SURVIVAL ANALYSIS OF FORECLOSURES BY HOUSING AND NEIGHBORHOOD CHARACTERISTICS

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Resha Shehari Aviena Dias

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Agribusiness and Applied Economics

November 2014

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Survival Analysis of Foreclosures by Housing and Neighborhood Characteristics

By

Resha Shehari Aviena Dias

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Lei Zhang
Chair

Dr. Ruilin Tian

Dr. Siew Lim

Approved:

11/21/14
Date

Dr. William Nganje
Department Chair

ABSTRACT

Losing a home affects not only the homeowner but the community at large. This thesis examines both housing and neighborhood characteristics to determine which characteristics have an impact on the speed of the sale of a foreclosed property. The findings from the Cox Proportional hazard model show which characteristics buyers value more when buying a foreclosed property. These characteristics can be different from the ones that are used to determine the sales price of a foreclosed property. Both financial institutions and real estate speculators can benefit from knowing which housing characteristics are preferred in a foreclosed property, allowing them to sell that property fast in the market.

ACKNOWLEDGEMENTS

First and foremost, I thank Almighty God for giving me the strength to complete another milestone in my life. I'd like to thank Dr. Lei Zhang for her guidance and patience given to me when completing this step. A special thank you goes out to my committee members, Dr. Siew Hoon Lim and Dr. Ruilin Tian for giving their valuable suggestions in making this a success. I would also like to thank my parents and my uncle Mr. Anton Pulle for believing in me and being constant pillars of strength. Finally, I would like to extend my sincere thanks to Peter for supporting me in every step of the way.

TABLE OF CONTENTS

ABSTRACT iii

ACKNOWLEDGEMENTS iv

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF APPENDIX TABLES ix

LIST OF APPENDIX FIGURES x

CHAPTER 1. INTRODUCTION 1

1.1. Problem Statement 4

1.2. Implications 5

CHAPTER 2. FORECLOSURES 7

2.1. Foreclosure Definition 7

2.2. Foreclosure Process 7

2.3. Texas Overview 12

2.4. Ownership Transfer Process 15

CHAPTER 3. LITERATURE REVIEW 17

3.1. Foreclosure Spillover Effects on Neighboring Property Values 17

3.2. Short Sale Spillover Effects 22

3.3. Non-Price Related Foreclosure Spillover Effects on the Neighborhood 23

CHAPTER 4. DATA AND METHODOLOGY 25

4.1. Data 25

4.2. Methodology 29

CHAPTER 5. RESULTS AND DISCUSSION 43

5.1. OLS Regression for Time to Resolution 43

5.2. Cox Proportional Hazard Model 45

5.3. Hedonic Pricing Model 58

CHAPTER 6. CONCLUSIONS 62

6.1. Limitations 63

6.2. Need for Future Study 64

REFERENCES 65

APPENDIX 68

LIST OF TABLES

Table Page

2.1. List of States that Follow Judicial Foreclosure Process 9

2.2. List of States that Follow Non-Judicial Foreclosure Process 10

2.3. List of States that Follow both Foreclosure Processes 11

4.1. The Summary Statistic Table for Housing Characteristics 25

4.2. The Summary Statistics for Neighborhood Characteristics 27

5.1. Coefficients of Time to Resolve OLS Regression 44

5.2. Cox Proportional Hazard Model Estimates 49

5.3. Survival Time Estimates 52

5.4. Results for Hedonic Regression (OLS) 58

LIST OF FIGURES

Figure Page

1.1. Chart of U.S Annual Foreclosure Activity from 2005- 2012 2

2.1. Map of Dallas County Location 12

2.2. School District Map of Dallas County 14

2.3. Foreclosure Ownership Transfer Process 16

4.1. Mean Sale Price of Foreclosures by Year 28

5.1. Kaplan- Meier Survival Estimate 51

5.2. Nelson-Aalen Cumulative Hazard Graph for Pool, Central Air Conditioning, Condition 3 and School District 5 53

5.3. The Smoothed Hazard Graph from Cox Proportional Hazards Regression 54

5.4. The Cumulative Cox Proportional Hazards Regression Graph 55

5.5. Cox Proportional Hazards Cumulative Graphs for Variables Pool, Central Air Conditioning, Condition 3 and School District 5 56

5.6. Cox-Snell Residual 57

LIST OF APPENDIX TABLES

Table Page

A.1. Summary Statistics 68

A.2. Time to Resolution Estimates 69

A.3. Cox Proportional Hazard Estimates and Ratios 72

A.4. Kaplan-Meier Survival Statistics 75

A.5. Hedonic Model Estimates 79

A.6. Correlation Matrix 82

LIST OF APPENDIX FIGURES

Figure | Page

A.1. Hazard Graphs.. 77

CHAPTER 1. INTRODUCTION

The "American Dream" has become one of home ownership and recently this dream has become more attainable for many (Rohe and Watson, 2007). However, to eventually own a home, most people have to take out a mortgage, a real estate loan with property as the underlying asset. Traditionally, mortgages were 15 to 30 year loans which would be paid off gradually; "mortgage" coming from the French term "death pledge", meaning that you had your life to pay it off. However, as home ownership started increasing in the United States of America in the beginning of 1995, the mortgages to support these ownerships became greatly customized (US Census Bureau, 2007).

This increase of homeownership was in part made possible by subprime mortgages that target customers with lower credit ratings using interest rates that are above prime lending rates. The financial crisis in 2008, heightened defaults in the sub-prime mortgages category, and as a result the housing bubble burst. The number of foreclosure filings spiked by a stunning 81% between 2007 and 2008. The number of foreclosures was further increased nationally by about 22% from nearly 2.3 million in 2008 to 2.8 million in 2010 (RealtyTrac, 2014). Figure 1.1 provides the annual foreclosure filings from 2005 to 2012. While the economy has gradually recovered, the number of foreclosure filings remained above the level prior to 2008.

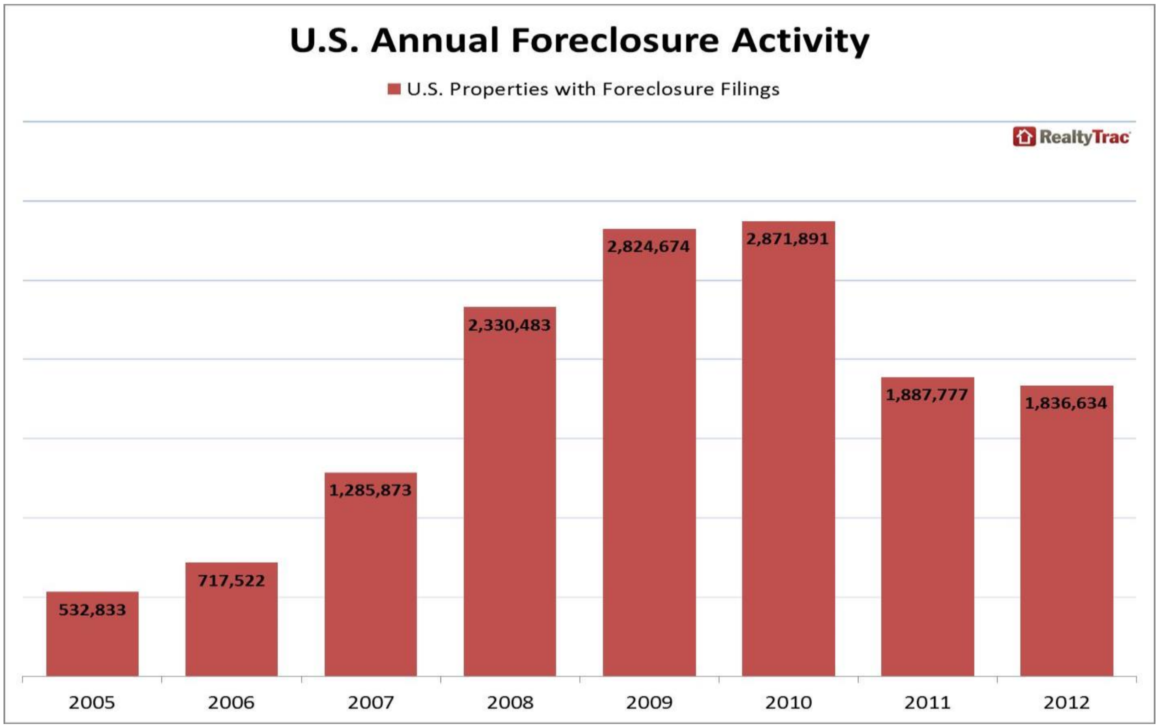

Figure1.1. Chart of U.S Annual Foreclosure Activity from 2005- 2012

Source: RealtyTrac.

One of the reasons why the financial downturn in 2008 had such a widespread effect was that these subprime and prime mortgages are important elements in the financial and the housing markets that are highly interdependent. To better understand the mutual downturn one must first understand how mortgages became traded investment assets. These debt obligations such as mortgage-backed securities are one important part of the housing bubble that burst in 2008. Mortgage-backed securities, or MBS, represent the cash flow payments that are from mortgage loans, mostly commercial and residential. Private companies, government agencies, or combinations of the two would then purchase the loans from banks and the loan originator. The next step is known as securitization, which is defined by the U.S. Securities and Exchange Commission as when “the entity then issues securities that represent claims on the principal and interest payments made by borrowers on the loans in the pool” (SEC, 2014).

Many of these securities were then issued by government organizations such as the Government National Mortgage Association (Ginnie Mae), the Federal National Mortgage Association (Fannie Mae, or the Federal Home Loan Mortgage Corporation (Freddie Mac). Others that were issued by private firms were called “private-label” mortgage securities. Due to the government packaging of the securities, many investors felt that these securities were all backed by the US government and therefore were not risky assets. However, only the Ginnie Mae securities were fully backed, but both Freddie Mac and Fannie Mae enjoyed the ability to borrow money from the US Treasury. Further spreading the use of mortgage-backed securities was the fact that they can exist in many different forms. What grew to be very popular with investors is a type of MBS called a collateralized mortgage obligation. These obligations, also known as mortgage derivatives, allowed different risk rated MBS to be mixed so that the investor would be exposed to a desired amount of risk (SEC, 2014).

Unfortunately, a problem arose when there was an economic downturn in 2008. The downturn caused people to lose their jobs or not have the ability to pay their mortgage. This then meant that people's properties were being foreclosed. However, while a bank plans for some foreclosures, they were often times left with more properties than they had expected. To make matters worse, the sudden surge in foreclosed and for sale houses made it so that many houses were now worth less than the bank had issued a mortgage for. This meant that not only was the bank stuck with a physical asset that was worth less than they paid for it, but also home owners who were still able to make payments were sometimes stuck paying a mortgage that was higher than the actual value of their house.

When payments are not made on the mortgage and the property is foreclosed, the costs incur to all parties who are involved in the transaction, including the individual in default, the lender, and households living in the neighborhood. Some studies find that average costs related to foreclosures are about $27,000 for the city and $10,000 for the neighborhood of the foreclosure (Moreno, 1995). Additionally, these costs will spread to city governments, which Apgar and Duda (2005) find can cost up to $30,000 per property in most cases. These neighborhood costs are the externalities associated with property values and quality of the neighborhood. Studies conducted on different aspects of foreclosures are beneficial and lead to better enacted policies that address this issue and reduce spillover effects to the society.

