housing markets into four clusters, West, South, Northeast and Midwest according to the geographic regions. Kasarda (1995) describes the associated features of housing prices between the Sunbelt (the West and South) area and the Frostbelt (the Northeast and the Midwest) area, and states that the geographic characteristics determine the housing prices differences among these groups.

Specifically, highly competitive housing markets are observed in the Sunbelt area, and less competitive housing markets are observed in the Frostbelt area. It is mainly caused by the geographic characteristics and regional advantages in this area. According to Kasarda (1995), the twenty-six cities in our data sample are classified into four groups as shown in Exhibit 9, and the grouping is different from the clustering analysis result. The discriminant analysis is used to compare the region-based grouping by Kasarda (1995) with the grouping of this paper. The conjecture of the discriminant analysis is to look at the relevant statistics of the region-based grouping and the grouping of this paper respectively. If the cluster analysis of this paper is more effective in capturing the housing price characteristics, its discriminant analysis results should outperform that of the region-based method.

There are several assessments of the validity of the grouping approaches, among which Eigenvalues and Wilks' Lambda are good measures. Particularly, the eigenvalues greater than unity or the chi-squared values of Wilks' Lambda statistically significant at certain confidence level specify good grouping results. The two cluster methods have eigenvalues of 23.749 and 11.358 respectively, suggesting that both have a high degree of success in classifying objects into groups. The Wilks' Lambdas of the two grouping methods are 0.0063298 and 0.0157364 respectively, with both p-values less than 0.0001. Therefore they are both statistically significant at the 5% confidence level and they are valid grouping approaches.

Another examination is based on the generalized squared distances, which measures the distances between means of groups. It indicates how different the groups are from one another, and the results are shown in Exhibit 10 and Exhibit 11. Exhibit 10 exhibits the generalized squared distances between groups defined by the clustering analysis of this paper, and Exhibit

11 displays the results of the region-based grouping method. The generalized squared distances between any two regions are always larger in Exhibit 10 than those in Exhibit 11. For instance, the generalized squared distance between Cluster 1 and Cluster 2 is 41.1523, while the distance between the Midwest Group and the Northeast Group is 1.43783. The difference suggests that averagely the groups determined by the clustering analysis have larger distances than the groups defined by the region-based method, so that the clustering analysis of this paper does a better job in characterizing the features of residential housing markets.

The fitness of the groupings can also be evaluated by the estimate of error rates (probabilities of misclassification). Exhibit 12 and Exhibit 13 show the number of observations and percent classified into region for two clustering methods. The region-based grouping correctly classifies 88.46 percent of cases with 3 misallocations, while the clustering analysis in this paper properly classified 96.15 percent of cases with only 1 misallocated object. With the comparisons, this clustering analysis outperforms the region-based grouping, although both have a high measure of success.

IV: Conclusion

This chapter seeks to measure and explain the volatility or the drift of house prices over time and space. Clustering analysis is employed to identify groups of residential housing markets in the United States and three submarkets are defined in U.S. residential housing market: a Central U.S group, a Coast group and an in-between group. It is observed that a distinctive geography of the house prices exists in 26 metropolitan cities, with low house prices in Central U.S. region and high house prices in coastal cities. This chapter also develops the discriminant analysis to test the fitness of the grouping, and it supports the results of the clustering analysis of this paper. Our analysis substantiates that the variation of house prices is embedded in, and influenced by, the environmental characteristics of the local housing markets. Associations across cities are not random and they are determined by the interregional economic differences in various cities. In this case, the characteristics of housing markets are subject to the economic development and industrialization in the areas. High economic growth leads to high house prices, partially because it results in more job opportunities, high income and young people's propensity to move.

Next chapter will incorporate the effect of housing attributes to explore the cross-sectional differences in residential housing prices.

Chapter Two: Hedonic Analysis Of U.S. Residential Housing Markets

I: Introduction

As discussed in Chapter One, the prices of owner-occupied housing exhibit strong geographic features, and the residential housing markets in large metropolitan areas can be classified into three clusters. The hedonic analysis approach has long been used as a standard tool to model the price of complex commodities, such as housing. This chapter uses the hedonic pricing model examining the relationship between residential housing prices and housing attributes in each cluster to explore the geographic pattern of housing prices.

Since the pioneer work of Lancaster (1966) and Rosen (1974), a large number of studies have been generated to explore the hedonic pricing model. Previous studies of the hedonic analysis state that property values are subject to its utility-generating characteristics, so that households' preferences to house features, such as dwelling quality, neighborhood quality and job opportunities substantially influence their purchase decisions. These preferences are translated into the marginal effects of housing attributes on housing prices. Early literature focus on the theoretical foundations as well as empirical applications of the hedonic model, while the recent studies concentrate on alternative techniques to mitigate the limitations of the hedonic model for better estimation. This chapter applies the hedonic pricing model to evaluate the effect of housing attributes on residential housing prices using the micro-data from U.S. Metropolitan Statistical Area American Housing Survey (AHS).

We will examine if the housing attributes effects on housing prices help to demonstrate the geographical diversity. First, this chapter assesses if the residential housing prices are determined by the marginal value of housing attributes. Second, this chapter examines whether the housing attributes effects on property values help to explain the geographical pattern of residential housing prices. The remainder of the chapter is organized as follows. Section II is a brief review of the literature on the hedonic analysis of housing markets. In section III, we apply the hedonic pricing model to assess the margin effects of housing attributes on housing prices of the pooled data and three clusters respectively. Section IV tests the differences of housing attribute effects among clusters. The last section is the conclusion.

II: Literature Review

A. Theoretical issues of hedonic model

Lancaster (1966) first introduces the idea that consumers acquire utility not only from goods themselves, but also from the features of goods. The utility-generating characteristics of goods can be used for the demand-side study of the market. This concept becomes the foundation of the hedonic pricing model.

Rosen (1974) focuses on both the demand and supply sides of the market. The theory postulates that the utility of differentiated products is attributable to the characteristics of the products, and the price of specific characteristic is determined by both consumers and producers of the goods.

Later studies pay great attention on various characteristics of residential housing as a good, and related theoretical modifications of the hedonic model are conducted. The special features of owner-occupied housing typically include: heterogeneity, nonlinearity and specification, the disequilibrium nature of housing markets as well as identification difficulties.

The residential housing is a typical example of heterogeneous products. The price of particular dwelling is determined not only by the effective prices of its attributes, but also implicitly by consumer preferences, non-market constraints, incomes etc. The problem has been well discussed by Fisher and Shell (1971) and later by Sheppard (1999).

Another important feature of housing market is the nonlinearity. In a linear supply and demand model, the price of a good is determined by the market, and individual consumers are usually price takers. As a result, the quantity of consumption has no impact on the price. In a nonlinear hedonic model, prices and quantities are correlated. Consumers choose to consume a certain amount of some characteristics, and the consumption decisions are influential to the characteristics prices. Papers such as Blomquist and Worley (1982), Diamond and Smith (1985) and Follain and Jimenz (1985) present prominent works on this topic.

The disequilibrium nature of housing markets has been well recognized. In particular, the supply and demand quantities do not generally equate to the quantity traded in the market. Bowden (1978) paper establishes an approach to solve the problem of disequilibrium by using only observations in or near the equilibrium for the hedonic estimation. Fair and Jaffee (1972) propose four methods such as the sample separation approach, and apply them to the housing market to solve for the absence of an equilibrium condition problem.

The hedonic pricing model has the potential difficulty in identifying and specifying relevant factors, as well as the potential problem of the imprecise estimation of coefficients. Specifically, there is no theoretical benchmark guiding which factors should be included in the model and how they are related to the price of the underlying product. The choice of dependent variable is generally the contract rent or the housing unit value. However, a controversial issue is the assessment of housing value. Most U.S data use the self-reported appraisals, which arouse special concerns to the accuracy of owner estimation. Kain and Quigley (1972) and Follain and Malpezzi (1981) conduct a test on the accuracy of owner assessments of housing values, and draw the conclusion that if the sample size is large enough, the biases are moderate and the estimate is robust. But the recent study by Goodman and Ittner (1992) find large biases, so that the reliability of the self-reported appraisals may be of great concern.

Recent sales price is another source of housing unit value, and it has the advantages of higher accuracy and independency. Some papers such as Galzlaff and Haurin (1997) make an essential discussion that the housing units recently traded may not be a good representative of the whole housing stock, so that the recent sales prices do not play an appropriate role.

Compared with the choice of dependent variable, the selection of independent variables draws more attention. A large number of housing characteristics have potential impacts on housing unit value, but which features should be included as independent variables is ambiguous. Moreover, the problem of omitted variables is also involved. Butler (1982) and Ozanne and Malpezzi (1985) test the effect of omitted variables on the bias and stability of the coefficient estimates, and infer that there is considerable bias associated. Some other papers such as Amemiya (1980) make constructive suggestion on how to select variables.

As there is no theoretical guide on the functional form of a hedonic regression, assorted opinions are given, such as Halverson and Pollakowski (1981) and Rosen (1974). The most common specification of functional form is linear or log-linear i.e. semi-log. Follain and Malpezzi (1980) favor log-linear specification for the following reasons. First, the semi-log form allows for the variation in the dollar value of a characteristic so that the price of one factor is connected to other components in the housing unit. For example, the linear model assumes that the second bedroom adds the same amount of value as the fifth one does, which seems to be implausible according to the diminishing marginal effect. On the contrary, the semi-log model allows for the value added varying proportionally with other features of the house. Second, the semi-log specification mitigates the intrinsic problem of heteroskedasticity (i.e. changing variance of the error term). Also semi-log model has more computational flexibility and feasibility and higher explanatory power than the linear model. In terms of flexibility, some economists like Christensen, Jorgensen and Lau (1971), Capozza, Green and Hendershott (1996, 1997) suggest the trans-log functional form.

B. Empirical applications of hedonic model

With the theoretical development of the hedonic model, some economists have been focusing on the application of the hedonic model to real housing markets.

One primary contribution of the hedonic model is to make improvements in housing price indexes. The representative works are Follain and Ozanne (1979), Follain and Malpezzi (1980), Malpezzi, Chun and Green (1998) and Thibodeau (1989, 1995). However, such “standard model” generally deduces the pattern of declining prices with the distance from the city center and ignores the submarket effect. For instance, Follain and Malpezzi (1981) paper focuses on the effects of multi-center and localized amenities, and concludes that these factors also play an important role on the fluctuation of housing prices.

Another important use of the hedonic model is based on its association with environmental quality. One interesting subject is whether house prices positively associate to environmental quality. Chesire and Sheppard (1995), Freeman (1979), Boyle and Kiel (2001) and Din, Hoesli and Bender (2000) conduct studies in this area.

Hedonic pricing method is also applied in other areas, such as racial difference in housing prices, market discrimination and saving patterns through the interpretation of coefficients. Kain and Quigley (1972) use the hedonic model to investigate whether racial discriminations exist in the housing market so that black households pay more than white households do for identical housing services. Chambers (1992) explores the house prices differentials faced by various races.

Numerous studies of the hedonic model have been undertaken in respect of policy making. The hedonic model is useful in the cost-benefit analysis of housing policies, such as subsidy and rent control; see Olsen and Barton (1983), Buchel and Hoesli (1995) and De Borger (1986).

Moreover, many studies have tried to relate the hedonic prices to the supply-demand model and capture demand parameters for individual housing attributes. Awan, Odling-Smee and Whitehead (1982) paper is a good example on this subject.

C. Cutting-edge development of the hedonic model

Even though many works on the hedonic model have been done over the past few decades, further studies are still in progress.

One fundamental strand is to collect more and better data. Malpezzi and Mayo (1994) paper makes some suggestions on how to improve housing data collection.

Regarding the functional form of the hedonic models, several approaches have been constructed to improve the hedonic model. To the traditional hedonic model, the data is expected to follow certain probability distribution (i.e. parametric approach), which is not always conceivable. To alleviate the restrictions of the parametric form, semiparametric and nonparametric approaches have been exploited by some studies with the representative papers such as Anglin and Ramazan (1996), Mason and Quigley (1996), Meese and Wallace (1991) and Pace (1993). The nonparametric model, on the contrary, does not require any strict distributional assumptions.