1.1. Problem Statement

Most scholars so far have discussed the foreclosure spillover effects having an impact on the neighborhood properties. Once a property is foreclosed, it needs to be transferred to a final user. This way, the property can be productive and the new owners can engage in maintaining the property which otherwise could result in vandalism (Campbell, 2011). If houses are not occupied for a long period of time it can further worsen blight and reduce the desirability of the property. For this test purposes, if a house is sold after being foreclosed, it is considered as a resolved property. An unresolved property is one that has not been sold after facing foreclosure. Not many scholars have investigated what takes place after a property is being foreclosed.

In that sense the gap exists in studying resolved and unresolved foreclosures. This study aims to fill that gap by focusing on factors that are likely to help resolved and unresolved properties. To be included the foreclosed property would be one that has been unresolved during a certain period by using a similar data set used by Zhang et al. (2014) for Dallas County, Texas. The data set consists of foreclosures and sales transactions from 2004-2012. The methodology will follow a similar survival analysis framework used by Towe and Lawley (2013). By using a Cox proportional hazard model, the likelihood of a property being resolved will be calculated by taking housing and neighborhood characteristics into consideration. This study will shed light on factors that help resolve foreclosed properties faster. A closer understanding about resolved and unresolved cases would help identify where the focus should be in housing markets and allow necessary actions to be taken accordingly.

1.2. Implications

Houses can be considered as illiquid assets if they are not occupied (Campbell et al, 2011). The liquidity of a house depends on many factors and is not limited to the price of the property alone (Kluger & Miller, 1990). When a borrower defaults, the ownership of the property is transferred to the bank/ financial institution that financed the mortgage loan. Financial institutions do not have an incentive to hold on to a vacant house and would attempt to sell it to another intermediate party or to a final user. Holding on to an unoccupied house does not bring the financial institution any revenue and this gives the motive to sell it as quickly as possible. Another drawback is the high maintenance costs associated with foreclosed properties. As discussed in previous sections, foreclosed houses are also prone to crime related activities and it brings a social disorder to the neighborhood (Immergluck & Smith, 2006). For all these reasons, financial institutions would like to transfer the ownership of a foreclosed property to another party.

Investigating the types of housing characteristics that help to sell a property once it is foreclosed is useful to financial institutions that deal with mortgages. This will help them lend money to parties who are interested in properties that are desirable and easier to sell in the event of a foreclosure. For example, if a property with a detached carport has a tendency of sitting in the market for a long time, then investing money on such properties can be avoided. Real estate speculators can also invest in foreclosed properties that have desirable characteristics as there is a higher chance of selling that property back in the market. In return, this identifies which housing and neighborhood characteristics are preferred by most home buyers in the Dallas County.

The rest of the paper will be as follows: Chapter 2 defines foreclosures and foreclosure practices, Chapter 3 looks into the previous work done by other scholars, Chapter 4 establishes the methodology, Chapter 5 discusses the results and finally Chapter 6 explains the limitations and the summary of the study.

CHAPTER 2. FORECLOSURES

2.1. Foreclosure Definition

A foreclosure can be defined as the process of taking possession of a mortgaged property after payments are not met, most often to be sold to recoup the loss. This process takes place once the home owner misses one or more monthly payments. Generally the foreclosure process would begin 3 -6 months after missing the first mortgage payment (U. S Department of Housing and Urban Development, 2014). Once the payments have been missed, the foreclosure process begins. The laws involved in the foreclosure process and the timeline of the actions taken usually depends on the individual state. According to the U.S department of Housing and Urban Development, it should be noted that the lender usually identifies borrower's short term financial crisis. Therefore, it is important to stay in touch with the lender in the event of a missed payment. Failing to make necessary arrangements would declare the borrower to be in default and the foreclosure process begins.

2.2. Foreclosure Process

In general the two main types of foreclosures are judicial and non-judicial. Both methods require public notices to be issued and all parties involved to be notified regarding the proceedings. One main difference between the two methods is that judicial foreclosures are mainly involved with the court. Under a judicial foreclosure, when a borrower misses one or more payments and show signs of defaulting on the mortgage, the lender can file a complaint with the court asking for permission to foreclose the property and take possession of it. The lender will record a notice in the public land records about the property that is being foreclosed, and will notify the potential buyers, creditors and other interested parties. The borrower can file a separate suit and save the property from being foreclosed. However, if the court decides that the borrower has defaulted, permission to carry out a foreclosure will be granted to the borrower. This allows the borrower to recover the left over amount of the mortgage.

Once permission is granted for a foreclosure, the court will authorize a sheriff's sale. This is an auction of the property which is held in a public place. It is open to anyone for bidding and the highest bidder legally becomes the owner of the property. Most of the time, the lender wins the highest bid and becomes the owner of the property. This will be transferred to a final party who will occupy the house and end the foreclosure sale. Section 2.4 of the paper discusses the process of a foreclosure sale and the ownership transfer in detail.

As previously mentioned, each state has different foreclosure practices. Once a borrower defaults on his/her mortgage, a redemption period helps the borrower to reclaim the house that has been foreclosed. The redemption period happens after a foreclosure takes place, where the borrower can pay the balance of the mortgage and all costs incurred during the foreclosure process to reclaim the house from the buyer (foreclosed sale). The availability of a redemption period mostly depends on whether the foreclosure process is judicial or non-judicial. For example, Texas is one of the states that do not allow the borrowers to have a redemption period. This is especially important because this study focuses on the foreclosure activities in the Dallas County of Texas. Given below in Table 1.1 is a list of states that practice the judicial foreclosure process. Both process and redemption periods are given in number of days.

Table 2.1. List of States that Follow Judicial Foreclosure Process

Judicial Foreclosures		
State	**Process Period**	**Redemption Period**
Connecticut	62	Court Decides
Delaware	170-210	None
Florida	135	None
Illinois	300	90
Indiana	261	None
Kansas	130	365
Kentucky	147	365
Louisiana	180	None
Maine	240	90
Maryland	46	Court Decides
Massachusetts	75	None
Nebraska	142	None
New jersey	270	10
New Mexico	180	30-270
New York	445	None
North Dakota	150	180-365
Ohio	217	None
Pennsylvania	270	None
South Carolina	150	None
Vermont	95	180-365

Source: RealtyTrac.com

Unlike judicial foreclosures, the second type of foreclosure process, the non-judicial foreclosure, does not have any interventions from the court. Most cases are handled by attorneys following a state mandated process. The requirements for this process are established by state statute which varies across states. Just as discussed in the judicial foreclosures sections, once the borrower defaults on the mortgage, a default letter will be mailed out to him/her. It is then recorded in the county’s recorder's office and is posted in public places. At the end of the notice period (this can differ according to the state), the foreclosure sale will take place and the highest bidder becomes the owner of the property (Mortgage Bankers Association, 2014). Given below in Table 1.2 is a list of states that only follow a non-judicial foreclosure process.

Table 2.2. List of States that Follow Non-Judicial Foreclosure Process

Non-judicial Foreclosures			
State	**Comment**	**Process Period**	**Redemption Period**
Michigan	Non-judicial Only	60	30-365
New Hampshire	Non-judicial Only	59	None
Tennessee	Non-judicial Only	40-45	730
Utah	Non-judicial Only	142	Court Decides
Washington D.C.	Trustee Sale Only	47	None
West Virginia	Trustee Sale Only	60-90	None

Source: RealtyTrac.com

Some states are not limited to just one foreclosure process. These states follow both judicial and non-judicial foreclosure processes. However even in these states, one process is followed more often than the other. In the table given below, the comment section specifies if both methods are used or one is preferred over the other.

Table 2.3. List of States that Follow both Foreclosure Processes

States with both Foreclosure Types			
State	**Comment**	**Process Period**	**Redemption Period**
Alabama	Judicial Rarely	49-74	365
Alaska	Judicial Rarely	105	365*
Arizona	Judicial Rarely	90+	30-180*
Arkansas	Both	70	365*
California	Judicial Rarely	117	365*
Colorado	Judicial Rarely	145	None
Georgia	Judicial Rarely	37	None
Hawaii	Both	220	None
Idaho	Trustee Sale	150	365
Iowa	Trustee Sale Voluntary	160	20
Minnesota	Non-judicial Mostly	90-100	180
Mississippi	Non-judicial Mostly	90	None
Missouri	Non-judicial Mostly	60	365
Montana	Trustee Sale Mostly	150	None
Nevada	Trustee Sale Mostly	116	None
North Carolina	Non-judicial Mostly	110	None
Oklahoma	Judicial Mostly	186	None
Oregon	Trustee Sale Mostly	150	180
Rhode Island	Non-judicial Mostly	62	None
South Dakota	Judicial Mostly	150	30-365
Texas	Non-judicial Mostly	27	None
Virginia	Trustee Sale Mostly	45	None
Wisconsin	Judicial Mostly	290	365
Wyoming	Non-judicial Mostly	60	90-365

*these redemption periods are only for judicial foreclosure process.

Source: RealtyTrac.com

2.3. Texas Overview

Texas was the 28th state to join the United States of America, and the only one to be an independent covering nation before joining. Following California, Texas is the second most populated state and following Alaska, it is the second largest state. Beyond these facts and figures, Texas is famous for “open spaces and cheap property” (Gopal, 2013). This has become increasingly important as Texas is facing a steady increase in its population. According to the Real Estate Center at Texas A& M and Atlas Van Lines (Texas A&M, 2014), Texas has had a surplus of inbound movers for the 9 years prior to 2014. This has led to 22.8% population growth from 1990 to 2000 and a 20.6% population growth from 2000 to 2010 (US Census, 2000; 2010). On the state level, it would appear that the Texas housing market has been doing quite well during the recession and is facing a quick recovery; however that may not be true for each of the 254 counties.

2.3.1. Dallas County

Figure 2.1. Map of Dallas County Location

Dallas County (highlighted in Figure 1.2.) is located in the North Eastern part of Texas, with the county seat and largest city being Dallas. The county has a population of nearly 2.4 million and it is the ninth most populated county in America (US Census, 2010). The county has over 800,000 households with many of those being located in the Dallas-Fort Worth-Arlington metropolitan area. The demographics of the county are approximately 38% Hispanic, 33% White, 22% Black, 5% Asian and 14% from other races. Dallas County is facing the same recovery as much of the country and the star of Texas. Along with this recovery Dallas County has benefited from an increase in house prices, with the mean price of a house sold in August 2014 being almost 15% higher than a year prior. Even with this, Dallas County continues to rank as one of the top affordable regions in the National Association of Realtors' Affordability Index (2014).

Dallas County has seventeen different school districts. In the latter part of the paper, the study will determine whether school districts have a role to play in determining house prices in foreclosed properties. According to an accountability report done in 2011 by Texas Education Agency, the school districts are ranked according to ranks given to each individual school in the district and averaging it to arrive at an overall rank. The four categories for the ranks are exemplary, recognized, academically accepted and academically unacceptable. Two school districts with exemplary ranks are Highland Park and Sunnyvale. The five school districts that have recognized ranks are Carrolton Farmers Branch, Cedar Hill, Mesquite, Coppell and Richardson. School districts with academically acceptable ranks are Dallas, Garland, Irving, Lancaster, Grand Prairie, Desoto and Duncanville. Figure 2.3, shows fifteen school districts of the Dallas County.

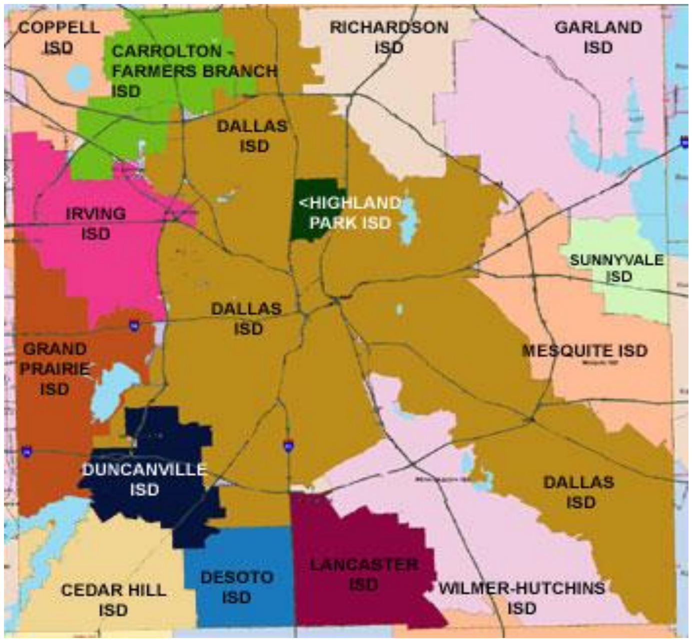

Figure 2.2. School District Map of Dallas County

Source: www.dallascountytexas.us

2.3.2. Texas Foreclosure Process

The foreclosure process in Texas has a very short timeline once the borrower is in default. Both judicial and non-judicial practice of foreclosure is practiced in the state. Typically the timeline expands to about 60 days, but with the exclusion of the right to redemption, homeowners can lose their house in less than 41 days. Judicial foreclosure process takes place if no power of sale clauses or deed of trust exists in a mortgage. The process is similar to the general judicial foreclosure process mentioned in the previous section. With a power of sale clause specifying the time and terms of a sale, the foreclosure would be conducted as a non-judicial foreclosure. In the event of having a power of sale clause with no indication of terms of sale, the process will be carried out in the following manner. First, prior to the foreclosure process a lender must mail the letter of demand to the borrower asking for the balanced payment to be made within 20 days. In the second stage, where the 20 day limit has expired, the foreclosure process must begin by filing it with a country clerk. In this stage, the notice must be sent to the borrower and posted at the county courthouse door. In the final stage, the foreclosure must take place on the first Tuesday of any month. At the auction where the property is sold, anyone can bid and the highest bidder would then receive the ownership of the foreclosed property.

2.4. Ownership Transfer Process

When homeowners miss or start making late mortgage payments, it can often lead to foreclosure. Once a house is foreclosed, the ownership of that house can be transferred to three possible parties. These parties are the bank, the investor and the final owner. Once a foreclosure is inevitable, a property is either auctioned or settled through a short sale. When a property is auctioned, the ownership will transfer in four different ways. The first possibility would be if a bank receives the ownership, the property will eventually be transferred to the final owner or to an investor. Because of the bank's role in the auctioning process, the property referred to as Real Estate Owned (REO). Another possibility is in the case of the investor receiving the ownership from the bank, the property will then be transferred to a final owner. The third potential way is that at the auctions held for a foreclosed property, an investor can bid the highest price and receive the ownership of the property. As stated in the previous scenario, the property will then be transferred to a final owner. Finally, the most straightforward ownership transfer takes place when the final owner itself bids the highest price at the auctions and becomes the final owner of the property. A visualization of this information is available in Figure 2.3.

In the case of a missed mortgage payment, the alternative method to foreclosure is a short sale. The short sale requires approval from the lender as there will be a difference in the mortgage balance payment. The shortage of the payment arises from the lower price that is placed on the property. The price of the property is lower than the remaining mortgage loan payments. If the lender identifies and confirms that the borrowers do not have enough funds to pay off the mortgage payments, the short sale avoids the house being foreclosed. Another incidence where a short sale will take place is when a homeowner wants to move to a different property. However, before a short sale the lender will determine whether a short sale is beneficial over a foreclosure or not. If all the elements are met and the short sale is successful, the ownership of the property will be transferred to the final owner. This process too can be seen in Figure 2.3.

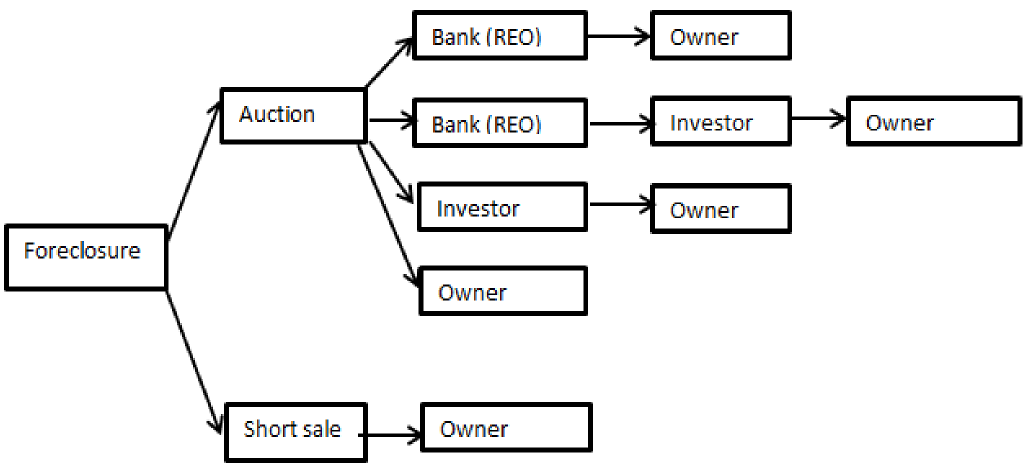

Figure 2.3. Foreclosure Ownership Transfer Process

The percentage of properties owned by a bank (REO) after the auctions is usually high. Out of all potential outcomes, most foreclosed properties after auctions are owned by the REO and it represents the majority of the foreclosure outcome. This paper mainly focuses on the first two scenarios (REO) represented by Figure 2.3, where the bank is directly involved.

CHAPTER 3. LITERATURE REVIEW

The financial crisis in 2008 heightened the foreclosure activities in most of the states. This increase in foreclosure activities gained attention by many scholars, leading them to further investigate foreclosures' impact on the society. In general, foreclosures have been identified as bringing in externalities to the neighborhoods. These externalities range from affecting the neighboring property values to general neighborhood spillovers such as vandalism and crime.

3.1. Foreclosure Spillover Effects on Neighboring Property Values

Spillovers occur when the presence of a foreclosed property starts to affect the houses in the surrounding area. Immergluck and Smith (2006) found external costs related to foreclosure activities on single – family property values. To research the effect they use a hedonic regression model to estimate the impact of foreclosures on the sales price of a property by using data of 1997-1998 foreclosure activities for Chicago. They find the increase of foreclosure activities within one eighth of a mile caused prices to decline by 0.9% of the house value. Furthermore, they describe the disorders foreclosures can bring in to a community through abandoned properties and criminal activities. The authors point out the need to study more about foreclosure related topics to enforce efficient policies to reduce the overall cost to the society.

The spillover studies conducted previously by Immergluck and Smith (2006) are extended by Lin, Rosenblatt and Yao (2009) who incorporate the spillover effects into a pricing based model based on the comparable properties theory. They use two major factors of discounts; the foreclosure sale and the weight placed on the foreclosure to evaluate the neighborhood property values. The Chicago Primary Metropolitan Statistical Area (PMSA) data were used for the study as it is a highly concentrated foreclosure area. The results of the study report that as the time and space between a foreclosed house and a chosen subject property increase, the spillover effects decrease. An 8.7% discount is the most severe number reported for the neighborhood property values within 300 feet and two years of liquidation. The values calculated for a sample of data for 2003, as opposed to the 2006 sample, report a price discount of only 5%. This is understandable as the housing market conditions in 2003 were better than the housing market conditions that prevailed in 2006.

Further exploring the spillover effects of foreclosures are Leonard and Murdoch (2009), who present how foreclosures can change the neighborhood quality. The authors improve upon previous studies on spillover effects by introducing a spatial auto-correlation model and analyzing properties in the foreclosure process. This is important as the neighborhood quality is neglected most often during the time a foreclosure is taking place. Results from this study show robust evidence for decreased prices in the selling prices of houses within 250 feet of foreclosed activities for both sold and unsold properties in Dallas County, Texas for 2006. These results are consistent with the hypothesis that residents of a neighborhood with foreclosed activities reduce spending on maintenance which contributes to a neighborhood quality decrease.

Many authors have explored the hypothesis that when the number of foreclosures in a neighborhood increases, housing prices in that area will decrease. Using a data set for foreclosure activities in St. Louis County, Missouri from 1998-2007, Rogers and Winter (2009) measure the impact of foreclosures on house prices. The time period used for the study allowed for more flexibility in terms of foreclosure effects by allowing them to be observed over time and across space. The hedonic model used contains spatial and temporal dimensions and control for unobserved neighborhood characteristics. However, any qualitative effects of foreclosure events are not accounted for in this model. The findings suggest that, as expected, the prices decline for neighboring houses, but these effects show a more localized trend. Another interesting result, along with the significant results for foreclosure impacts, is how the marginal impact for a foreclosure declined rather than increasing, in both cross section and over time. Unfortunately, the reasons for this decline were not explored by the authors.

In their paper, Kobie and Lee (2011), address the impact of residential foreclosures on single family residential property values by improving upon the work of Immergluck and Smith (2006). One particular issue Kobie and Lee address is the use of a straight line distance approach which ignores the structural details about each observation such as highways, parks etc. Therefore, the proximity measures are evaluated by introducing the face block unit. A face block unit includes only houses that are on the same street, to the left and the right and on the other side of the street from one intersection to another (Kobie and Lee, 2011). Using this improved method the authors explore two main hypotheses. First, a foreclosed property on the same face block as a house that is "for sale" will drive down the price of the latter house and second, that the duration of the foreclosure process will have an impact on the houses in the same face block unit. To test these hypotheses the authors use data from 2005-2007 for Cuyahoga County, Ohio, as the number of foreclosure activities has been increasing in that area

When running the model, Kobie and Lee divided the data into three categories of city, suburban and county level foreclosures. As an overall result, the first hypothesis held true for both the suburbs and the county categories and the negative impact from foreclosed properties in the same face block is proven to be significant due to delayed maintenance activities. The second hypothesis held true only for county level model, confirming that properties that had foreclosure processes stretched out for more than 12 months would have a negative impact on sales prices. From the three models (Cleveland, Suburban and Cuyahoga), the Cuyahoga model further pointed out the importance of classifying foreclosure properties as “pre-foreclosures” and “post foreclosures” as results can be different before and after an auction sale is held.