A development of the traditional hedonic model is the application of the Bayesian technique on the hedonic estimation, as illustrated by Atkinson and Crocker (1987) and Knight, Hill and Sirmans (1992). The Bayesian method transfers the prior estimates of the price of housing attributes into posterior estimates to increase the accuracy of the estimation and solve the collinearity problem.

Another significant development is the use of spatial data and spatial autocorrelation; see Basu and Thibodeau (1998, 2002) and Pace and Gilley (1997). Specifically, the spatial autoregression model conducted by Pace and Gilley (1997) incorporates the area configuration of the data so that the estimated errors are reduced by 44% relative to the simple regression by the conventional hedonic model.

III: Hedonic Analysis Of Housing Attributes Effect

D. Model specification

The hedonic regression employs a semi-logarithmic functional form, which is based on the hedonic model of Thibodeau (1995). The model regresses the log of the dependent variable on a linear combination of independent variables, as specified by the functional form below.

$$V = e^{X\beta + \varepsilon}$$

where,

V = a vector of the housing unit values

X = a matrix of three categories of housing attributes, including household characteristics, dwelling variables and neighborhood quality variables etc. The components of housing characteristics are defined in Exhibit 15.

β = a vector of estimated coefficients of independent variables

ε = the model errors with the assumption that $\varepsilon \sim N(0, \sigma^2 I)$, where σ^2 is the residual variance and I is the identity matrix.

Based on the previous literature of hedonic study, the price of a good is determined by the marginal value of the features of this good. The linear regression form of the hedonic pricing model is specified as below.

$PValue_{ijt} = \beta_{i0} + \beta_{i1}Income_{ijt} + \beta_{i2}Crowding_{ijt} + \beta_{i3}Year2 + \beta_{i4}Year3 + \beta_{i5}Year4$[1] $+$

$\beta_{i6}Size_{ijt} + \beta_{i7}Rooms_{ijt} + \beta_{i8}Detached_{ijt} + \beta_{i9}Age_{ijt} + \beta_{i10}Garage_{ijt} + \beta_{i11}Basement_{ijt} +$

$\beta_{i12}Heating1_{ijt} + \beta_{i13}Heating2_{ijt} + \beta_{i14}AC_{ijt} + \beta_{i15}BuildingProblem_{ijt} +$

[1] To metropolitan areas with only three time periods of observations, the regression does not contain the $\beta_{i8}YEAR4$ term.

$\beta_{i16}HallwayProblem_{ijt}$ + $\beta_{i17}LackFeature_{ijt}$ + $\beta_{i18}Breakdown_{ijt}$ + $\beta_{i19}Quality_{ijt}$ + $\beta_{i20}Rat_{ijt}$ + $\beta_{i21}Abandon_{ijt}$ + $\beta_{i22}Junk_{ijt}$ + $\beta_{i23}Crime_{ijt}$ + $\beta_{i24}Noise_{ijt}$ + δ_{ijt}

(1)

Previous studies reveal that some macroeconomic factors, particularly unemployment rate, make an important impact on residential housing prices. To further explore the determination of the housing unit values, the housing prices are regressed on the selected housing attributes as well as unemployment rate with the mixed effect model as below.

$$Y_{ijt} = \beta_{i0} + \sum_{25}^{k=1} \beta_{ik} x_{kijt} + \theta z_{it} + e_{ijt}$$

In this model, x_{ijt} is the household level measurement of housing attributes, and z_{it} is the city-level measurement of unemployment rate. The underlying assumptions on the random effect factor are $z_i \sim N(0, \sigma_z^2)$, x_{ijt} and z_{it} are uncorrelated and z_{it} is not correlated with the errors. The functional form of the mixed effect model is

$PValue_{ijt}$ = β_{i0} + $\theta Unemployment_{it}$ + $\beta_{i1}Income_{ijt}$ + $\beta_{i2}Crowding_{ijt}$ + $\beta_{i3}Year2$ + $\beta_{i4}Year3$ + $\beta_{i5}Year4$[2] + $\beta_{i6}Size_{ijt}$ + $\beta_{i7}Rooms_{ijt}$ + $\beta_{i8}Detached_{ijt}$ + $\beta_{i9}Age_{ijt}$ + $\beta_{i10}Garage_{ijt}$ + $\beta_{i11}Basement_{ijt}$ + $\beta_{i12}Heating1_{ijt}$ + $\beta_{i13}Heating2_{ijt}$ + $\beta_{i14}AC_{ijt}$ + $\beta_{i15}BuildingProblem_{ijt}$ + $\beta_{i16}HallwayProblem_{ijt}$ + $\beta_{i17}LackFeature_{ijt}$ + $\beta_{i18}Breakdown_{ijt}$ + $\beta_{i19}Quality_{ijt}$ + $\beta_{i20}Rat_{ijt}$ + $\beta_{i21}Abandon_{ijt}$ + $\beta_{i22}Junk_{ijt}$ + $\beta_{i23}Crime_{ijt}$ + $\beta_{i24}Noise_{ijt}$ + ζ_{ijt} (2)

The subscript are defined as below,

i: city

j: household

[2] To metropolitan areas with only three time periods of observations, the regression does not contain the $\beta_{i8}YEAR4$ term.

t: time (from 1985 to 2004)

E. Estimation results

The analysis uses Annual Housing Surveys (AHS) for 26 metropolitan statistical areas (MSAs) data from 1985 through 2004, and the 26 cities (as shown in Exhibit 14) are grouped into three clusters according to the results of Chapter 1 (see Exhibit 7). The descriptions of variables in the hedonic regression model are listed in Exhibit 15. The variables include housing attributes, household income, and the macroeconomic variable, unemployment rate.

The results of the linear regression model and mixed model estimates are presented in Exhibit 16 and Exhibit 17 respectively. We analyze the estimation results on the aspects of prediction power, precision of coefficient estimates and variation of the estimates to examine the performance of the two hedonic models respectively. Exhibit 16 shows that F statistics in three regressions are 127.61, 91.6 and 123.49 respectively with p-value all less than 0.001. Therefore, linear regression model makes a good fitness. Another important measure of the explanatory power is the R-square statistics, and the results present moderate R-square values in all three linear regression models. The linear regressions of all three clusters do a good job in capturing the effects of housing attributes on property values. In the mixed model, Chi-square statistic is the result of the null model likelihood ratio test, which indicates a significant improvement over the null model consisting of no random effects and a homogeneous residual error. In Exhibit 17, all three Chi-square statistics are statistically significant, so that mixed model is also a good fit to the data.

The signs and magnitudes of individual coefficient estimates represent the effects of the variables on the property values, and the significances of the effects are examined by t statistics and p-value. We also investigate the sectional pattern of each factor to see if the effects of housing attributes are subject to regional difference.

The analysis results in Chapter One indicate that the property values of Cluster One cities are generally low and the property values of Cluster Three cities are higher. This implication is supported by the intercept values of three regressions. The intercepts of three linear regressions are 9.22424, 9.68102 and 10.15587 respectively, which means that without the effects of other factors, the baseline price of cluster one is $5870.85 lower than that of cluster two, and $15601.80 lower than that of cluster three. The intercept terms of mixed effect models display the same pattern.

Ideally high household incomes lead to high demand for housing, so that the real estate prices rise. The housing value presumably moves in the same direction with the household income. As shown by the results of both the linear regression models and the mixed effect models, the coefficient estimates of *Income* has a positive sign and the variable is generally significant at 10% confidence level.

Next factor considered in two models is the crowding level of the dwelling, which is the ratio of the number of persons in the household to the total number of rooms. According to Malpezzi, Ozanne and Thibodeau (1980), this factor is negatively connected to the property value for the reason that crowded dwellings depreciate fast due to the heavy use. Most results in two models are accordant to this presumption but some coefficient estimates are statistically insignificant.

The estimate shows that the year dummy variables have significant effects on housing values. For example, in the linear regression model, the coefficient estimates of *Year2* in cluster one is 0.16227, implying that the dwellings surveyed in the first time period have 0.16227 unit higher logged housing value than the dwellings surveyed in later periods. The coefficient estimates are positive in all three clusters, and the coefficient estimates in later periods are always greater than those in early periods. This suggests an increasing trend of residential housing prices over time in large cities. At the same time, it can also be due to the development, amenities and attractiveness changes of the metropolitan area over time. Large metropolitan areas generally display high economic developments, offer more job opportunities and attract people from other areas. As an effect, rapid increases in population result in high demands on housing, but the housing supplies do not increase at the same pace due to the limited land resources. In this case, unbalanced supply-demand changes lead to an appreciation of residential housing prices.

Unit size typically makes a positive impact on the property value. Controlling for other effects, housing units with large sizes are more attractive to households, and the associated housing price is high. Same as our expectation, the results exhibit a positive coefficient of the unit size variable in all three clusters.

Theoretically dwellings with more rooms have higher values, so that the coefficient of the variable, *Rooms*, should have a positive sign. All coefficient estimates in both models have a correct sign and they are statistically significant at 10% confidence level.

Variable *Detached* is a structural variable. Housing units in a detached building conceptually have higher values then those in other structures. The estimation results of both models are consistent with our hypothesis, and they show that the effects of structural types on housing values are significant in Cluster One and Cluster Three.

Building age is usually a big concern to households. Aged buildings are more likely to have low quality and potential quality problems. In this case, the increase in building age will result in the decrease of property valued. Expectedly, the coefficient estimates of *Age* are negative in all clusters in our results.

Garage and basement are also two important components to dwellings. Theoretically housing units with garage and basement have higher property values. The coefficient estimate of the *Garage* variable is largely consistent with our expectation and the effect of this variable on housing values is significant. The coefficient estimates of *Basement* demonstrate an anticipated result with all three coefficient estimates having positive signs and no regional pattern is presented.

The heating equipment is also critical to housing prices. *Heating1* denotes the dwellings with central heating equipments. *Heating2* represents the built-in heating facilities in the housing units. Heating equipment is a plus to properties in the North, while housing values in other areas should be loosely connected to this factor. The conjecture is justified by the estimate results. Though the coefficient estimates of *Heating1* and *Heating2* are positive, the estimates are statistically insignificant in some regressions.

Another dwelling equipment factor is *AC*, which stands for the dwelling with central air conditioning equipment. Ordinarily the property values of housing units are positively associated with the central air equipment. The positive coefficient estimates of *AC* in both models support this proposition.

Variables *BuildingProblem*, *HallwayProblem*, *LackFeature* and *Breakdown* represent various problems related to the dwelling quality, and the existence of quality problems drives down the property values. Dwellings in any cluster are negatively impacted by these effects, so that all coefficient estimates of these four variables have negative signs.

The quality of neighborhood is an important concern to households in terms of housing evaluation. The data ranks the neighborhood quality by numbers from 1 to 10, and larger values specify better neighborhood around the dwellings. All estimation results of *Quality* variable in two models have positive signs and are statistically significant.

The measurements of neighborhood quality, *Rat*, *Abandon*, *Junk*, *Crime* and *Noise* are also taken into consideration. These variables signify if the neighborhood has rats, abandoned buildings, junk, crime or noise or not. Certainly these components make negative impacts on housing values so that negative signs are expected on the coefficient estimates. The results shows that all coefficient estimates have correct signs and the magnitudes of the effects fluctuate across regions to certain extend. For instance, some metropolitan areas may be extraordinarily sensitive to the presence of crime, resulting in a large impact of crime rate on property values.

Theoretically, the unemployment rate is expected to be negatively related to residential housing prices. When there are fewer job opportunities in the city, people are less affordable to buying a house, so that the housing demand is depressed. The housing supply is relatively inelastic in the short run. As an effect, the equilibrium marketing prices of housing will go down. Since the feature of the unemployment rate data, the unemployment factor is included in the mixed effect model as a random factor. The negative coefficients of the unemployment rate factor in the mixed effect model justify the negative association between the unemployment rate and the housing prices.

In both models, explanatory variables exhibit expected signs and most of the estimates are statistically significant. We observe considerate differences of housing attributes effects between clusters, which verifies the conclusion of the cluster analysis in Chapter One. Exhibit 18 shows the results of the pooled data with the linear regression model. The coefficient estimates present similar features, verifying the relationship between property values and housing attributes. In next section, we will further investigate the cross-sectional differences of Hedonic coefficient estimates.

IV: Analysis Of Hedonic Coefficient Estimates

A. F-test of the equality of the Hedonic coefficient estimates

Thus far we have reported the effects of housing attributes on property values in three clusters defined by Chapter One. The wide variety in the housing attribute effects across clusters is displayed in Section III. In this section, we explore if the cross-section variability of housing prices is associated with the inequality of marginal effects of housing attributes among clusters. F-test is used to assess the equality of the hedonic coefficient estimates.

In the panel data setting, the F-test provides the easiest and the most robust way to test the cross equation restrictions on the parameters in different equations using system OLS without assuming homoskedasticity or serial independence of the errors. The multiple hypotheses on the coefficient estimates of the hedonic regressions in Section III are:

H_0: $\beta_{1,1} = \beta_{2,1} = \beta_{3,1}$ (i.e., the coefficient estimates on variable, *Income*, are identical across three clusters)

H_1: the equation above doesn't hold

H_0: $\beta_{1,2} = \beta_{2,2} = \beta_{3,2}$ (i.e., the coefficient estimates on variable, *Crowding*, are identical across three clusters)

H_1: the equation above doesn't hold

…

H_0: $\beta_{1,24} = \beta_{2,24} = \beta_{3,24}$ (i.e., the coefficient estimates on variable, *Noise*, are identical across three clusters)

H_1: the equation above does not hold

Exhibit 19 illustrates the result of the F-test on the equality of the hedonic coefficient estimates across clusters. The F-statistics of the coefficient estimates of the variable, *Income*, is 1.19, and the p-value is 0.3029, greater 0.1. It implies that at 10% confidence level, we fail to reject the null hypothesis and can draw the conclusion that the coefficient estimates of *Income* are not substantially different across clusters. In other words, the effects of household incomes on property values do not differentiate significantly among clusters. In general, the effects of time, unit size, unit age, and some quality variables are significantly different across clusters, which is consistent with our supposition that the graphical pattern of housing prices are associated with the regional disparity of housing attributes effects.

V: Conclusion

This chapter constructs a hedonic regression analysis on the three clusters of U.S residential housing markets to capture the effects of housing attributes on owner-occupied housing prices. The analysis results show that the housing attributes significantly affect the dwelling prices, and geographical disparity of housing prices across regions is attributable to housing features effects. Owner-occupied housing has the nature of consumption good. The utility of housing services to households depends on the margin value of housing attributes to housing consumers. Therefore, housing attributes can explain the determination of dwelling prices to a large extend. Residential housing is also treated as an investment good by some households. In this case, the value of residential housing relies on the potential values and returns of the housing attributes as an investment.

Furthermore, the F-test is used to assess the cross-section equality of housing attributes effects. Based on the analysis results, the margin values of some housing attributes are substantially different across clusters.

Chapter Three: Analysis Of Tenure Choice And Timing Of Tenure Transition In U.S. Residential Housing Markets

I: Introduction

The tenure choice between renting and owning, as well as the timing of the tenure transition from renting to owning long has been an interesting topic to economists. The most controversial questions on this respect are: What affect household's decision between renting and owning a house? When do the renters decide to make the change from renting to owning?

Housing serves both as consumption good and an investment good, therefore the tenure choice and the tenure transition are linked to many factors. Previous studies show that for families concentrating on the consumption consideration of housing, the step to owner-occupation primarily depends on household income, family composition change and the demographic characteristics of the household.

Some housing economists stress that when households make the decision of tenure choice, they consider the investment aspect of housing. They treat housing stock as a hedge against the inflation and the risks of financial investment. In this case, the tenure change is more related to the fluctuations of house prices, inflation rate, unemployment rate etc. Therefore the household's tenure-change decision is based on the potential value increase in housing assets while resell the dwelling.

This study extends the past research by focusing on the relationship between household tenure choice and household demographics, as well as the time-sequence of household demographics on timing of the household tenure transition. In particular, we try to answer two questions. Is the household tenure choice substantially determined by household characteristics and demographics? Do household characteristics and demographics affect the timing of household's move from renting to ownership?

The rest of the paper is structured as below. The second section is a literature view on the representative papers of tenure choice and the timing of tenure transition. Section III investigates how the rent-own choice is influenced by household characteristics and demographics as well as some other factors. In the fourth section, we follow the descriptive interpretation of household tenure choice with a proportional hazard regression analysis to assess the timing of tenure transition in the light of household characteristics and demographics. The conclusion is in Section IV.

II: Literature Review

The original study on housing tenure choice is established by Henderson and Ioannides (1983). They conduct the utility maximization model to derive the housing tenure choice of households.

Since the initial investigation on housing tenure choice by Henderson and Ioannides (1983), there is an extensive literature in this area. And they generally fall in several major aspects.

Some studies explain the determination of housing tenure choice with the conception of user-cost, and suggest that home-owning provides hedge against fluctuation in future housing costs. The typical use-cost approach compares the costs of owning and renting, and examines the effect of various housing expenditure components, such as the interest rate, the inflation rate, taxes on household's decision. Rosen *et al.* (1984) analyze the relationship between housing tenure choice and housing price uncertainty under the impact of income tax. They find that housing price variance substantially reduce the homeownership rate. The capital gain taxation increases the homeownership ratio, while property tax reduction and interest payment deductions depress the aggregate proportion of homeowners. They conduct an empirical test to year 1956 to 1979 data. The results support their theoretical model and partially explain the fluctuation of housing ownership rate in the late 1970s.

Hilber (2005) uses the American Housing Survey (AHS) data to test the effect of neighborhood externality risk on the homeownership probability. The test relies on location-specific dummies to capture the variations in neighborhood qualities across markets.

The author argues that the neighborhood externality risk largely increases the housing investment risk and reduces the homeownership rate.

Sinai and Souleles (2001) take into account both the rent risk and housing price risk. They assert that home owning provides a way to hedge against the risk of fluctuations in future rent payments, but it is also affected by potential asset price risk. As an effect, the housing demand increases with rent variance, controlling for house price variability. The hedging function of home owning is capitalized into the house prices; so that the price-to-rent ratio is high in places with high rent variance.

Another group of housing studies emphasize how income uncertainty determines the housing tenure choice. A few studies are more explicit. For example, Haurin (1991) explores the relationship between income uncertainty and home ownership. Specifically, the paper derives income uncertainty from the coefficient of income variation over time, and assesses whether the estimated income risk affect the probability of homeownership. The paper finds that increasing income risk reduces the likelihood of home ownership.

Robst *et al.* (1999) uses the University of Michigan's Panel Study of Income Dynamics (PSID) data to portray the association of household income to homeownership. In this paper, income uncertainty is measured by the variation in residual earnings from three different earnings functions. The results from three earnings function are consistent, and the study verifies that income variability has strong implication on housing purchase decision.

Davidoff (2006) aggregates both the housing costs and household income to discuss the tenure choice. He finds that covariance of housing costs and income significantly affects the housing investment, so that unsecured household income and housing price variability tend to distort investment decisions of households to a large extent.

The behavior model integrates the household characteristics to explain housing tenure decision. For instance, Rudel, T.K. (1987) employs the Annual Housing Survey data to describe the relationship between household demographic characteristics and the transition from rented to owner occupied housing. The author claims that household composition information, such as household size, family growth and wealth help to predict the homeownership rate.

Henderson and Ioannides (1987) build an investment-consumption model of housing demand with various household capital constraints and demographic characteristics of owner-occupiers. They conclude that for both renters and homeowners, the housing demand is wealth inelastic and price inelastic. Households with incomes tilted toward the present are more likely to own. The impact of age and education on tenure choice is observed. Younger people with less education and lower income have a smaller intension to own a house. Moreover, large transitory income also makes a positive effect on the probability of housing purchases.

Clark & Deurloo (1994) present a longitudinal study on housing tenure choice and demonstrate that tenure changes are closely linked to family composition changes of households. Specifically, the increasing stability of forming a couple or a family plays an important role in household's decision between renting and owning. The Clark & Deurloo (1994) paper indicates that the housing tenure change is also influenced by location and economic contexts.

Ample studies have been made on the effect of credit constraints on the tenure decision of households. Linneman and Wachter (1989) test the impact of wealth and income constraints on homeownership propensities. They find that households subject to wealth or income reduce homeownership propensities, and wealth constraint plays a more important role. Contractually, easing the access to mortgage credit can result in an observed increase in the owner occupancy rate.

Ortalo-Magne and Rady (1999) use the life-cycle model to investigate the boom-bust cycle in British residential real estate market from the mid-1980s to the mid-1990s. They claim that the availability of mortgage finance determines the owner occupancy rates across different age groups, and the credit constraints are particularly crucial to the owner occupancy rate of young households. The downturn of residential housing markets can be illustrated by negative income and credit market shocks, while the housing market boom in the early 1980s is associated with financial deregulation.

Chiuri and Jappelli (2003) examine the trend of owner occupancy rates to different age groups across 14 OECD countries. They observe the negative relationship of homeowner occupancy and credit constraints across age groups.

Past research has explored the timing of tenure transition in a limited manner and the representative works are Plaut (1987), Clark and Deurloo (1994), Moriizumi and Naoi (2006). The theoretic work by Plaut (1987) emphasizes the dual nature of housing as consumption good and an investment good. Therefore, the tenure transition is affected by a variety of factors, like non-housing wealth, mortgage interest rate, housing rental rate, return on financial assets, and variance of housing prices. The impact of each component on timing of the transition is discussed by the paper.

Clark and Deurloo (1994) generate an empirical analysis on the tenure change and timing of tenure transition. They conclude that on the micro level, the tenure change is closely linked to the change of family composition, especially to the family type change and income change. On the macro level, the tenure change is affected by locational factors and economic contexts, e.g. the price and amount of new construction, mortgage rate etc.

Moriizumi and Naoi (2006) explore the relationship between income and the timing of homeownership in Japan using the discrete-time hazard model. They find that income uncertainty defers the household's decision on the change to homeownership. However, if the private transfer of wealth is available to the household, the negative effect of income uncertainty on the timing of homeownership will be markedly alleviated.

III: Analysis Of Household Tenure Choice

A. Tenure Choice Model

Henderson and Ioannides (1983) built a two-period utility model to discuss the household's choice between renting and owning. They concluded that with the existence of rental externality, owning dominates renting without any other disturbances. The paper also uses the two-period behavior model to analyze the relationship between tenure choice and wealth. Housing has the dual role of consumption good and an investment good, so that households' tenure choice is subject to both their consumption demand on housing service and their investment demand on housing stock. They show that households with less wealth or more wealth in the early stage of their lifetime tend to rent. Moreover, at certain income level, due to the rental externality, households are likely to distort their investment and consumption decision and own rather than rent.

Based on Henderson and Ioannides (1983), in a simple tenure choice model, the household makes the decisions on whether to rent or own a house to maximize the expected discounted value of the multi-period future utility, without the effect of mortgage rate, property tax and moving cost.

Let $U(c, h)$ be the utility function of the household, where c is the consumption of non-housing goods and services, and h is the measure of the quality of housing services.

Owners choose the optimum value of $\{c^*, h_c^*, s^*, u^*\}$ to maximize their lifetime utility,

$$U(c, h) + V(w) \tag{1}$$

Subject to, $h = h_c f(u);\ f'>0, f''<0$

$y_1 = c + ph_c + s$

$w = y_2 + s(1+r) + ph_c - T(u)h_c;\ T'>0,\ T''>0$

U(.) is the utility derived from the consumption of goods and services *(c, h)* in current period, and *V(.)* is the utility derived from the future flows of wealth *w*. y_1 is the income in the current period; h_c is the housing stock owned and *u* is the rate of utilization; *s* is the saving of the household; *r* is the market rate of interest or investment return; *p* is the unit price of housing stock.