Campbell, Giglio and Pathak (2011) examine the differences between forced sales and unforced sales. Forced sales are more likely to occur after a house is foreclosed, close to the time of death of an owner or when facing bankruptcy. The authors find that in the state of Massachusetts forced sale houses have a lower price than other house sales. Results indicate that the largest discount rate of 27% applied to foreclosed house sales with the lowest at 3% for bankruptcy related sales.

In an attempt to investigate the existence of heterogeneity in the neighborhood price effects, Zhang and Leonard (2014) apply a quantile regression method and find that negative price externalities caused by foreclosures are highest in lower-priced homes. By using the quantile regression method, they account for differences across lower-priced houses and higher-priced houses for different time periods as well as distances. They examine the issue by using foreclosure and sales data from 2007-2009 for Dallas County. Having 4 different rings ranging from 0 to 1500ft, they incorporate the distance factor and find that within 250 ft the negative price externalities are the greatest. The study further points out why preventing lower-priced foreclosures should have more policies and how it can help reduce negative externalities to the neighborhood.

Many studies discuss the impact of foreclosures on neighboring property values by using a hedonic regression model. Harding et al (2009) focus on a possible contagion effect with regard to foreclosed properties and their externalities. They use a repeat sales approach which reduces the omitted variable problem often encountered with hedonic models. The method is an extension to the original method suggested by Bailey et al., (1963) with Schwartz et al., (2003) and Harding et al., (2007). The model specifically controls for property characteristics to prevent having any differences in the repeat sales regression. A negative contagion effect of 1% was observed from nearby distressed properties. As the distance between a foreclosed property and a non-distressed property increased, the contagion effect declined suggesting that beyond a distance of 500 feet, no significant contagion effects can be observed. This is determined by looking at repeat sales details for Atlanta, Charlotte, Columbus, Las Vegas, Los Angeles, Memphis, and St. Louis from 1989-2007. The timing of the foreclosure sale is important as the results find an increasing contagion discount when the foreclosure was inevitable for properties located within 300 feet for a foreclosed property. Although smaller values are obtained by this research, it supports the idea that immediate neighbors are affected by negative externalities proposed by prior literature.

Another study that examines the contagion effect of neighboring foreclosures is done by Towe and Lawley in 2013. An increase in the trend of strategic default means more people are walking away from their home mortgages even though the payments can be made. Towe and Lawley (2013) test out a possible contagion effect of neighboring foreclosures based on social interactions by using spatially explicit parcel level data in Maryland from 2006- 2009. Panel data is used in an attempt to avoid simultaneity problems that arise as a result of including social interactions. As seen in other papers, the contagion effect discussed in this paper was highly localized. The authors find that the increase of the foreclosure hazard ratio of 18% was due to a 1% increase of a neighbor in foreclosure. Neighbors of a property are defined as the 13 nearest houses and when the number of foreclosures increased from zero to four within the neighborhood, increase of foreclosure hazard ratio was 4.4% to 6.5%. Thus, the multiplier effects of the likelihood of neighbors defaulting on their loans after observing others in foreclosure can lead to a series of foreclosure activities within neighborhoods

3.2. Short Sale Spillover Effects

Currently, more studies have focused on non-distressed sales; however, Daneshvary, Clauretie and Kader (2011) approach the spillover effects from a distressed residential properties perspective. To achieve this, they use 2008 data for Las Vegas, Nevada because most of the sale transactions of this data were REO sales. The approach is different as they attempt to estimate the size of the discount for houses sold using a short sale method. The three types of distressed properties considered in this study are short sales, sales in the process for foreclosure, and sales after lender's repossession (REO). Some of the conditions previous studies did not control for variables such as status of the sold properties, physical condition of a property, and endogenous time on the market that affect transaction prices are all controlled in this study.

While a short sale can suffer from its own pricing discount factor, the study done by Daneshvary et al (2011) proves that it does not produce a spillover effect as other distressed properties do. Although short sale properties do not, both REO and properties that are in the middle of a foreclosure process tend to have similar spillover effects. The marginal spillover effect is reported as 1% for REO sales within 0.1 miles from a non-distressed home. When the distance increases from 0.1 to 0.5 miles, the spillovers decrease from 0.7% to 0.4%. Furthermore, the discount for distressed properties shows a significantly high own-price discount values ranging from 10%- 19%.

Also addressing toxic spillover effects is the paper by Danaeshvary and Clauretie (2010. In their paper, the authors use a hedonic pricing model to look at the effect of foreclosures on nearby houses. As previous studies have done, they develop a distance weight matrix in order to determine the role that distance of a foreclosed property plays on house sales price. The results of the study show that a REO property within 3 months and within .1 miles has a -2.9% effect on nearby houses, but decreases to -1.9% and 1.3% for .25 and .5 mile radius respectively. However, they note that when market trend and unobservable neighborhood characteristics are controlled, the role of an REO diminishes.

3.3. Non-Price Related Foreclosure Spillover Effects on the Neighborhood

Not all research related foreclosures have focused on spillover price effects. For example, Immergluck and Smith (2006) examine the relationship between crime and foreclosures. The study looks at both violent and property crimes in the neighborhood for Chicago using crime and neighborhood characteristics variables. Skogan (1990) explains how abandoned buildings can harm a neighborhood in different ways. In the same ways, foreclosed houses may also appeal to people for criminal and drug related activities. This is especially true for lower-income neighborhoods, when compared to the number of criminal activities that take place in high-income neighborhoods (Immergluck and Smith, 2006). Immergluck and Smith find significant results for an increase in violent crimes in the neighborhood as a result of foreclosures. When other things are held equal, this increase is about 2.33% for every 0.01% increase in foreclosure rates in a tract. However, the authors do not find significant results for an increase in property crimes as a result of foreclosure activities

The effects of foreclosures can be much more detrimental than just forced eviction. Lower property values, reduction in a local property tax base, increase in crime related activities, and a disruption in social ties are some of the many spillover effects (Lee 2008). As a result of foreclosure, houses that are poorly maintained are prone to attract vandalism and crime related activities. This is supported by Ellen, Lacoe and Sharygin (2012) with similar findings about the impact foreclosures have on neighboring block faces. Properties that are sold through REO method are more likely to face these spillovers. The results do not suggest any increase in new crimes, but once a foreclosure hits a neighborhood monitoring them closely can prevent some crime related activities.

CHAPTER 4. DATA AND METHODOLOGY

4.1. Data

The data used for this study contains foreclosure and sales data for Dallas County, Texas. The main source for the sales data is University of Texas at Dallas Real Estate Research Database. It includes a detailed list of variables such as the account number, latitude and longitude coordinates, housing characteristics, historic sale prices and sales data for the Dallas County. The summary statistics of some of the major housing characteristics are listed in Table 4.1.

Table 4.1. The Summary Statistic Table for Housing Characteristics

Variable	**Description**	**Mean**	**Std. Dev.**	**Min**	**Max**
Attached Garage	House with attached garage	0.782176	0.4127737	0	1
Attached Carport	House with attached carport	0.0468333	0.211285	0	1
Detached Carport	House with detached carport	0.0220654	0.1468985	0	1
Pool	House with swimming pool	0.0805036	0.2720756	0	1
Slab	House with slab foundation	0.7583619	0.4280829	0	1
Central Heat	House with central heat	0.9145682	0.2795276	0	1
Central AC	House with central ac	0.9000699	0.2999115	0	1
Baths	Number of baths	2.028075	0.6889137	0	10
Lot Area	Size of lot area in thousands of square feet	9.57534	9.944437	1.034877	468.1696
Living Area	Size of living area in thousands of square feet	1.796294	0.7967797	0.528	11.014
Fireplace	House with a fire place	0.6453962	0.4784007	0	1
Story 1	House with one story	0.7645619	0.4242791	0	1
Story1.5	House with one and a half stories	0.097927	0.2972206	0	1
Story 2	House with two stories	0.1364619	0.3432839	0	1
Effective Age	Years since last refurbished/ built	36.00232	20.10938	6	115

Further information comes from other sources, such as the list of foreclosures from 2005-2010 comes from the RealtyTrac database. RealtyTrac is a commercial database that contains many U.S real estate and other related data. In general, it has information and data on preforeclosure, auctions, bank owned (REO) and all recently sold properties. The list of foreclosure data collected from RealtyTrac was geocoded and spatially merged in to the real estate database. This step was necessary in identifying resolved foreclosures in the database. A resolved foreclosure in the data base was recognized as a market sale that took place after the foreclosure, but before July, 2012.The unresolved foreclosure in this case, was where no market sale information was found before July 2012 for a particular foreclosed property. Thus, the status of these properties after July, 2012 is not known.

Looking at the data, the highest lot area from the data set is reported as 468,169.6 sq ft while the lowest is 1034.877 sq ft. The total living area of a house has the lowest value of 528 sq ft and the highest being 11,014 sq ft. The most modern house was only 6 years old, while the oldest house was 115.

The next section of data, which includes neighborhood characteristics, is collected from American Community Survey 2006-2010 from the US Census data. The main categories of these variables are the race, age and poverty levels. The highest number of below poverty percentage is seen in the neighborhood is 83.96% and the highest number of school aged children (5-17) is 47.12%. A description of variables used is listed in the summary statistics table for neighborhood characteristics in Table 4.2.

Table 4.2. The Summary Statistics for Neighborhood Characteristics

Variable	Description	Mean	Std.Dev.	Min	Max
White	The percentage of white people in the neighborhood	0.3217089	0.2443916	0	1
School Age	The percentage of households with school age children(5-17)	0.2109185	0.0700638	0	0.4712203
Community Older than 60	The percentage of people who are 60 and above in the neighborhood	0.0809088	0.0639499	0	0.5962733
Below poverty	The percentage of people are below the poverty line in the neighborhood	0.1373073	0.1280769	0	0.8396861

The use of Cox proportional hazard model requires a main time variable; the time variable used in this study is the time to resolution (t2r). This variable is calculated by taking the time difference between a foreclosure and the sale of that particular property. For any property that was not sold, July 2012 was given as the date of the sale to calculate the variable t2r. Even though a sale date is given for unsold properties for calculation purposes, the property is still considered as an unresolved case. Hazard models take the occurrence of an event into consideration. For this study the event is defined as the sale of the property, which is technically considered as a failure event in survival analysis.

Another section of variables for this study is the inclusion of school districts. Out of the 17 school districts for the Dallas County, 14 of them are included in the study to control for any socio-geographic effects. The included sales districts are Carrollton-Farmers B ISD, Dallas ISD, Cedar Hill ISD, Garland ISD, Highland Park ISD, Irving ISD, Lancaster ISD, Mesquite ISD, Coppell ISD, Grand Prairie ISD, Richardson, Desoto ISD and Duncanville ISD are included. Some of the school districts were dropped as the observations were too small to obtain a significant result. Finally, the Sunnyvale ISD is the base case for these dummy variables so the results will measure the effect of the school district against Sunnyvale.

To further round out the characteristics and potential factors that might affect foreclosure sales seven house condition dummies are also included. These dummies are used to capture house quality effects on the time to resolution with the two ends of the housing quality being Excellent and Unsound. Furthermore, five foreclosure year dummy variables are used to account for yearly foreclosure effects with the base year being 2005. Finally, the season dummy variable of winter will account for any seasonal effects as well. A detailed list of all dummy variables can be found in the appendix section.

When looking at the sales prices for the foreclosed houses, the total number of houses sold for the time period from 2005-2012 is 22,699 with a mean price of $112,781.6 dollars. The highest price for a foreclosed property reported is $3,999,500 and the lowest is $1,200 dollars. Given below is a graph that breaks down the average sales price for a foreclosed house by the year it was sold.

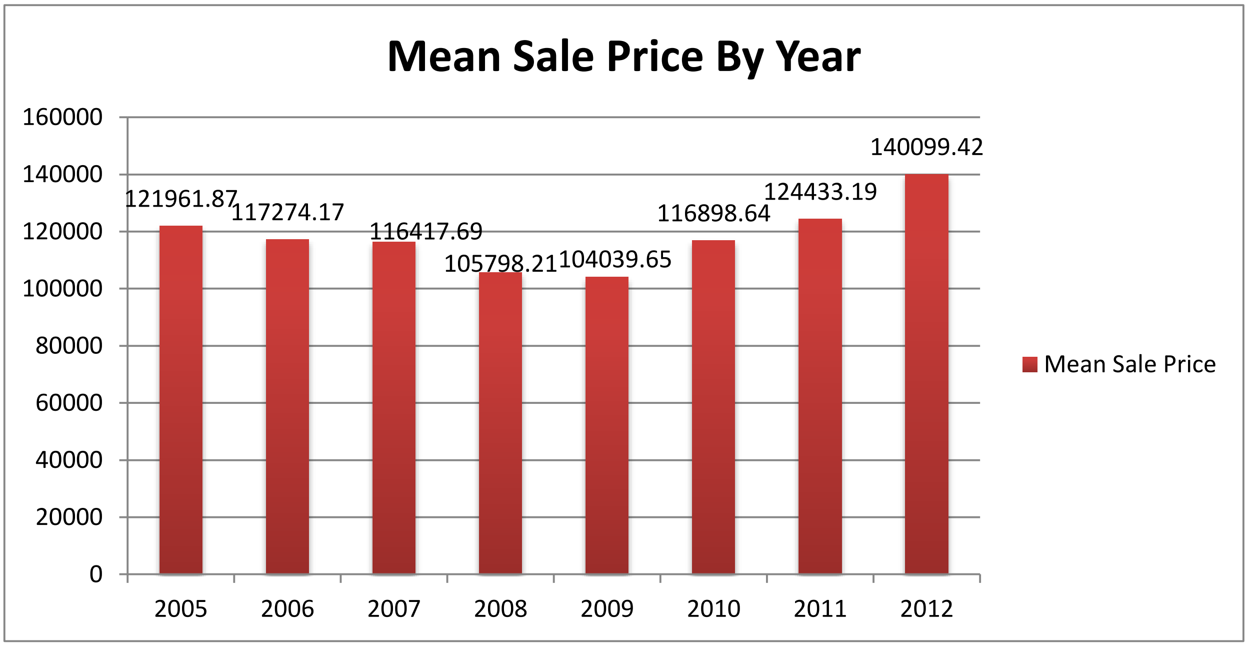

Figure 4.1. Mean Sale Price of Foreclosures by Year

As seen on the graph, there is a decline in the average sales price in 2008 and 2009, but it increases in 2010. By 2012 the average sales price for foreclosed sales increase up to $140,099. The total number of foreclosed houses sold has the highest number in 2008, which is 5996.

4.2. Methodology

In this section, the two main methods used in the study will be discussed. The first method will look at factors that affect time to resolution by using the Cox proportional hazard model. The second estimation used in the study which focuses on factors that impact sales price of foreclosed properties, will use the traditional hedonic regression method.

The Ordinary Least Squares (OLS) is a statistical tool that is used to find the best model to explain variations in the data. It is the line that best fits the observed data by minimizing the distance between observed data and predicted data. Using this method, the independent variables of housing characteristics, neighborhood characteristics and dummy variables for school districts, condition of the house, foreclosure year and winter can be regressed on the dependent variable of time to resolution.

The model can be written as

$$t2r = \alpha + \beta C + \delta N + \gamma D + \varepsilon \quad (4.1)$$

Where t2r is the dependent variable; the time to resolution, C is a vector of housing characteristics, N is a vector of neighborhood characteristics, D represents all the dummy variables including house condition dummy for all conditions, school district dummy for all districts ,foreclosure year dummy and a seasonal dummy for winter.

However, using an OLS method assumes that residuals are distributed normally. This assumption of normality is not suitable when analyzing survival data. The time to an event may not be the same in every case. Therefore, OLS regressions when associated with time distribution data can run into nonsymmetric and bimodal problems. This is where survival analysis proves to be a good method to analyze survival data. The survival method has two main advantages when compared with the traditional OLS method. First, survival methods can distinguish between uncensored and right censored data. Second, it can account for covariates than can change over time. Because of this, OLS regressions are considered as time invariant models. Survival methods can easily incorporate time varying variables when compared to traditional methods.

4.2.1. Survival Analysis

Survival analysis is a method used in many fields, where the survival time until the occurrence of an event can be calculated. It is widely used in the medical field to calculate death and failure rates. In the social science field, this method is applied to calculate the survival or risk rates associated with a particular event taking place.

When an individual or any other applicable entity is chosen to conduct a survival analysis, the transition from one stage to the other is considered as the "event". The starting point of the event can be determined according to the type of the study and varies across different types. It should be noted that this starting point is not necessarily the point of origin. For example, a foreclosure rate for a particular county can be calculated from 2010-2013 using the survival method. If a different time period is chosen, that means it would have a different starting point. The foreclosure is the event that is being observed during the time period.

The mathematical components of survival analysis can be expressed in the following way.

$$F(t) = \int_0^t f(u)d(u) = \Pr(T \leq t) \tag{4.2}$$

T would be defined as a positive random variable denoting survival time (it is considered to be continuous). The actual survival time is denoted as *t*. The probability distribution characterized by a probability density function *f (t)* and cumulative distribution function, *F (t)*

Where the survival time T is less than or equal to some value t, the density function *f(t)*

$$f(t) = \frac{dF(t)}{d(t)} = F^{'}(t) \tag{4.3}$$

The survivor function *S(t)* can be expressed as

$$S(t) = 1 - F(t) = \Pr(\,T\ \leq t) \tag{4.4}$$

Survivor function is the probability of surviving longer than *t.* This is the probability that there is no failure event occurring prior to *t.* Survivor function, in the beginning (t=0) takes the value of 1, and it decreases toward zero as *t* goes to infinity. Therefore survivor function can be described as a monotone non-increasing function in time.

The hazard function which is commonly known as failure rate, conditional failure rate and force of mortality is the instantaneous failure rate, given that a subject has survived until that time. It is the probability that the failure event will occur in an interval, on the condition that the subject has survived to the beginning of a particular interval when divided by the width of the interval. Thus, it can be expressed as

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t + \ \Delta t > T > t | T > t)}{\Delta t} \tag{4.5}$$

The hazard rate is the relationship between the failure time and the survival function which can be expressed as

$$h\ (t) = \frac{f(t)}{S(t)} \tag{4.6}$$

where $h\ (t)$ is the hazard rate, $f(t)$ is the density function and $S(t)$ is the survivor function. Hazard rates can vary from of zero to infinity. A zero hazard rate indicates that the subject faces no risk of experiencing the event and infinity indicates the certainty of the facing an event at that instant itself. The hazard rates can increase, decrease, be constant or even take other shapes.

Another important function that can be derived from the hazard function in survival analysis is the cumulative hazard. Knowing the hazard function gives the ability to determine most of the other functions. The, cumulative hazard can be written as

$$H(t) = \int_{0}^{t} h(u)du \tag{4.7}$$

This enables us to write survival and hazard functions in a more convenient manner. Cumulative hazard function is the total number of expected failure rates for a subject, and it has unit number of failures.

Since the time period is dependent upon the needs of a researcher, there is no mention of the data before and after the time that is specified. This leads to the concept of censoring. Censoring occurs when the survival times are not observed beyond the time period that is mentioned in a study. Not all subjects will experience the event. If a subject hasn't experienced the event by the end time, it is called as "right censored". There is a chance where the subject may have experienced this before the starting time of the study. Such cases are known as "left truncation". If the censored data is omitted from the study, it leads to a biased data set. To avoid the problem of selection bias, binary method together with the likelihood method can be used. However, the survival analysis method has the capability of handling censored data than other methods.

In survival methods the three main models can be considered as the non-parametric, parametric and the semi-parametric models. Each has different assumptions made about the hazard ratio and relationship of covariates in the model. The selection of the method depends on the researcher's particular needs for a study.

4.2.1.1. Non-Parametric Model

In non-parametric method, there are no assumptions made about the distribution of the dependent variable. The estimates are calculated by using the survivor function. The survivor function gives the probability of surviving until an event takes place. As there are no assumptions made about the shape of the hazard ratio and how covariates affect the shape of the hazard ratio, this method has the advantage of understanding more about the data itself and produce more descriptive results. One of the main disadvantages of the method is the inability of comparing many groups and including multiple covariates. The most common non-parametric method is the Kaplan-Meier method.

4.2.1.1.1. Kaplan-Meier

Originally this method was developed by Kaplan and Meier in 1958.Kaplan-Meier (KM) method can take right censoring into account and empirically calculate the probability of surviving past a certain time. Thus, it can describe the survivorship of a population that is being studied. Most studies use this method to compare two study population and their respective survival times. The graphical presentation helps to understand how the survival times change. As discussed earlier, it cannot accommodate covariates and control for time varying variables. Survivor function for the Kaplan-Meier can be written as

$$\hat{S} = \prod_{j\,|t_j\ \leq t} \left(\frac{n_j - d_j}{n_j}\right) \tag{4.8}$$

where $t_1, t_2 \ldots, t_k$ are the failure times, where for each j, n_j is the number of individuals at risk just before t_j, and where d_j is the number of individual who failed at time t_j.

4.2.1.2. Parametric Models

These models can directly exhibit the time dependency when using event history data by specifying a distribution function for the failure rates. This means that the shape of the hazard rate is specified in these models. Unless the shape of the hazard line can be justified with a strong theoretical background, the parametric model will not provide accurate results for any given sample size. Shape of the baseline hazard function directly depends on the covariates due to its parameterization. Therefore, the time dependency of the data changes with the number and the nature of the covariates included in the model. The advantage of this method is the ability to get more precise parameter estimates if the assumptions made about the model are correct. This on the other hand becomes a disadvantage when the correct hazard function is not specified. Furthermore, it will lead to biased parameters estimates and it can be very sensitive in the presence of omitted variables. Examples for some of the parametric models are Exponential, Weibull, Gompertz, Log –Logistic and Log-Normal.