Renters face the choice variables $\{\tilde{c},\ \tilde{h}_c,\ \tilde{s},\ \tilde{u}\}$ to maximize their lifetime utility,

$$U(c, h) + V(w) \tag{2}$$

Subject to, $h = h_c f(u),\ f'>0,\ f''<0$

$y_1 = c + Rh_c + s$

$w = y_2 + s(1+r) - \tau(u)h_c;\ \tau'>0;\ \tau''>0$

In equilibrium market, there are owners of rental housing who hold housing equity as investment. Their forgone interests on housing equity must equal the housing profits.

$$(1+r)R - rp = T(u^*) - \tau(u^*) \tag{3}$$

From the first order conditions of (1) and (2), we have

$$f'(u)\frac{U_2(u)}{U_1(u)}(1+r) = T'(u)\text{, where } u = u^* \text{ or } \tilde{u}$$

The left side of the equation is the marginal benefits of increasing housing utilization *u* to households, and the right side is the cost of housing utilization. To renters, if the rate of utilization of owning u^* generates higher marginal benefit than that of renting $\tilde{u}$, they will choose to own. Otherwise, they will keep renting. In the market equilibrium, the landlord's u^* is equal to $\tilde{u}$. The rate of utilization *u* is subject to household features and demographics, so that household tenure choice is determined by household characteristic and their demographics. This section depicts the relationship between household tenure choice and household characteristics.

B. Data

The data used in the empirical analysis came from the American Housing Survey (AHS), a national survey of more than 50,000 repeatedly evaluated homes and their residents. The U.S. Bureau of the Census conducts the American Housing Survey on a regular basis for the U.S. Department of Housing and Urban Development. This analysis is based on the Metropolitan Statistical Areas (MSAs) survey in seventeen cities (Atlanta, Buffalo, Cleveland, Columbus, Hartford, Indianapolis, Kansas City, Memphis, Miami, Milwaukee, New Orleans, Oklahoma City, Pittsburgh, Portland, San Antonio, Seattle, St. Louis) in year 2002 and year 2004. All housing units in these cities with complete data on the variables used in the models are included in our sample (n = 66,895).

C. Model Specification

We investigate how housing tenure preferences depend on household features and demographics in the surveyed metropolitan areas using the logistic model. A nonlinear estimation is established to measure the probability that a household will choose to own or rent, in order to assess how well various response variables can explain household's tenure choice decision.

The dependent variable is the tenure status, which is defined as whether a home is owned or rented by its residents. The explanatory factors include a large array of variables, including regional information, unemployment rate, household credit constraint, household demographics. All variables represented in the analyses are exhibited in Exhibit 20, and Exhibit 21 displays the basic statistics of these variables.

The homeownership model includes regional information (South, Midwest, Northeast and West) as control variables. Differentiating among regions is necessary as Chapter One indicates that locational effect plays an important role in household's decision of housing consumption. As shown in Exhibit 21, among all sampled homes, 46,132 have home-owned housing, which counts for 68.96%. 35.71% of the homes are in Midwest areas, 17% live in the Northeast region, and West and South have 12.61% and 34.68% of the sampled homes.

The effects of economic variables, such as income uncertainty on household tenure choice are highly emphasized by previous studies, such as Moriizumi and Naoi (2006), Rosen *et al.* (1984), Haurin (1991), so these factors are incorporated in the model. Specifically the unemployment rate of each surveyed metropolitan area is used to proximate the income uncertainty as suggested by Moriizumi and Naoi (2006).

Clark and Deurloo (1994) highlight the importance of household characteristics to the analysis of tenure changes. This model contains household income and household composition variables such as, the primary family's race, crowding, age, education level, marital status and gender.

Based on our implicit assumption of dichotomy in housing tenure choice decisions, the household faces a discrete choice between renting and owning. The dependent variable, household tenure status is denoted by H, and the dual possibilities of H are:

H=1 if household chooses to own the home

H=0 if household chooses to rent the home

The probability of owning the home is represented by *P(H=1)*. Given the vector of factors for the probability of homeownership discussed earlier (X_K), the logistic model for the probability is

$$P(H = 1 : X_K) = L(H = 1 : X, \beta) = \frac{1}{1+e^{-z}}$$

$$z = \sum \beta_K X_K$$

where

L(.) = the likelihood function

X = vector of measured explanatory variables

β = vector of coefficients associated with the explanatory variables (X)

z is a function of explanatory variables (X) that affect tenure choice, defined as below

$$z = \beta_0 + \beta_1 *Midwest + \beta_2 *Northeast + \beta_3 *West + \beta_4 *Unemployment + \beta_5 *Income + \beta_6 *Race + \beta_7 *Crowding + \beta_8 *Age + \beta_9 *Education + \beta_{10} *Marital + \beta_{11} *Gender$$

D. Results

The result of the estimation of homeownership rate, *P(H=1)* is presented by Exhibit 22. The likelihood ratio test, efficient score test and Wald's test assess the joint significance of the explanatory variables. The results show that all three test statistics have p-value less than 0.001, so the logistic model makes a good fit. The R-square value is 0.4346, which indicates that the response variables are attributable to a considerable proportion of variation in the probability of homeownership. Most of the coefficients have signs consistent with our expectations.

The coefficient estimates of region variables indicate that households in the West are the least likely to own houses among all regions and people in the South have the highest intension to owner-occupy. Families in the Midwest have a similar probability of homeownership to households in the South, while people in the Northeast are less likely to own. However, the region effect is not statistically significant in Midwest areas.

Unemployment has negative coefficient estimate, indicating that the unemployment rate damps people's willingness to own houses. In the areas with high unemployment rate, it's less easy for people to find decent jobs. Moreover, unemployment rate usually associates with job insecurity and high turnover rate. Both effects lead to a low probability of owner-occupation.

The results also show that high incomes encourage the homeownership, and families with more incomes are more likely to own homes. This result supports Henderson and Ioannides (1983) in that households with more wealth incline to owning than renting.

The household compositions also have important impacts on the likelihood of homeownership. For instance, the variable *Race* has the coefficient estimate of -0.5613, implying that Black families have a lower probability of homeownership compared to households of other races.

Variable, *Crowding* is a measurement of the number of persons per room. Crowded families are usually low-income families, so that negative coefficient estimate of *Crowing* specifies a negative relationship between the crowding level of the household and the probability of homeownership.

The coefficient estimate of household demographics variable, *Age*, has a positive sign, suggesting that older people tend to own a house rather than rent. *Education* also has a positive coefficient estimate, so that high education level increases the current magnitudes of housing purchases.

Marriage triggers the tenure choice of owning over renting, so the observed coefficient estimate of *Marital* is positive. Gender factor is also influential to household tenure choices. Males are more likely to purchase a house.

IV: Analysis Of Timing Of Tenure Transition

A. Data and Model Specification

The data for the timing of tenure transition analysis is also abstracted from the American Housing Survey (AHS). The AHS dataset in year 2002 and 2004 comprise information about the recent home moving. The year and month households moved into recent housing units are provided, and the year and month of previous move before the current one are also contained. Moreover, an indicator variable identifies if the previous housing unit was owned or rented by the household.

To depict the timing of a move from rent to own, the spell duration until a home purchase is defined as the interval from the start of the state as a renter until the eventual move to purchase a home. Thus the spell of survival time as a renter household is calculated by the difference between the time of previous state (renting) and the time of current state (owning) and we focus on the households who rented the previous housing unit and own the current dwelling.

The timing of tenure transition is modeled by the Cox proportional hazard model, which is generally used to explain the effect of explanatory variables on survival times in survival analysis approach. The proportional-hazard model can be expressed as

$$h(t) = h(t; X, \beta) = h_0(t)exp(z)$$

where

$h(t)$ = the hazard function of the survival time

$h_0(t)$ = the baseline hazard function

X = vector of measured explanatory variables

β = vector of coefficients associated with the explanatory variables (X)

z is a function of explanatory variables (X) that affect survival time, defined as below

$z = \beta_0 + \beta_1*Midwest + \beta_2*Northeast + \beta_3*West + \beta_4*Unemployment + \beta_5*Income +$

$\beta_6*Race + \beta_7*Crowding + \beta_8*Age + \beta_9*Education + \beta_{10}*Marital + \beta_{11}*Gender$

Then the hazard function is

$$S(t, X) = S_0(t)^{\exp(z)} \; S(t, X) = S_0(t)*exp(z)$$

where

$S(t)$ = the hazard function

$S_0(t)$ = the baseline hazard function

$$S_0(t) = \exp(-\int_0^t h_0(u)du)$$

The estimate of coefficients (β) is obtained by maximizing the partial likelihood function,

$$L(\beta) = \prod_{Y_i uncensored} \frac{\exp(X_i\beta)}{\sum_{Y_j \geq Y_i} \exp(X_j\beta)}$$

The hazard function represents the notions of risk. A small value of the hazard implies that the expected length of time until the event occurs will be short. In this analysis, a small value of the hazard corresponds to a short duration from rent to own. The coefficients of the explanatory variables illustrate the changes in the log of the hazard caused by each unit changes in the response variable, controlling for other variables.

The data set consists of 15,646 observations from seventeen metropolitan areas in year 2002 and year 2004. The dependent variable is the time duration for tenure move from rent to own. Similar to the analysis of tenure choice, the independent variables are comprise of several groups, including regional factors, unemployment rate, household incomes and household demographics. A positive coefficient suggests that the variable accelerates the move to the ownership, and a negative coefficient means that the factor defers the move to homeowner status.

B. Results

The empirical result of the proportional-hazard regression model for move from rent to own is shown by Exhibit 23. The test for the global null hypothesis is to test the general fitness of the model, and the null hypothesis is that there is no difference between the survival times of the owning and renting groups. All three test statistics show that the survival curves for the owning and renting groups are not identical. Overall, the proportional-hazard model generates a good fit to the analysis.

In the proportional-hazard model, the hazard ratio (i.e. risk ratio) of an explanatory variable is defined as the exponentiation of the regression coefficient for this variable. For instance, the hazard ratio of *Race* is 0.695, implying that the hazard function for *Race = 1* (Black family) is smaller than that for *Race = 0* (Non-black family). In other words, the duration from rent to own for black families is longer than that for non-black families.

Based on the results shown in Exhibit 23, regional variables make a significant impact on hazards to a large extend. All three variables, *Midwest*, *Northeast* and *West* have negative coefficient estimates and hazard ratio less than one, implying that all three regions have long durations from rent to own compared with the south region. Moreover, the *Northeast* variable has the smallest hazard ratio, so that the expected length of time for tenure transition in Northwest is the longest among all regions.

Unemployment has negative coefficient estimate and hazard ratio less than one, implying that high unemployment rates depress the opportunities for movement.

The coefficient estimate of *Income* is 0.34455 and hazard ratio is 1.411. High-income households are 1.411 times more likely to change from rent to own that that of the low-income households. Income plays an important facilitating role in the likelihood of tenure change from rent to move largely on the respect of affordability of owner-occupancy.

The sign of *Crowding* is negative, so that crowded households have low willingness to own, and this effect may associate with household income and wealth as discussed in previous section.

Age factor is also influential to the likelihood of tenure transition. The hazard ratio of *Age* is 0.973, suggesting that the tendency for old people to move from renting to owning is lower than young people.

High education increases the propensity for renters to move to ownership as the coefficient estimate of *Education* is positive. The close link between buying a house and education level may be substantiated by income effect. Generally individuals with high education have high income, which is also a favorable condition for moving to homeownership.

Consistent with our expectations, marital status tends to be positively related to the propensity to move to ownership. As discussed in Clark and Deurloo (1994), homeownership decision is directly linked to people's positions and changes in their family life cycle. Family formation is a critical stimulant to housing transitions.

The hazard ratio of *Gender* is 1.069. If the head of the family is male, the household's decision of owning a house is simulated.

The results above imply that the decision to move from rent to own is determined by both economic factors, like job market, and regional factors, and the household compositions.

V: Conclusion

This chapter presents an analysis to the household tenure choice and timing of tenure transition in the residential housing markets of seventeen metropolitan areas in year 2002 and 2004. The study incorporates not only the widely used factors such as household income, unemployment rate, but also considers more sophisticated information such as housing attributes and household compositions to provide a enriched understanding to the tenure choice and tenure transition. The results are able to show that household demographics help to demonstrate household tenure choice decisions and moving decisions. Specifically, household demographics are closely linked to the probabilities of owner-occupancy, and are critical to household's decision of a move to ownership.

Appendix

Exhibit 1: Average Prices Of Owner-Occupied Housing (In 2004 Dollars) And Percent Change In Housing Prices Of 26 Metropolitan Areas, 1985 - 2004

Metropolitan Areas	Average Price of Owner-Occupied Housing ($)	Percentage Change of Housing Prices
Atlanta	219,683	155%
Baltimore	180,177	48%
Birmingham	131,716	69%
Boston	206,609	72%
Buffalo	92,146	85%
Cincinnati	118,165	85%
Cleveland	125,713	178%
Columbus	193,362	301%
Hartford	211,029	91%
Indianapolis	183,697	191%
Kansas City	136,660	141%
Memphis	174,997	227%
Miami	150,556	99%
Milwaukee	120,162	146%
Minneapolis	131,730	82%
New Orleans	116,914	117%
Oklahoma City	154,034	244%
Pittsburgh	132,649	143%
Portland	210,864	236%
Providence	143,247	9%
Rochester	110,233	48%
San Antonio	148,204	90%
San Francisco	393,530	145%
Seattle	271,276	249%
St. Louis	132,212	191%
Washington	269,173	94%

Exhibit 2: Average Prices Of Owner-Occupied Housing (In 2004 Dollars) Of 26 Metropolitan Areas In Different Time Periods From 1985 To 2004

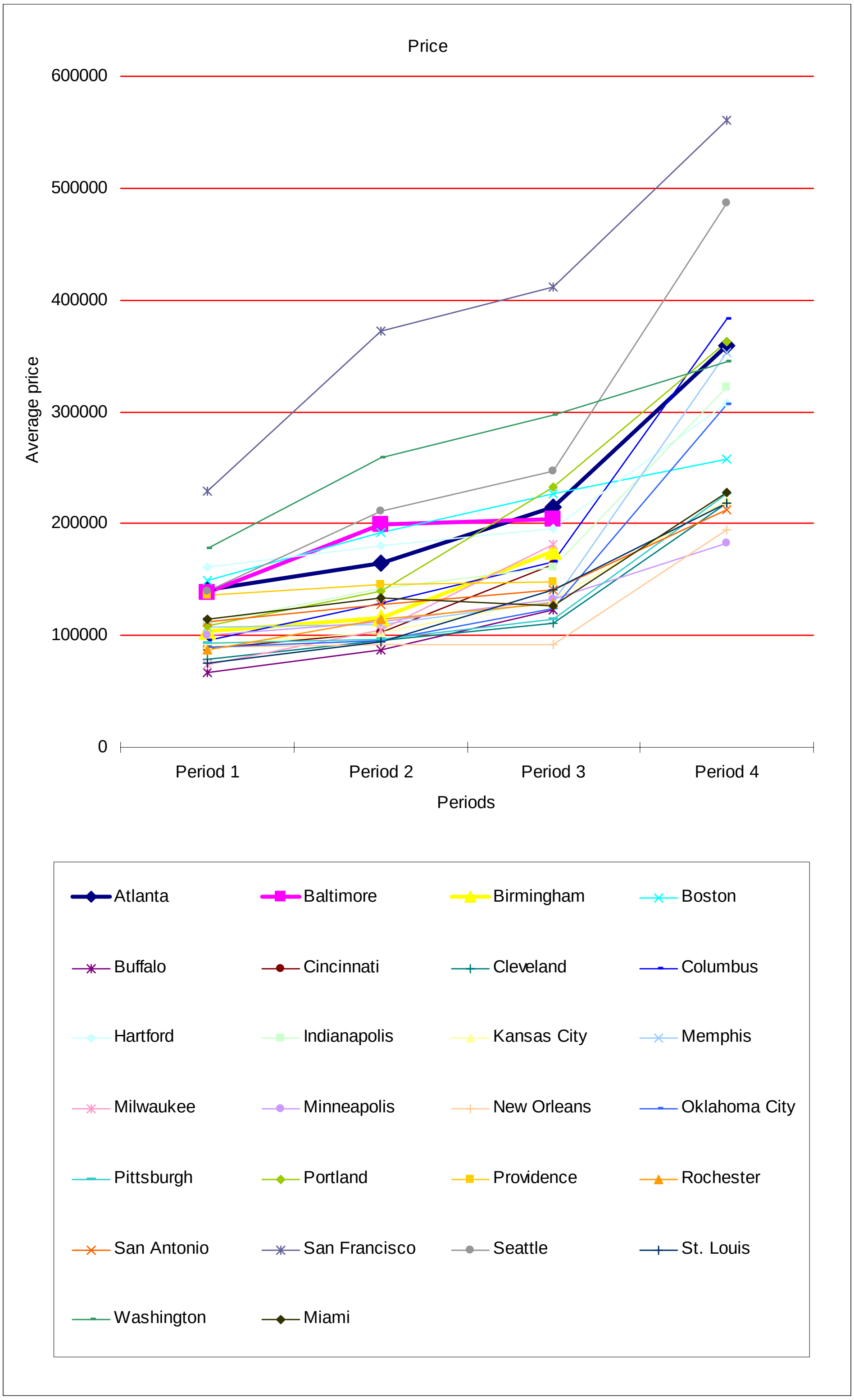

Exhibit 3: Descriptions Of Variables In Clustering Analysis

Variable	Categories and Descriptions
Property value	Current value of unit
Unemployment rate[3]	Rate of unemployment
Tax payment	Yearly real estate taxes payment
Mortgage rate	Current interest rate on primary mortgage (in %)
Household income	Expected household income in next twelve months
Unit size	Size of the unit (in square feet)
Rooms	Number of rooms in the unit (including bedrooms, bathrooms, living rooms, kitchens, family rooms, office, and other rooms)
Crowding	Number of persons per room
Unit quality rating	Rating of unit as a place to live (scale from 1(worst) to 10(best))
Neighborhood quality rating	Rating of neighborhood as a place to live (scale from 1(worst) to 10(best))

[3] The unemployment rate data is extracted from Bureau of Labor Statistics.

Exhibit 4: Cluster Generation History And Semipartial R-Square, R-Square Values Using Ward's Cluster Analysis

Number of Clusters	Cluster Joined		Frequency	SPRSQ	RSQ	PSF
25	ind	kan	2	0.0046	0.995	9.8
24	buf	cle	2	0.0055	0.99	9.2
23	min	mem	3	0.0068	0.983	8.9
22	atl	por	2	0.008	0.975	8.5
21	bos	mia	2	0.0085	0.967	7.6
20	bal	CL25	2	0.0086	0.958	7.3
19	CL22	CL23	2	0.0118	0.946	7.0
18	CL21	sea	4	0.0129	0.933	7.0
17	har	pro	4	0.0134	0.92	6.8
16	CL18	was	5	0.0136	0.906	6.7
15	CL20	pit	3	0.0165	0.89	6.7
14	CL24	mil	5	0.0229	0.867	6.6
13	bir	cin	2	0.023	0.844	6.5
12	CL19	okl	6	0.0238	0.82	6.4
11	CL18	CL16	3	0.0284	0.792	6.2
10	CL14	nor	2	0.0293	0.763	6.1
9	CL15	stl	6	0.0323	0.73	6.1
8	CL10	CL13	2	0.042	0.688	6.0
7	CL12	roc	6	0.0461	0.642	5.9
6	CL11	sfr	4	0.0514	0.581	5.8
5	CL7	san	8	0.0574	0.503	5.6
4	CL5	CL17	14	0.0708	0.423	5.5
3	CL9	CL8	18	0.1021	0.331	8.2
2	CL4	CL3	24	0.1715	0.209	6.4
1	CL2	CL6	27	0.2092	0	.

Exhibit 5: Plot Of Number Of Clusters Versus Semipartial R-Square Using Ward's Cluster Analysis

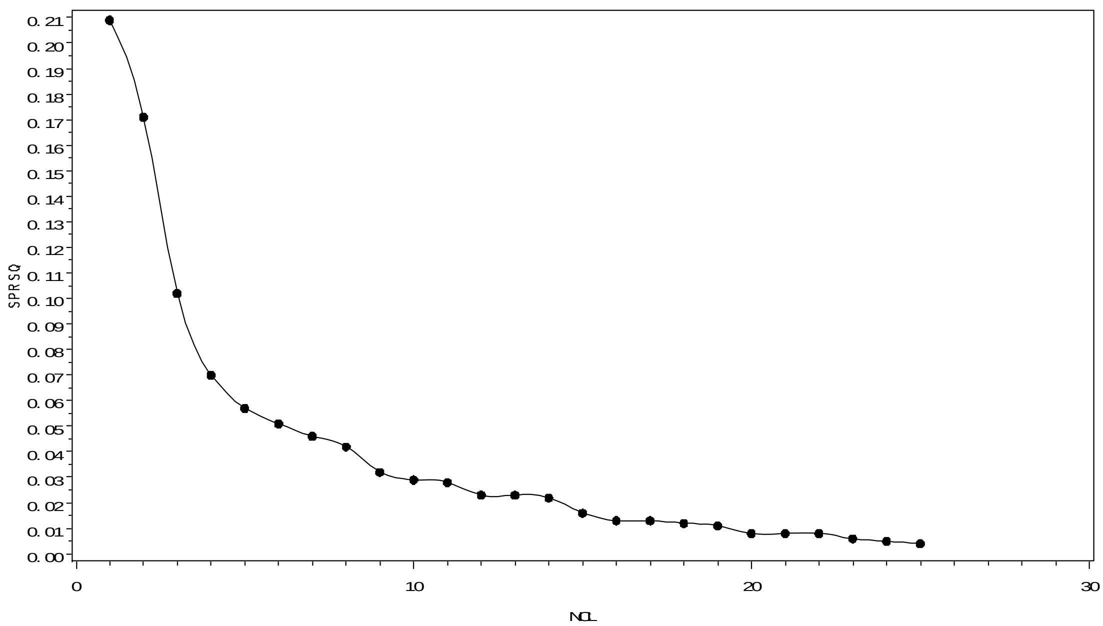

Exhibit 6: Tree Diagram Of Clusters Versus Semipartial R-Square Using Ward's Cluster Analysis

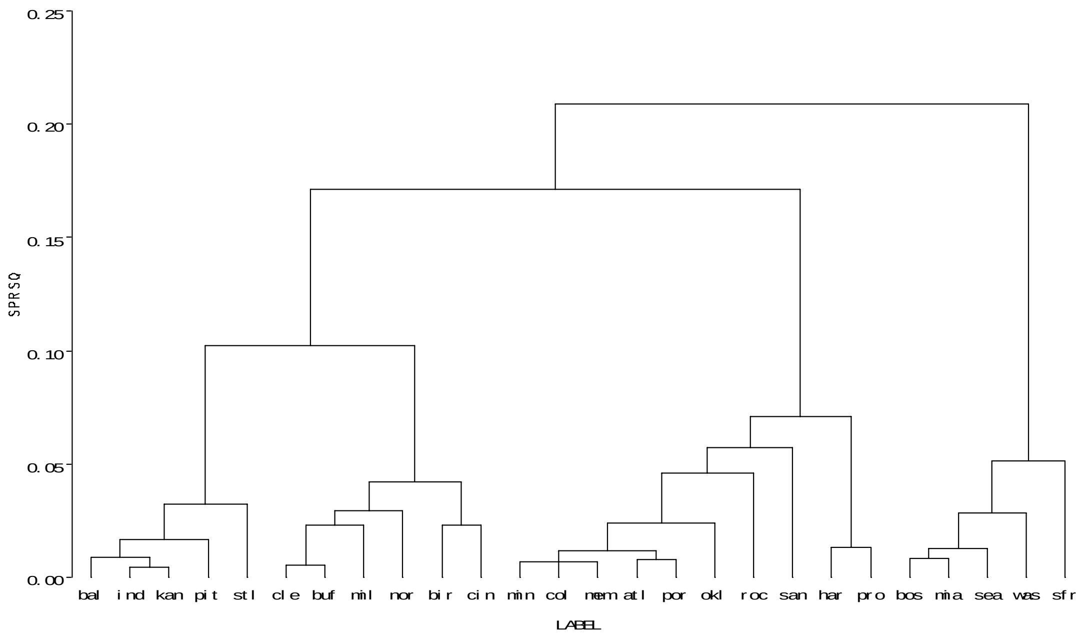

Exhibit 7: K-Means Clusters With Three Groups Specified

Cluster	Cities
1	Baltimore, Birmingham, Buffalo, Cincinnati, Cleveland, Indianapolis, Kansas City, Milwaukee, New Orleans, Pittsburgh, St. Louis
2	Atlanta, Hartford, Minneapolis, Columbus, Memphis, Oklahoma City, Portland, Providence, Rochester, San Antonio
3	Boston, Miami, San Francisco, Seattle, Washington D.C.