4.2.1.2.1. Exponential Model

Exponential model is a very simple model where the baseline hazard rate is flat. This simply means based on the covariates the risk of an event taking place is same at all time-points. The shape of the hazard line suggests a constant hazard rate. The exponential hazard can be expressed as

$$h_0\,(t) = \exp(a). \tag{4.9}$$

If this is incorporated in to the exponential model then it gives the following equation

$$h\,(t\,|x_j) = \exp(a +\ x_j b). \tag{4.10}$$

This can be implied in terms of hazard and survivor function as

$$H_0(t) = \exp(a)t, \tag{4.11}$$

and

$$S_0(t) = \exp\{-\exp(a)t\} \tag{4.12}$$

In return, the equations can be rewritten for both hazard and survivor methods as

$$H(t|x_j\, b) = \exp(a + x_j\, b\ )\ t \tag{4.13}$$

and

$$S(t|x_j\, b) = \exp\{-exp\ \exp(a + x_j\, b\ )\ t\} \tag{4.14}$$

One of the major disadvantages of using the exponential method is that beyond a certain time period survival time observed still has an exponential distribution. Another flaw of this method is how a single parameter fully determines the distribution. Once the mean is calculated, the variance is considered to be fixed and it cannot be estimated separately from data.

4.2.1.2.2. Weibull Model

The Weibull model is more flexible when compared with the exponential model. The major characteristic of this model is the monotonic nature of the baseline hazard function. It can be monotonically increasing, decreasing or flat with respect to time. An exponential model also gives a flat baseline hazard rate which can be classified as a special case of the Weibull method. The Weibull model allows the hazard to change with respect to time. If the increasing and decreasing hazard rate can be justified for a particular event taking place, then the use of Weibull method can be justified.

The baseline hazard rate of the Weibull method can be expressed as

$$h_0\,(t) = pt^{p-1}\exp(a). \tag{4.15}$$

By substituting this into the hazard model, the equation can be written as

$$H(t|x_j\,) = pt^{p-1}\exp(a + \; x_j b). \tag{4.16}$$

This implies

$$H_0(t) = \exp(a)t^p, \tag{4.17}$$

and

$$S_0(t) = \exp\{-\exp(a)t^p\}. \tag{4.18}$$

Finally, the complete hazard and survival models can be written as

$$H(t|x_j\,b) = \exp(a + \; x_j b)t^p, \tag{4.19}$$

and

$$S(t|x_j\,b) = \exp\{\,-exp\;(a + x_j b)t^p\;\} \tag{4.20}$$

where a is a positive scale parameter and p represents the shape of the parameter as the hazard rate depends on the value of it. When p is greater than 1 the hazard rate is monotonically increasing with time. A less value than p indicates a monotonically decreasing hazard rate with time and when p equals 1 the hazard is flat with a constant value of a .

The Weibull model can be parameterized in terms of its covariates by using a linear model for a random variable T,

$$\log(T) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \beta_j x_{ij} + \sigma\epsilon \tag{4.21}$$

4.2.1.2.3. Gompertz Model

This duration model has a hazard function which is an exponential function of the duration times. The hazard rate for Gompertz model can be monotonically increasing, decreasing or flat. The baseline hazard can be written as

$$h_o(t) = \exp(\gamma t)\exp(a). \tag{4.22}$$

The parameter γ is responsible for controlling the shape of the baseline hazard and this can be either increasing or decreasing exponentially with time. When the sign of the γ is positive, the hazard function will increase with time and a negative γ will decrease the hazard rate with time. When the γ equals zero, the hazard rate is equal to the one that is seen on an exponential distribution. The Gompertz model is capable of restricting the γ to be strictly positive so that when t goes to infinity, the survivor function will exponentially decrease to a nonzero constant. This assures that the survivor function reaching zero as t approaches infinity. More traditional use of Gompertz method has a tendency to not restrict γ because in survival studies subjects are not followed forever and the study usually has a date that it ends.

When the baseline hazard assumed for the Gompertz model is substituted into the hazard model it can be written as

$$h(t|x_j) = \exp(\gamma t)\exp(a + x_j\, b). \tag{4.23}$$

Furthermore, both hazard and survivor functions can be written respectively as

$$H(t|x_j) = \left({}^{1}/_{\gamma}\right)\{\exp(\gamma t) - 1\}\exp(a + x_j b), \tag{4.24}$$

and

$$S(t|x_j b) = \exp[-\{\exp(a + \ x_j b)/\gamma\}\,\{\exp(\gamma t) - 1\}] \tag{4.25}$$

4.2.1.2.4. Log-Normal Regression

The next two models of Log-Normal and Log-Logistic do not assume monotonic hazard rates and allow for nonmonotonic hazard rate. The two models produce similar results for both hazard and survival functions. In the log-normal model, ln (τ) is known to follow a normal distribution. The desirable feature of this distribution is that the nonmonotonic hazard function increases and then it decreases. When the model is written, it would not be written in the usual method where the baseline hazard is expressed separately. It will be expressed as an Accelerated Failure Time (AFT) formula.

The log-normal model can be written as an AFT

$$\tau_j = \exp(-x_j b) t_j. \tag{4.26}$$

The assumption for the lognormal regression can be written as

$$\tau_j \sim lognormal\ (\mu, \sigma). \tag{4.27}$$

When this assumption is used to express the model for log-normal with previous model specifications it can be expressed as

$$\ln(\ t_i) = a + \ x_j b + \ u_j. \tag{4.28}$$

At this point it is no longer a proportional hazard model. However, the inverted U-shaped for log-normal is a good fit for data of repeated nature.

When u_j is assumed to follow a normal distribution, the resulting equations can be written as

$$E(\ln(t_j)|x_j b) = a + \ x_j \mathrm{b}, \tag{4.29}$$

$$E\ (t_j|x_j b) = \exp(a + \ x_j b + \ {}^{\sigma^2}/_{2}, \tag{4.30}$$

$$median\ (\ t_j|x_j b) = \exp(\ a + x_j b). \tag{4.31}$$

4.2.1.2.5. Log-Logistic Regression

The log-logistic distribution is similar to the log-normal distribution, but the ln (τ) is assumed to follow a logistic distribution. There are several advantages of the log-logistics distribution over the log-normal distribution. One of the advantages of it is that log-logistic model has simpler mathematical expressions for both hazards and survivor functions and these expressions do not include integrals.

The two parameters γ and λ determines the distribution, where parameter λ controls the scale of the distribution, and parameter γ controls its shape. If the γ is lower than 1, then the hazard function increases to reach a single maximum at $t = (1/\lambda)\ (1/\gamma - 1)^{\gamma}$ and then it will decrease. If γ is greater than or equal to 1, then the hazard rate is monotonically decreasing. From the AFT formulation it is expressed as

$$\tau_j = \exp(-x_j b)t_j. \tag{4.32}$$

For the log-logistic regression it is assumed that

$$\tau_j \sim loglogistic\ (\lambda, \gamma). \tag{4.33}$$

When λ is substituted with a for consistency purposes, the log-logistic implies the following equation

$$\ln(t_j) = a +\ x_j +\ u_j, \tag{4.34}$$

when u_j assumed to have logistic distribution, the model looks like

$$E\{\ln\left(t_j\right)|x_j b\} = a +\ x_j b. \tag{4.35}$$

One important characteristic that differentiates the log-logistic and log-normal from the Weibull model is the proportional hazards property that applies for the Weibull model.

4.2.1.3. Semi-Parametric Model

As discussed in the previous sections the parametric models yield monotonic hazard rates that are more focused on time. If the distribution of the duration time is known, parametric models are desired. However, many theories and hypotheses require the study of relationship between a dependent variable and the covariates. The form of the distribution is not a major concern.

4.2.1.3.1. The Cox Proportional Hazards Model

The Cox model, first introduced in 1972, is a method where the covariates can be estimated while leaving the duration dependency unspecified. Due to this quality and the increased model flexibility, the semi-parametric Cox model is more desirable than the other parametric models. Because of the increased flexibility of the model, it is called as a semi-parametric model.

Towe and Lawley (2013) use the foreclosure hazard function method in their paper to address reflection problem discussed by previous authors. This study will use a similar method to calculate the foreclosure resolved rate. The Cox proportional hazards model when used as a regression will use the explanatory variables and give out a hazard function. As for the coefficient results, a positive coefficient would indicate the higher chance of the hazard and the negative coefficients would indicate the opposite. The Cox method has the ability to produce results by using several variables at a time. The hazard function can be defined as the probability of an event taking place.

Hazard rate for the i^{th} observation can be written as

$$h_i = h_0(t)\exp(\beta' x). \tag{4.36}$$

The baseline hazard function is denoted by h_0 and covariates and regression parameters are denoted by $\beta' x$. Both Weibull and Cox model have a hazard rate that is proportional. Therefore the ratio for can be expressed as

$$\frac{h_i(t)}{h_j(t)} = exp(\beta'(x_i - x_j)). \tag{4.37}$$

The above equation shows the property of both models having a fixed proportion across time. Even though both Weibull and Cox models have a proportional hazard form, Cox model differentiates itself by having a baseline hazard rate which is unknown and unparameterized. Thus it is called a semi-parametric model. Since it is unparameterized, the regression model does not have an intercept term. In its scalar form the Cox is written as

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \ldots \beta_k x_{ki}) h_0(t) \tag{4.38}$$

The above equation can be rewritten by using the log of hazard model ratios which gives the following equation:

$$\log\{\frac{h(t)}{(h_0(t)}\} = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots \beta_k x_{ki} \tag{4.39}$$

With the given properties of the Cox model, this study will use the hazard model to examine the rate at which a foreclosed house in a regular housing market can be solved by time *t*, given that the house has not been sold until *t*. The hazard rate *h* depends on the housing characteristics C and Neighborhood characteristics N:

$$h(t, C, N) = h_0(t)\exp(\beta_1 C + \ \beta_2 N). \tag{4.40}$$

4.2.2. Hedonic Regression

Hedonic regression is a widely used method in many real estate studies. It gives the ability to estimate the value of a property by including many different categories of components. The value of a house can depend on many factors such as its characteristics, neighborhood characteristics and time of the sale etc. As a result the regression can estimate the impact of each and every individual characteristic. For example, it can include many housing characteristics such as the number of bedrooms, square footage, living area and more. In economics it is used to reveal a consumer's willingness to pay for a certain good. The characteristics included in the regression will reveal how the value of each characteristic will change the dependent variable.

Consistent with most foreclosure studies done by many authors (Immergluck and Smith 2006; Harding et al.,2009; Leonard and Murdoch, 2009; Lin et al.,2009; Rogers and Winter 2009; Daneshvary) this study will attempt to estimate the impact of housing and neighborhood characteristic on sales price of a foreclosed property. So it moves away from the time to resolution variable and now takes the natural log of the sales price as the dependent variable. It is then regressed on housing, neighborhood characteristics, and dummy variables such as school districts, year of the foreclosure and the winter dummy variable. The equation can be written as

$$\ln(P) = \alpha + \beta C + \delta N + \gamma D + \varepsilon, \tag{4.41}$$

where the dependent variable is the natural log of the foreclosed sales price (P), C is a vector of housing characteristics, N is a vector of neighborhood characteristics and D represents all the dummy variables including house conditions , school districts, foreclosure years and a seasonal dummy.