Exhibit 8: Characteristics Of Owner-Occupied Dwellings Belongings To The Cities In The Clusters

Cluster	Average Property Price ($)	Unemployment rate (%)	Tax payment ($)	Mortgage rate (%)	Income ($)
1	168,831	0.0538	1,508.59	9.49%	44,476
2	194,347	0.0506	1,737.70	9.27%	51,742
3	210,574	0.0442	1,805.80	9.21%	52,595
Cluster	**Size (Sq ft)**	**Rooms**	**Crowding (Persons/room)**	**Unit Rating**	**Neighborhood Rating**
1	2,229	7.78	0.288	8.183	7.871
2	2,167	7.63	0.289	8.102	7.659
3	2,041	7.38	0.292	8.036	7.423

Exhibit 9: The Regional Allocations Of 26 Metropolitan Areas Based On Kasarda (1995)

Midwest	Cincinnati, Cleveland, Columbus, Indianapolis, Kansas City, Milwaukee, Minneapolis, St Louis
Northeast	Boston, Buffalo, Hartford, Pittsburgh, Providence, Rochester
South	Atlanta, Baltimore, Birmingham, Memphis, Miami, New Orleans, Oklahoma City, San Antonio Washington
West	Portland, Seattle, San Francisco

Exhibit 10: The Generalized Squared Distances Between Groups Defined By Clustering Analysis Of This Paper

Generalized Squared Distance to Cluster			
Cluster	Cluster 1	Cluster 2	Cluster 3
Cluster 1	0	41.1523	30.7925
Cluster 2	41.1523	0	10.9202
Cluster 3	30.7925	10.9202	0

Exhibit 11: The Generalized Squared Distances Between Groups Defined By Region-Based Grouping Method

Generalized Squared Distance to Region				
Region	Midwest	Northeast	South	West
Midwest	0	1.43783	13.05104	27.44764
Northeast	1.43783	0	5.29394	11.55841
South	13.05104	5.29394	0	6.50109
West	27.44764	11.55841	6.50109	0

Exhibit 12: Number Of Observations And Percent Classified Into Region For The Clustering Analysis Of This Paper

Cluster	Cluster 1	Cluster 2	Cluster 3	Total
Cluster 1	10	1	0	11
	90.91	9.09	0	100
Cluster 2	0	10	0	10
	0	100	0	100
Cluster 3	0	0	5	5
	0	0	100	100
Total	10	11	5	26
	38.46	42.31	19.23	100
Priors	0.3333	0.3333	0.3333	

Exhibit 13: Number Of Observations And Percent Classified Into Region For Region-Based Grouping Method

Region	Midwest	Northeast	South	West	Total
Midwest	5	2	1	0	8
	62.50	25.00	12.50	0	100
Northeast	0	6	0	0	6
	0	100	0	0	100
South	0	1	8	0	9
	0	11.11	88.89	0	100
West	0	0	0	3	3
	0	0	0	100	100
Total	5	9	9	3	26
	19.23	34.62	34.62	11.53	100
Priors	0.25	0.25	0.25	0.25	

Exhibit 14: Metropolitan AHS Geography (Year 1985-2004)

Metropolitan Area	Post-1985 Surveys
Atlanta, GA	1987, 1991, 1996, 2004
Baltimore, MD	1987, 1991, 1998
Birmingham, AL	1988, 1992, 1998
Boston, MA	1985, 1989, 1993, 1998
Buffalo, NY	1988, 1994, 2002
Cincinnati, OH	1986, 1990, 1998
Cleveland, OH	1988, 1992, 1996, 2004
Columbus, OH	1987, 1991, 1995, 2002
Hartford, CT	1987, 1991, 1996, 2004
Indianapolis, IN	1988, 1992, 1996, 2004
Kansas City, MO	1986, 1990, 1995, 2002
Memphis, TN	1988, 1992, 1996, 2004
Miami, FL	1986, 1990, 1995, 2002
Milwaukee, WI	1988, 1994, 2002
Minneapolis, MN	1985, 1989, 1993, 1998
New Orleans, LA	1986, 1990, 1995, 2004
Oklahoma City, OK	1988, 1992, 1996, 2004
Pittsburgh, PA	1986, 1990, 1995, 2004
Portland, OR	1986, 1990, 1995, 2002
Providence, RI	1988, 1992, 1998
Rochester, NY	1986, 1990, 1998
San Antonio, TX	1986, 1990, 1995, 2004
San Francisco, CA	1985, 1989, 1993, 1998
Seattle, WA	1987, 1991, 1996, 2004
St. Louis, MO	1987, 1991, 1996, 2004
Washington, DC	1985, 1989, 1993, 1998

Exhibit 15: Descriptions Of Variables In Hedonic Pricing Model

Variable	Categories and Descriptions
I. Dependent variable	
PValue	Log of the reported selling price if the property sold within last twelve months, otherwise the owner's estimate of the current value of the property[4]
II. Independent variables	
Unemployment[5]	Rate of unemployment
Income	Log of household income
Crowding	Number of persons per room
$Year_i$, i=1,..,4	The dummy variable equals 1 if the dwelling was surveyed in the first year of the four (or three) time periods and 0 otherwise. One category is omitted to avoid perfect multi-collinearity. The omitted survey year is the last year of the four (or three) time periods.
Structural variables	
Size	Size of the unit (in square feet)
Rooms	The total number rooms in the unit (including bedrooms, bathrooms, living rooms, kitchens, family rooms, office, and other rooms0
Detached	Structural type dummy variable, equals 1 if the dwelling was in one unit detached building and 0 otherwise
Age	Dummy variable of aged building; equals 1 if the dwelling was built before 1960; and 0 otherwise
Garage	Dummy variable, equals 1 if there is garage in house or building and 0 otherwise
Basement	Dummy variable, equals 1 if there is basement in house or building and 0 otherwise
Heating1	Dummy variable; equals 1 if any of the following heating equipments is present in the building: forced warm-air furnace with ducts and vents to individual rooms, steam or hot water system with radiators, other system using steam or hot water, electric heat pump present in the unit; and 0 otherwise
Heating2	Dummy variable; equals 1 if any of the following heating equipments is present in the building: built-in electric

[4] In the survey years before 1998, the property values were recorded in intervals. We recoded these intervals to their midpoints as suggested by Malpezzi, Ozanne and Thibodeau (1980).
[5] The unemployment rate data is extracted from Bureau of Labor Statistics.

	baseboard heating or electric coils in floors, ceilings, or walls, floor, wall, or other pipeless furnace built into the building; and 0 otherwise
AC	Dummy variable, equals 1 if the central air conditioning present and 0 present
BuildingProblem	Dummy variable of building problem (1 if the unit has any of the following problems: basement leaks, roof leaks, open cracks or holes in walls or ceilings, holes in floor, or broken plaster or peeling paint over an area exceeding one square foot; 0 otherwise)
HallwayProblem	Dummy variable of public hallway problem (1 if the unit has any of the following problems: absence of light fixtures in public halls, hazardous steps on common stairs or stair railings not firmly attached; 0 otherwise)
LackFeature	Dummy variable of lack of important features (1 if the unit has any of the following deficiencies: lacks plumbing; lacks complete kitchen facilities; sewer system is a chemical toilet, privy, outhouse, facilities in another structure or some other sewage/toilet facilities; wiring in house not concealed or some rooms lack working electrical outlets; 0 otherwise)
Breakdown	Dummy variable of equipment breakdowns in last 90 days (1 if the unit has any of the following equipment breakdowns: water breakdown, flush toilet breakdown, sewage system breakdown, or fuses blown or circuit breakers tripped unit; 0 otherwise)
Neighborhood variables	
Quality	Rating of neighborhood as a place to live (scale from 1(poor) to 10(excellent))
Abandon	Dummy variable of abandoned buildings, equals 1 if abandoned buildings are observed within 300 feet of the dwelling and 0 otherwise
Junk	Dummy variable of junk, equals 1 if respondents observed accumulation of trash, litter or junk in neighborhood and 1 otherwise
Crime	Dummy variable of crime, equals 1 if street/neighborhood crime present and 0 otherwise
Noise	Dummy variable of noise, equals 1 if noise in neighborhood is bothersome and 0 otherwise

Exhibit 16: Hedonic Estimate Of Housing Attributes Effects On Housing Prices:

Linear Regression Model

	Cluster 1		**Cluster 2**		**Cluster 3**	
N	3543		3046		1662	
F statistic	127.61		91.6		123.49	
R-Square	0.4685		0.5312		0.5277	
Covariate	Coefficient	Std Err	Coefficient	Std Err	Coefficient	Std Err
Intercept	9.22424	0.15023	9.68102	0.15401	10.15587*	0.20462
Income	0.12843*	0.01446	0.11364*	0.01505	0.113*	0.0188
Crowding	-0.08055	0.06624	-0.17376*	0.07166	-0.11931	0.09783
Year2	0.16227*	0.02573	0.05376*	0.02755	0.26711*	0.03923
Year3	0.25965*	0.02907	0.15128*	0.02971	0.31332*	0.03972
Year4	0.39389*	0.04103	0.48985*	0.04	0.46097*	0.05192
Size	0.00982	0.00628	0.03742*	0.00687	0.00844	0.00945
Rooms	0.06983*	0.00635	0.07729*	0.0061	0.03627*	0.0088
Detached	0.23349*	0.03449	0.00745	0.03032	0.10339*	0.0495
Age	-0.0256	0.02079	-0.054*	0.02275	-0.02178	0.03187
Garage	0.12233*	0.02212	0.07986*	0.02221	0.22629*	0.0291
Basement	0.13508*	0.03468	0.09739*	0.0301	0.01006	0.05799
Heating 1	0.09111*	0.03814	0.03303	0.03616	0.17479*	0.05341
Heating 2	0.00528	0.04202	0.16588*	0.03701	0.08973*	0.05336
AC	0.21527*	0.02293	0.05143*	0.02643	0.18526*	0.0416
BuildingProblem	-0.05387*	0.02305	-0.08742*	0.02498	-0.05014	0.0351
HallwayProblem	-0.16673*	0.02802	-0.10849*	0.02739	-0.0598*	0.03618
LackFeature	-0.12416*	0.04551	-0.12313*	0.04794	-0.01784	0.07066
Breakdown	-0.00473	0.02386	-0.04568*	0.02383	-0.0398	0.03396
Quality	0.03667*	0.00474	0.01949*	0.0046	0.03034*	0.00764
Rat	-0.18519*	0.0699	-0.01155	0.07271	-0.11311	0.08782
Abandon	-0.29144*	0.04626	-0.17525*	0.05795	-0.04948	0.07091
Junk	-0.14274*	0.02577	-0.04498*	0.02671	-0.07439*	0.03521
Crime	-0.04981*	0.0239	-0.02769	0.02483	-0.02968	0.03421
Noise	-0.05367	0.03952	-0.05846	0.03955	-0.06496	0.04632

* Statistically significant at 10% confidence level

Exhibit 17: Hedonic Estimate Of Housing Attributes Effects On Housing Prices:

Mixed Effect Model

	Cluster 1		Cluster 2		Cluster 3	
N	3543		3046		1662	
Chi-square statistics	29.05		195.63		26.18	
Covariate	Coefficient	Std Err	Coefficient	Std Err	Coefficient	Std Err
Random effect						
Unemployment	-3.9651*	0.6932	-8.8119*	0.6111	-7.195*	1.3161
Fixed effect						
Intercept	8.9575	0.1566	9.2094	0.1525	10.5814	0.2172
Income	0.1304*	0.01439	0.1092*	0.01455	0.1036*	0.01871
Crowding	-0.08346	0.06593	-0.1323*	0.06934	-0.1004	0.09701
Year2	0.1383*	0.02595	0.02471	0.02719	0.3649*	0.04048
Year3	0.2558*	0.02894	0.1922*	0.02887	0.4039*	0.04623
Year4	0.3641*	0.04117	0.4904*	0.03867	0.457*	0.05145
Size	0.01251*	0.006267	0.03249*	0.006649	0.01091	0.009378
Rooms	0.06862*	0.006321	0.0763*	0.005895	0.03482*	0.008726
Detached	0.2693*	0.03489	0.04682	0.02944	0.09703*	0.04906
Age	-0.03066	0.02071	-0.04414*	0.02201	-0.04499	0.03187
Garage	0.1293*	0.02205	0.1002*	0.02152	0.2031*	0.02915
Basement	0.1201*	0.03461	0.1321*	0.0292	0.01905	0.05749
Heating 1	0.05375	0.03852	0.04152	0.03535	0.1351*	0.05342
Heating 2	0.03652	0.04218	0.1241*	0.0359	0.08991*	0.05287
AC	0.2044*	0.0229	0.03272	0.02559	0.1478*	0.04179
BuildingProblem	-0.04977*	0.02295	-0.07066*	0.02418	-0.04923	0.03478
HallwayProblem	-0.1749*	0.02793	-0.07191*	0.02661	-0.06766*	0.03588
LackFeature	-0.1335*	0.04532	-0.1463*	0.04638	-0.01155	0.07003
Breakdown	0.008523	0.02375	-0.01733	0.02313	-0.02089	0.03383
Quality	0.0369*	0.004716	0.01711*	0.004453	0.03115*	0.007573
Rat	-0.2208*	0.06985	-0.03005	0.07036	-0.1009	0.08705
Abandon	-0.2818*	0.04608	-0.174*	0.05603	-0.05258	0.07027
Junk	-0.1574*	0.02578	-0.04281*	0.02583	-0.07918*	0.0349
Crime	-0.04595*	0.0238	-0.02648	0.02401	-0.02985	0.0339
Noise	-0.04844	0.03935	-0.04132	0.03826	-0.0652	0.0459

* Statistically significant at 10% confidence level

Exhibit 18: Hedonic Estimate Of Housing Attributes Effects On Housing Prices With Pooled Data: Linear Regression Model

Covariate	Coefficient	Std Err
Intercept	8.98149	0.10526
Income	0.17811*	0.01011
Crowding	0.01587	0.04878
Year2	0.11901*	0.01904
Year3	0.25721*	0.02025
Year4	0.60255*	0.02705
Size	0.01604*	0.00464
Rooms	0.07423*	0.00438
Detached	0.06354*	0.02269
Age	-0.06664*	0.01534
Garage	0.07485*	0.01532
Basement	0.06367*	0.02327
Heating 1	-0.09041*	0.02623
Heating 2	0.15293*	0.02749
AC	0.01785	0.01743
BuildingProblem	-0.08925*	0.01708
HallwayProblem	-0.12733*	0.01887
LackFeature	-0.11837*	0.03357
Breakdown	0.00993	0.01687
Quality	0.03558*	0.00341
Rat	-0.05323	0.04904
Abandon	-0.25263*	0.03587
Junk	-0.08474*	0.01838
Crime	-0.04838*	0.01723
Noise	0.05243*	0.02679

* Statistically significant at 10% confidence level

Exhibit 19: F-Test To The Equality Of The Coefficient Estimates Of The Hedonic Regressions

Independent variables	F value	Pr > F
Income	1.19	0.3029
Crowding*	3.43	0.0324
Year2*	10.38	0.0001
Year3*	5.38	0.0046
Year4	0.84	0.4324
Size*	9.12	0.0001
Rooms*	13.33	0.0001
Detached	0.31	0.733
Age*	10.55	0.0001
Garage*	4.96	0.007
Basement*	4.19	0.0152
Heating 1	2.03	0.1311
Heating 2*	29.09	0.0001
AC	0.04	0.9563
BuildingProblem	0.68	0.5044
HallwayProblem*	4.63	0.0098
LackFeature*	3.53	0.0295
Breakdown	1.09	0.3373
Quality	0.97	0.3803
Rat*	10.11	0.0001
Abandon	1.2	0.3007
Junk*	3.31	0.0365
Crime	1.61	0.2006
Noise	1.19	0.3029

* Statistically significant at 10% confidence level

Exhibit 20: Descriptions Of Variables In Tenure Choice Analysis

Variable	Categories and Descriptions
I. Dependent variable	
Tenure	Dummy variable of the ownership of the dwelling; Equals 1 if the household owns the home; Equals 0 if they rent the home
II. Independent variables	
(South)	The dwelling is in the South area of the United States
Midwest	The dwelling is in the Midwest area of the United States
Northeast	The dwelling is in the Northeast area of the United States
West	The dwelling is in the West area of the United States
Unemployment	Rate of unemployment
Income	Expected household income in next twelve months
Household demographics	
Race	Dummy variable of the race of householder; equals 1 if black only; and 0 otherwise
Crowding	Number of persons per room
Age	The primary family's age
Education	The primary family's education
Marital	The primary family's marital status; equals 1 if married; and 0 otherwise
Gender	The primary family's gender; equals 1 if male; and 0 female

Exhibit 21: Sample Statistics Of Tenure Choice Analysis

Variables	Mean	Count
Homeownership	68.96%	46,132
South	34.68%	23,198
Midwest	35.71%	23,889
Northeast	17.00%	11,372
West	12.61%	8,436
Unemployment rate	5.84%	
Income	$6,756.36	
Race	13.67%	
Crowding	0.36	
Age	47.58	
Education	40.51	
Marital	46.89% married	
Gender	53.61% male	

Exhibit 22: Logit Estimation For Probability Of Homeownership

Variable	**Coefficient Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > Chi-Square**
Midwest	-0.0388	0.0251	2.3818	0.1228
Northeast	-0.1151**	0.0311	13.6512	0.0002
West	-0.2245**	0.0396	32.1405	<.0001
Unemployment	-9.1681**	1.5717	34.0277	<.0001
Income	0.3747**	0.0299	156.6415	<.0001
Race	-0.5613**	0.0278	408.6093	<.0001
Crowding	-2.2087**	0.0623	1257.287	<.0001
Age	0.0395**	0.000696	3214.313	<.0001
Education	0.1139**	0.00365	972.2254	<.0001
Marital	0.9713**	0.0128	5790.191	<.0001
Gender	0.0144	0.0107	1.7926	0.1806
Pseudo-R^2			0.4346	
Likelihood ratio (LR)			745.25	<.0001
Score			427.49	
Wald			668.39	<.0001
* Significant at 5%				
** Significant at 10%				

Exhibit 23: Proportional-Hazard Regression Model For Move From Rent To Own

Variable	Coefficient Estimate	Standard Error	Chi-Square	Pr > Chi-Square	Hazard Ratio
Midwest	-0.07881**	0.03122	6.3736	0.0116	0.924
Northeast	-0.25003**	0.04163	36.0642	<.0001	0.779
West	-0.09468**	0.04758	3.9597	0.0466	0.910
Unemployment	-7.67452**	1.89096	16.4718	<.0001	0.000
Income	0.34455**	0.0361	91.0826	<.0001	1.411
Race	-0.36418**	0.04413	68.0995	<.0001	0.695
Crowding	-1.09664**	0.08919	151.1737	<.0001	0.334
Age	-0.02777**	0.000998	774.1298	<.0001	0.973
Education	0.07426**	0.00488	231.469	<.0001	1.077
Marital	0.75022**	0.03054	603.5341	<.0001	2.117
Gender	0.06652**	0.02661	6.2475	0.0124	1.069
Likelihood ratio (LR)				2396.7978	<0.0001
Score				2227.3314	<0.0001
Wald				2188.1997	<0.0001
* Significant at 5%					
** Significant at 10%					

References

- Abraham, J. M., Goetzmann, W. N., and Wachter, S. M. (1994) "Homogeneous Groupings of Metropolitan Housing Markets", *Journal of Housing Economics*, **3**(3), 186–206
- Abraham, Jesse. M and Patric H. Hendershott (1996) "Bubbles in Metropolitan Housing Markets", *Journal of Housing Research*, 7, pp191-206
- Adair, A. S., J. N. Berry, W. S. McGreal (1996), "Hedonic Modelling, Housing Submarkets and Residential Valuation", *Journal of Property Research*, 13(1), pp67–83
- Amemiya, Takeshi (1980) "Selection of Regressors", *International Economic Review*, 21, pp331-354
- Anas, Alex and S. J. Eum (1984) "Hedonic Analysis of a Housing Market in Disequilibrium", *Journal of Urban Economics*, 15
- Anglin, Paul M. and Gencay Ramazan (1996) "Semiparametric Estimation of a Hedonic Price Function", *Journal of Applied Econometric*, 11(6), pp633-648
- Atkinson, Scott E. and Thomas D. Crocker (1987) "A Bayesian Approach to Assessing the Robustness of Hedonic Property Value Studies", *Journal of Applied Econometrics*, 2, pp27-45
- Awan, K., J. Odling-Smee and M. Whitehead (1982) "Household Attributes and the Demand for Private Rental Housing", Economica, 49, pp183-200
- Barth, J., Kraft, J. and Wiese, P. (1975), "A Portfolio-Theoretic Approach to Industrial Diversification and Regional Employment", *Journal of Regional Science*, 15(1), pp9-15
- Basu, Sabyasachi and Thomas G. Thibodeau (1998) "Analysis of Spatial Autocorrelation in

House Price", *Journal of Real Estate Finance and Economics*, 17(1), pp61-85

- Blomquist, Glenn and Lawrence Worley (1982) "Specifying the Demand for Housing Charactersitcs: The Exogeneity Issue", in D.B. Diamond and G. Tolley (eds.), *The Economics of Urban Amentities*, Academic Press, 1982
- Bowden, Roger J. *The Econometrics of Disequilibrium*, North Holland, 1978
- Boyle, Melissa A. and Katherien A. Kiel (2001) "A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities", *Journal of Real Estate Literature*, 9(2), pp117-144
- Buchel, S and M. Hoesli (1995) "A Hedonic Analysis of Rent and Rental Revenue in the Subsidized and Unsubsidized Housing Sectors in Geneva", *Urban Studies*, 32(7), pp1199-1213
- Butler, Richard V. (1982) "The Specification of Hedonic Indexes for Urban Housing", *Land Economics*, 58(1), pp96-108
- Chambers, Daniel (1992) "The Racial Housing Price Differential and Racially Transitional Neighborhoods", *Journal of Urban Economics*, 32(2), pp214-232
- Chesire, Paul and Stephen Sheppard (1995) "On the Price of Land and the Value of Amenity", Econometrica, 62, pp247-267
- Chiuri, M.C. and Jappelli, T. (2003) "Financial Market Imperfections and Home Ownership: A Comparative Study", *European Economic Review*, 47(5), pp857-875
- Clark, W.A.C. and Deurloo, M.C. (1994) "Tenure Changes in the Context of Micro-level Family and Macro-level Economic Shifts", *Urban Studies*, 31(1), pp1-21
- Conroy, M. E. (1974), "Alternative Strategies for Regional Industrial Diversification",