CHAPTER 5. RESULTS AND DISCUSSION

5.1. OLS Regression for Time to Resolution

The main dependent variable of this study, time to resolution, was regressed against all independent variables of housing, neighborhood and dummies to obtain the coefficient values from the OLS method. The negative coefficient values show that there is a negative impact on the dependent variable. In this case, it means that a negative coefficient will help decrease the time it is vacant and help sell the property faster. The positive coefficient values will increase the time that it is vacant and not help a property to be sold fast. All the results for OLS regression for time to resolution is reported in Table 5.1.

Housing characteristics such as the number of stories (story1, story 1.5 and story 2) help to sell the property faster, once it is foreclosed. Characteristics such as slab foundation, central heat, living area, fireplace and the efficiency age of the house all slow down the process of resolution. Only one school district dummy is significant and a property in Highland Park will increase the time of the property being vacant by as high as 9 months.

An increase in the percentage of school children in the neighborhood increases the speed of a property being resolved by 2.6 months. According to the results the percentage of white people in a neighborhood increases the time a property is vacant and it is about 2.04 months. All the foreclosure year dummy variables are significant and they all help to reduce the time a property is vacant. The winter dummy variable is significant at 1% and hinders the sale of a foreclosed property by 1.03 months. The OLS regression reports an R-squared of 0.11only. The OLS method used here only looks at uncensored data (only the foreclosed properties that were sold). It fails to take the censoring nature of the data and the time variant variables into account when estimating the model. The results obtained from the Cox proportional hazard model provide a more accurate set of coefficients for the variables and the likelihood of a sale for a foreclosed property taking place, as it has the capacity to account for censoring and time variant data.

Table 5.1. Coefficients of Time to Resolve OLS Regression

Time to Resolve OLS- Abridged			
Variable	**Coefficient**	**95% Confidence Interval**	
Constant	22.87205*** (4.73883)	13.58362	32.16048
Attached Garage	-0.195586 (0.3084479)	-0.8001651	0.4089931
Attached Carport	0.1878966 (0.5075227)	-0.8068828	1.182676
Detached Carport	0.2415704 (0.7618613)	-1.25173	1.734871
Pool	-0.2184126 (0.3658205)	-0.935446	0.4986208
Slab Foundation	0.8380121** (0.3284502)	0.1942272	1.481797
Central Heat	1.08114* (0.6326794)	-0.1589551	2.321235
Central Air Conditioning	-0.9177252 (0.5972709)	-2.088417	0.2529668
Baths	0.4033117 (0.2598347)	-0.1059821	0.9126055
Lot Area	-0.0133282 (0.0111259)	-0.0351357	0.0084794
Living Area	0.5028554** (0.2330589)	0.046044	0.9596668
Fireplace	0.9281334*** (0.2618168)	0.4149545	1.441312
1 Story	-10.22056*** (3.719653)	-17.51133	-2.929782
1.5 Story	-10.04148*** (3.717873)	-17.32876	-2.754192
2 Story	-9.723288*** (3.707461)	-16.99017	-2.456411
Effective Age	0.0385671*** (0.0086951)	0.0215241	0.0556101
Community Percentage of School Aged Children	-2.669003* (1.567407)	-5.741228	0.4032218
Community Percentage Over 60	-0.8624788 (1.673052)	-4.141775	2.416817
Community Percentage Below Poverty	1.094134 (1.002906)	-0.8716297	3.059898

Table 5.1. Coefficients of Time to Resolve OLS Regression (Continued)

Time to Resolve OLS- Abridged (continued)			
Variable	**Coefficient**	**95% Confidence Interval**	
Community Percentage White	2.043418*** (0.5492129)	0.9669228	3.119913
Condition Dummy	Yes		
School District Dummy	Yes		
Seasonal Dummy	Yes		
Foreclosure Year Dummy	Yes		
Number of Observations	22698		
R-squared	0.1189		
Adj R-squared	0.1172		
F-Statistic	69.47		

5.2. Cox Proportional Hazard Model

As discussed in Chapter 4 section 4.2.1.3, the Cox proportional method was used to obtain the hazard ratios for housing and neighborhood characteristics. The results are reported with the relevant coefficient and hazard ratios in Table 5.2.When interpreting, a hazard ratio above 1 implies that the independent variable increases the chance of that particular property being resolved faster. A hazard ratio below 1 implies that independent variable decreases the chance of a particular property being resolved. For example, a hazard ratio of 0.94 for a binary variable means that the presence of that characteristic will decrease the chance of a property being resolved by 6%. If a continuous variable, such as living area, it can be interpreted as the increase of an additional square footage would lead to an increase/decrease in the time to resolution. When coefficients are presented, it can easily be converted to its hazard ratio by taking the exponential value of the coefficient.

5.2.1. Housing Characteristics

The results from the Cox model for most housing characteristics are significant. Characteristics such as an attached garage, a pool, central heat, central air conditioning, amount of living area, one story, one and a half stories and two stories are all significant and increase the chance of a property being resolved faster. For example, the hazard ratio of 1.0967 for an attached garage implies that the presence of an attached garage will increase the likelihood of the house being sold by 10%. Kluger & Miller (1990) use a Cox proportional method for their study to observe the sale probability of regular houses in Columbus Ohio by housing characteristics. They report how having an extra bedroom can decrease the expected time on market for a property. Even though this study will not have a separate bedroom variable, living area and number of bedrooms are highly correlated enough that living area captures the bedroom variable details as well. The results from the Cox model found that when the living area is increased by 1000 square feet, the likelihood of being resolved increases by 12%. But these results cannot be directly compared to the results of Kluger & Miller (1990) even though they use a similar methodology.

Housing characteristics such as attached carport, detached carport, houses with a slab foundation, size of the lot, and effective age are all significant, but would hinder the chance of a property being sold. Even though lot area decreases the time to resolution, the hazard ratio of it only records a value that is less than 1%. In that sense, it only has a minimal impact on the dependent variable. Kluger & Miller (1990) also find that the lot area has a negative impact on the time on market for a property. They also find a minimal impact for lot area, which is less than 1% (the reported coefficient value was -0.000048 and the exponential value of that gives a hazard ratio of 0.999952). However, these values are for general properties that are up for sale and not for foreclosed properties specifically. But their research still helps to understand about desirable housing characteristics in the real estate market.

The results obtained for the variables story 1, story 1.5 and story 2 indicate that they are twice as likely to be sold in the market when compared to houses with stories 2.5 and above. The hazard ratio of 2.75 for story 1 shows that a house with one story will sell 2.75 faster than the base case in the study (story 2.5 and above). As the number of stories increase, the hazard ratio also decreases, but this is only a small change. It appears that the number of stories is a major factor when deciding to buy a house and especially when it comes to a foreclosed property.

The hazard ratio of 0.9317 for attached carport implies that the presence of an attached carport will decrease the chance of the house being sold by 7%. The effective age variable implies that each additional year added to the property will decrease the chance of that property being resolved by less than 1%. Once again, as the value is less than 1% it can be implied that the impact would be minimal. It also shows that if a house is newer in age or has been refurbished recently, it will likely be resolved faster.

The condition dummy variable capturing the quality of the house has some interesting results. The conditions that are significant are very good, fair and poor. All three conditions help to reduce the time to resolution. A house in very good condition will help resolve a property by 12%, a fair condition by 21% and a poor condition by 15%. Intuitively one would expect above average houses to resolve faster than the houses below average would. According to the results, a poor conditioned house is likely to be resolved faster than a house in a very good condition. One possibility could be the price associated with each condition, where a low quality house would be sold at a price that is affordable to a buyer or if a buyer plans to refurbish and sell the house, a low quality house may be more desirable.

Towe and Lowley (2013) in their study find evidence of a low quality house increasing the hazard of foreclosure by 66%. They also find that both good and very good quality houses are less likely to foreclose. According to this study, properties with fair conditions stand a better chance of being resolved.

Table 5.2. Cox Proportional Hazard Model Estimates

Cox Proportional Hazard Model - Abridged				
Variable	**Coefficient**	**Hazard Ratio**	**95% Confidence Interval**	
Attached Garage	0.092386***	1.096788	1.051254	1.144295
	(0.0237284)			
Attached Carport	-0.070677**	0.9317628	0.8693409	0.9986668
	(0.0329655)			
Detatched Carport	-0.3062342***	0.7362142	0.6625904	0.8180187
	(0.0395776)			
Pool	0.0667578***	1.069037	1.016555	1.124227
	(0.0274563)			
Slab Foundation	-0.0974065***	0.9071872	0.8672384	0.9489761
	(0.0208448)			
Central Heat	0.1728756***	1.188718	1.089382	1.297112
	(0.0529261)			
Central Air Conditioning	0.193649***	1.21367	1.117157	1.318522
	(0.0513108)			
Baths	-0.0192058	0.9809774	0.9473483	1.0158
	(0.017459)			
Lot Area	-0.0019085**	0.9980934	0.9965213	0.9996679
	(0.0008027)			
Living Area	0.1141445***	1.120914	1.087418	1.155442
	(0.0173508)			
Fireplace	0.0119975	1.01207	0.9761918	1.049266
	(0.0186378)			
1 Story	1.013217***	2.754447	1.678482	4.520141
	(0.6961101)			
1.5 Story	0.9126388***	2.490887	1.517327	4.08911
	(0.6299626)			
2 Story	0.8952382***	2.447919	1.493019	4.01355
	(0.6175339)			
Effective Age	-0.0029136***	0.9970907	0.9958647	0.9983182
	(0.0006259)			
Community Percentage of School Aged Children	0.4175565***	1.518247	1.225541	1.880862
	(0.1659056)			
Community Percentage Over 60	-0.1956522*	0.8222982	0.6518004	1.037395
	(0.097488)			
Community Percentage Below Poverty	-0.4019829***	0.6689922	0.5830795	0.7675635
	(0.0469153)			

Table 5.2. Cox Proportional Hazard Model Estimates (Continued)

Cox Proportional Hazard Model - Abridged				
Variable	**Coefficient**	**Hazard Ratio**	**95% Confidence Interval**	
Community Percentage White	0.0405178	1.04135	0.9662476	1.122289
	(0.0397702)			
Condition Dummy	Yes			
School District Dummy	Yes			
Seasonal Dummy	Yes			
Foreclosure Year Dummy	Yes			
Number of Observations	30581			
Number of Failures	21828			
Log-Likelihood	-212569.35			

5.2.2. Neighborhood Characteristics

The neighborhood characteristics tested for the Cox method are significant with the exception of the race variable white. Racial composition still acts as a significant characteristic, when buyers make decisions related to purchasing a house. However, this study did not find any evidence to support that notion. Both school age and below poverty variables are significant at 1%, while the community greater than 60 variable is significant at 10%. The presence of school aged children when increased by 1% will increase the chance of resolution by 52%. Both variables of community greater than 60 and below poverty decrease the chance of being resolved.

Most school district dummies used in the study are insignificant except for school district 5 and 7. These school districts are Highland Park and Lancaster Independent School District. The hazard ratio for both indicates that a foreclosed property in Highland Park and Lancaster Independent School District will decrease the chance for resolution by 56% and 32% respectively. Foreclosed year dummy variables are all significant and report the chance of resolution by more than 100% when compared to the base year of 2005. The seasonal dummy variable winter indicates that most houses foreclosed in winter are likely to be resolved 2% faster.