Journal of Regional Science, 14, pp31-46

- Davidoff, T., (2006) “Labor Income, Housing Prices and Homeownership”, *Journal of Urban Economics*, 59(2), pp209-235
- De Borger, Bruno (1986) “Estimating the Benefits of Public Housing Programs: A Characteristics Approach”, *Journal of Regional Science*, 26, pp761-773
- Diamond, Douglas B. Jr. and Barton Smith (1985) “Simultaneity in the Market for Housing Characteritics”, *Journal of Urban Economics*, 17, pp280-292
- Din Allan, Martin Hoesli and Andre Bender (2000) “Environmental Values and Real Estate Prices”, *Urban Studies*, 38(100), pp1989-2000
- Elton, E. J. and M. J. Gruber (1970) “Homogeneous Groups and the Testing of Economic Hypotheses”, *Journal of Financial and Quantitative Analysis*, 4(5), pp581-602
- Fair, Ray C. and Dwight M. Jaffee (1972) “Methods of Estimation for Markets in Disequilibrium”, *Econometrica*, 40, pp497-514
- Fisher, Franklin and Karl Shell, “Taste and Quality Change in the Pure Theory of the True Cost of Living Index” in Zvi Griliches (ed.), *Price Indexes and Quality Change*, Harvard University Press, 1971
- Follain, James R. and Emmanuel Jimenz (1985) “Estimating the Demand for Housing Characteristics: A Survey and Critique”, *Regional Science and Urban Economics*, 15(1), pp77-107
- Follain, James R. and Larry Ozanne (1979) “Place to Place Indexes of the Price of Housing”, The Urban Institute: Washington, DC
- Follain, James R. and Stephen Malpezzi (1980) “Estimates of Housing Inflation for

Thirty-Nice SMSAs: An Alternative to the Consumer Price Index”, *Annals of Regional Science*, pp41-56

- Follain, James R. and Stephen Malpezzi (1981) “Are Occupants Accurate Appraisers?” *Review of Public Data Use*, 9(1), pp47-55
- Follain, James R. and Stephen Malpezzi (1981) “The Flight to the Suburbs: Insight from an Analysis of Central City versus Suburban Housing Costs”, *Journal of Urban Economics*, 9(3), pp381-398
- Freeman, A. Myrick (1979) “Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues”, *Scandanavian Journal of Economics*, 81(2), pp154-173
- Gatzlaff, Dean H. and Donald R. Haurin (1997) “Sample Selection Bias and Repeat-Sales Index Estimates”, *Journal of Real Estate Finance and Economics*, 14, pp35-50
- Goetzmann, W. N. (1993), “The Single Family Home in the Investment Portfolio”, *Journal of Real Estate Finance and Economics*, 6(3), pp201-222
- Goetzmann, W. N., and Wachter, S. M. (1995a) “Clustering Methods for Real Estate Portfolios”, *Real Estate Economics*, **23**(3), 271–310
- Goetzmann, W. N., and Wachter, S. M. (1995b) “The Global Real Estate Crash: Evidence from an International Database”, in *A Global Perspective on Real Estate Cycles* by Stephen J. Brown and Crocker H. Liu, Boston: Kluwer Academic Publishers (2001)
- Goodman, John L. Jr. and John B. Ittner (1992) “The Accuracy of Home Owners’ Estimates of House Value”, *Journal of Housing Economics*, 2(4), pp339-357
- Guiso, L. and T. Jappelli (2002) “Private Transfers, Borrowing Constraints and the Timing

of Homeownership”, *Journal of Money, Credit and Banking*, 34(2), pp315-339

- Gyourko, Joseph and Richard Voith (1992) “Local Market and National Components in House Price Appreciation”, *Journal of Urban economics*, 32, pp52-69
- Halverson, Robert and Henry O. Pollakowski (1981) “Choice of Functional Form for Hedonic Price Equations”, *Journal of Urban Economics*, 10(1), pp37-49
- Harvey S., Rosen, Kenneth T. and Holtz-Eakin, D., (1984) "Housing Tenure, Uncertainty, and Taxation", *The Review of Economics and Statistics*, MIT Press, vol. 66(3), pp405-416
- Haurin, D., (1991) “Income variability, homeownership and housing demand”, *Journal of Housing Economics*, 1, pp60–74
- Henderson, J.V. and Ioannides, Y. M., (1983), “A model of housing tenure choice”, *American Economic Review*, 73(1), pp98–113
- Henderson, J.V. and Ioannides, Y.M. (1987) “Owner Occupancy: Investment Vs Consumption Demand”, *Journal of Urban Economics*, 21, pp 228-241
- Hilber, Christian A.L., (2005) "Neighborhood externality risk and the homeownership status of properties", *Journal of Urban Economics,* Elsevier, vol. 57(2), pp213-241
- Hoesli, M., Lizieri, C., and Macgregor, B. D. (1997) “The Spatial Dimensions of the Investment Performance of UK Commercial Property”, *Urban Studies*, 34(9), 1475–1494
- Jud, G Donoal and Daniel T. Winkler (2002) “The Dynamics of Metropolitan Housing Prices”, *The Journal of Real Estate Research*, 23, 1/2, pp29-45
- Kain, John F. and John M. Quigley (1972) “Housing Market Discrimination, Home Ownership and Savings Behavior”, *American Economic Review*, 62, pp263-277
- Kasarda, John D. (1995) “Industrial Restructuring and the Changing Location of Jobs”, in:

R. Farley, (Ed.) *State of the Union: America in the 1990s, Volume 1: Economic Trends*, pp. 215-267. New York: Russell Sage Foundation

- Kennickell, Arthur B., Martha Starr-McCluer and Brian J. Surette “Recent Changes in U.S. Family Finances: Results from the 1998 Survey of Consumer Finances”, *Federal Reserve Bulletin*, Jan 2000, 1-29
- Knight, J.R., R. Carter Hill and C.F. Sirmans (1992) “Biased Prediction of Housing Values”, *Real Estate Economics*, 20(3), pp427-456
- Kort, J. R. (1981), “Regional Economic Instability and Industrial Diversification in the U.S.”, *Land Economics*, 57(4), pp596-608
- Lancaster, K. (1966) “A New Approach to Consumer Theory”, *Journal of Political Economy*, 74, 132-157
- Linneman, Peter and Wachter, Susan M., (1989) “The impacts of borrowing constraints on homeownership”, *Real Estate Economics*, 17(4), 389–402.
- Malpezzi, Stephen (2002) “Hedonic Pricing Models: A Selective and Applied Review”, In: O’Sullivan, T., Gibb, K. (Eds.), *Housing Economics and Public Policy*. Blackwell, Malder, MA.
- Malpezzi, Stephen, Gregory Chun and Richard Green (1998) “New Place-to-Place Housing Price Indexes for U.S. Metropolitan Areas, and Their Determinants: An Application of Housing Indicators”, *Real Estate Economics*, 26(2), pp235-275
- Malpezzi, Stephen, Larry Ozanne and Thomas Thibodeau (1980) “Characteristic Prices of Housing in Fifty-Nine Metropolitan Areas”, Research Report, Washington, DC: The Urban Institute

- Mason, Carl and John M. Quigley (1996) “Non-Parametric Housing Prices”, *Housing Studies*, 11(3), pp373-385
- Meese, Richard and Nancy Wallace (1991) “Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Price Indices”, *Real Estate Economics*, 19(3), pp308-332
- Morrizumi, Yoko and Naoi, Michio (2006) “Income Uncertainty, Private Transfer and the Timing of Homeownership: Evidence from Retrospective Panel Data in Japan”, *Keio University Market Quality Research Project*, pp1-17
- Olsen, Edgar O. (1972) “An Econometric Analysis of Rent Control”, *Journal of Political Economy*, 80, pp1081-1100
- Olsen, Edgar O. and David M. Barton (1983) “The Benefits and Costs of Public Housing in New York City”, *Journal of Public Economics*, 20, pp299-332
- Ortalo-Magne, Francois and Rady, Sven (1999) “Boom in, bust out: young households and the housing price cycle”, *European Economic Review*, 43(4-6), pp755–766
- Ortalo-Magne, Francois and Rady, Sven (2002) “Tenure Choice and the Riskiness of Non-housing Consumption”, *Journal of Housing Economics*, 11, pp266-279
- Ozanne, Larry and Stephen Malpezzi (1985) “The Efficacy of Hedonic Estimation with the Annual Housing Survey: Evidences from the Demand Experiment”, *Journal of Economic and Social Measurement*, 13(2), pp153-172
- Pace, R. Kelley (1993) “Nonparametric Methods with Applications to Hedonic Model”, *Journal of Real Estate Finance and Economics*, 7(3), pp185-204
- Pace, R. Kelley and Otis W. Gilley (1997) “Using the Spatial Configuration of the Data to

Improve Estimation", *Journal of Real Estate Finance and Economics*, 14(3), pp333-340

- Ortalo-Magne, Francois and Rady, Sven (1999) "Boom in, bust out: young households and the housing price cycle", *European Economic Review*, 43(4-6), pp755–766
- Ortalo-Magne, Francois and Rady, Sven (2002) "Tenure Choice and the Riskiness of Non-housing Consumption", *Journal of Housing Economics*, 11, pp266-279
- Plaut, Steven E. (1987) "The Timing of Housing Tenure Transition", *Journal of Urban Economics*, 21, pp312-322
- Robst, J., Deitz R. and McGoldrick, K. (1999) "Income variability, uncertainty and housing tenure choice", *Regional Science and Urban Economics*, 29(2), pp219-229
- Rosen, Sherwin (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economy*, 82(1), pp34-55
- Rosen, Harvey S. and David J. Fullerton (1977) "A Note on Local Tax Rates, Public Benefits Levels, and Property Values", *Journal of Political Economy*, 85(2), pp443-440
- Rosen, H.S., Rosen, K.T., and Holtz-Eakin, D. (1984) "Housing tenure, uncertainty and taxation", *The Review of Economics and Statistics*, 66, pp405–416
- Rudel, T.K. (1987) "Housing price inflation, family growth and the move from rented to owner-occupied housing", *Urban Studies*, 24, pp258-267
- Schnare, Ann B., Raymond J. Struyk (1976), "Segmentation in Urban Housing Markets", *Journal of Urban Economics*, 3, pp146–166
- Sheppard, Stephen (1999) "Hedonic Analysis of Housing Markets", in Paul C. Chesire and Edwin S. Mills (eds.), *Handbook of Regional and Urban Economics*, volume 3, Elsevier
- Sinai, Todd and Souleles, Nicholas S. (2005) "Owner-Occupied Housing as a Hedge

Against Rent Risk", *The Quarterly Journal of Economics*, Vol. 120, No. 2, pp763-789

- Thibodeau, Thomas G. (1989) "Housing Price Indexes from the 1974-1983 SIMSA Annual Housing Surveys", *Real Estate Economics*, 17(1), pp100-117
- Thibodeau, Thomas G. (1995) "Housing Price Indexes from the 1984-1992 MSA Annual Housing Surveys", *Journal of Housing Research*, 6(3), pp439-481
- Tsatsaronis, Kostas and Haibin Zhu (2004) "What Drives Housing Price Dynamics: Cross-Country Evidence", *BIS Quarterly Review*, March 2004, pp65-78

VITA

Yu Lu, was born in China in 1979. After earned B.A. in Accounting and Finance degree, she started her graduate study at the University of Missouri – Columbia in Fall 2003. She earned Master in Statistics degree in Winter 2008 and expects to graduate with PhD in Economic degree in Fall 2009.

She currently works in a database marketing company in Connecticut. She has broad interests in economics, statistics, finance and marketing.