5.2.3. Kaplan-Meier Survival Estimates

By using the non-parametric approach, the survival estimates can be obtained. The most popular method is the Kaplan- Meier survival estimates. Figure 5.1 displays the Kaplan-Meier survival estimates for the foreclosed properties.

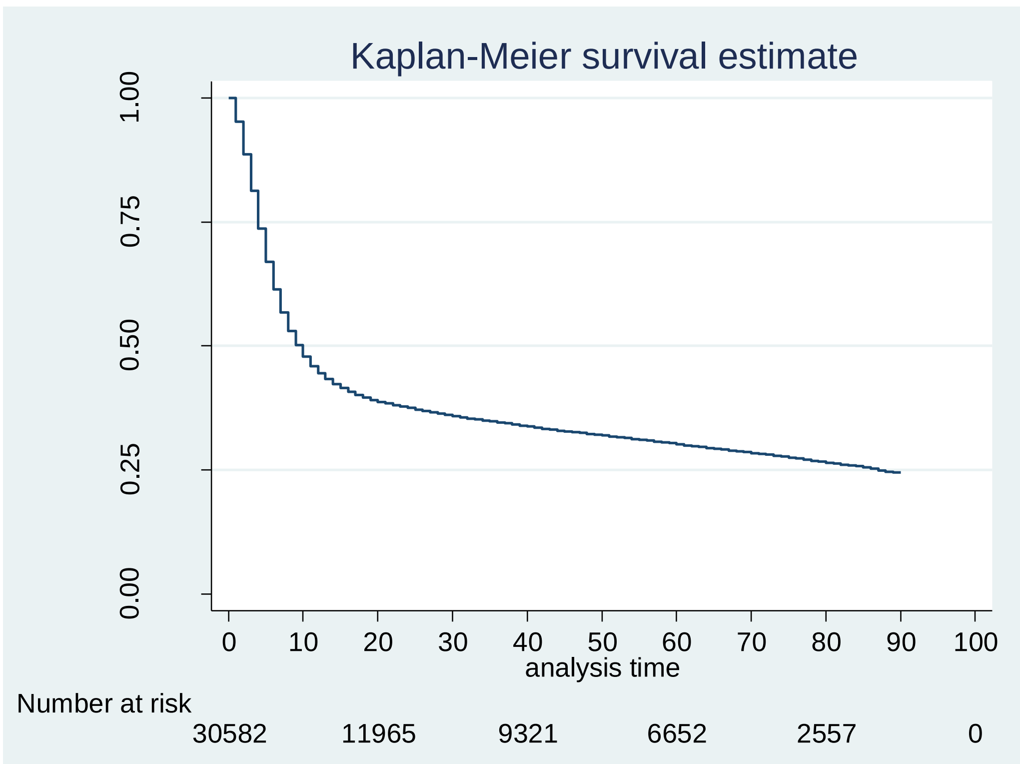

Figure 5.1. Kaplan- Meier Survival Estimate

This graph displays the survival curve for the data set and the number of subjects at risk for each time period. On the X axis, the time to resolution is displayed in number of months and on the Y axis the survivor function is displayed. When t=0, the total number of subjects at risk is 30, 582. This is number of subjects that face the risk of facing the event of being sold. In the beginning the survival rate is 1. At time 1 (t=1), 1474 subjects fail, indicating that they have faced the event of being resolved. Therefore at t=1 the survival function can be calculated as 0.9518. Given in Table 5.3 is the summary for survival times in periods of 10, and a full table of survival functions is attached in the appendix section.

Table 5.3. Survival Time Estimates

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
1	30582	1474	0	0.9518	0.0012	0.9493	0.9541
10	15328	700	0	0.4783	0.0029	0.4727	0.4839
20	11965	128	0	0.3871	0.0028	0.3816	0.3925
30	10828	76	93	0.358	0.0027	0.3526	0.3633
40	9321	57	111	0.3372	0.0027	0.3319	0.3426
50	8070	45	126	0.319	0.0027	0.3138	0.3243
60	6652	46	103	0.3014	0.0027	0.2962	0.3067
70	5305	29	50	0.284	0.0027	0.2788	0.2893
80	2557	18	273	0.2644	0.0028	0.2590	0.2698
90	7	0	7	0.245	0.0034	0.2384	0.2517

5.2.4. Nelson-Aalen Cumulative Hazards

The cumulative hazards are obtained by using a non-parametric method of the Nelson-Aalen estimator. In non-parametric studies there are no assumptions made and this helps to compare how two groups will have different hazard rates. Given below are four graphs for variables of pool, central air condition, house with a fair condition and school district 5 (Highland Park) showing the cumulative hazard rate for variables used in the study. Both pool 5.2 (a) and central air conditioning 5.2 (b) variables show that houses featuring them are more likely to face the event of being resolved. Cumulatively, there is a big difference over time and one can say that pool and central air condition are desirable characteristics for a house in Dallas County, as one might imagine. The house characteristic condition 3, which is the fair condition5.2 (c) show that over time the likelihood of the event happening is not that far apart. While School district 5, Highland Park, shows that a house in that school district is less likely to be resolved when compared to a house located in another school district of the county.

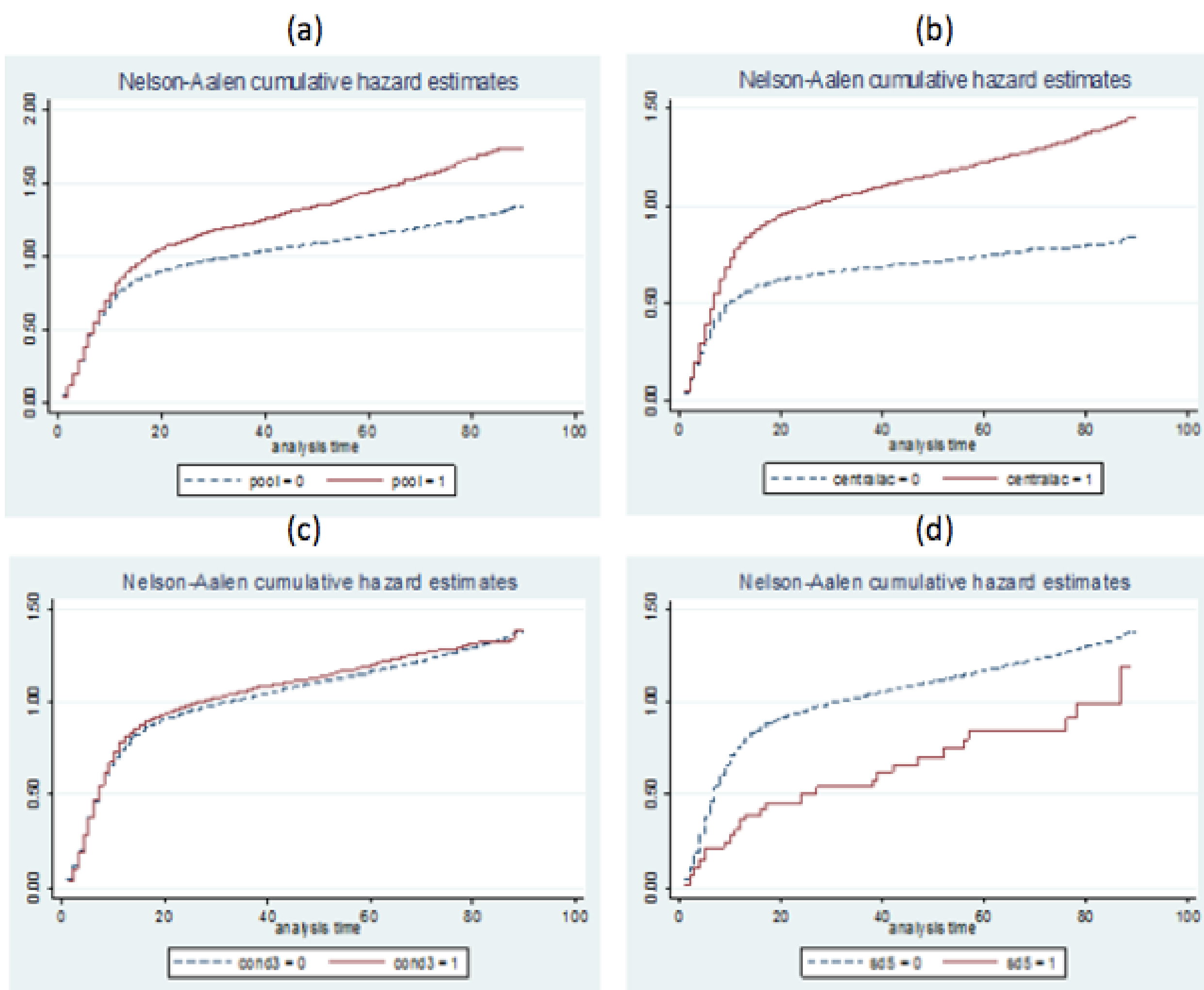

Figure 5.2. Nelson-Aalen Cumulative Hazard Graph for Pool, Central Air Conditioning, Condition 3 and School District 5

Once the variables are fitted into the Cox proportional model, the graphs can be obtained for survival, hazard and cumulative hazard functions. The hazard function in this case, Figure 5.3, decreases over time. It decreases until 40 months and increases slightly thereafter. By the time it reaches 40 months, the chance of being sold in the market decreases, increases a little bit until it reaches 70 months and finally starts to decrease again.

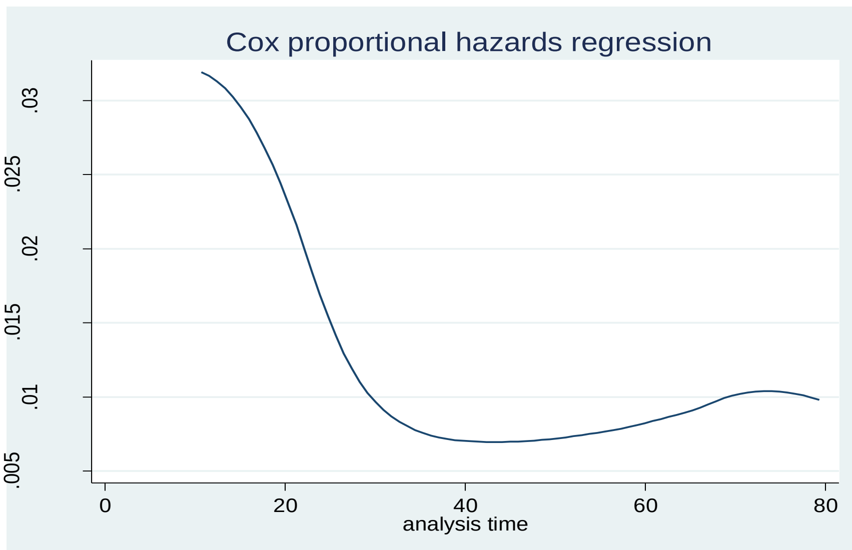

Figure 5.3. The Smoothed Hazard Graph from Cox Proportional Hazards Regression

The cumulative hazard graph, Figure 5.4, considers all the hazard ratios for the properties and graphs them against time. The graph for cumulative hazards increases at a decreasing rate until it reaches the 20 months and thereafter follows a small more consistent increase.

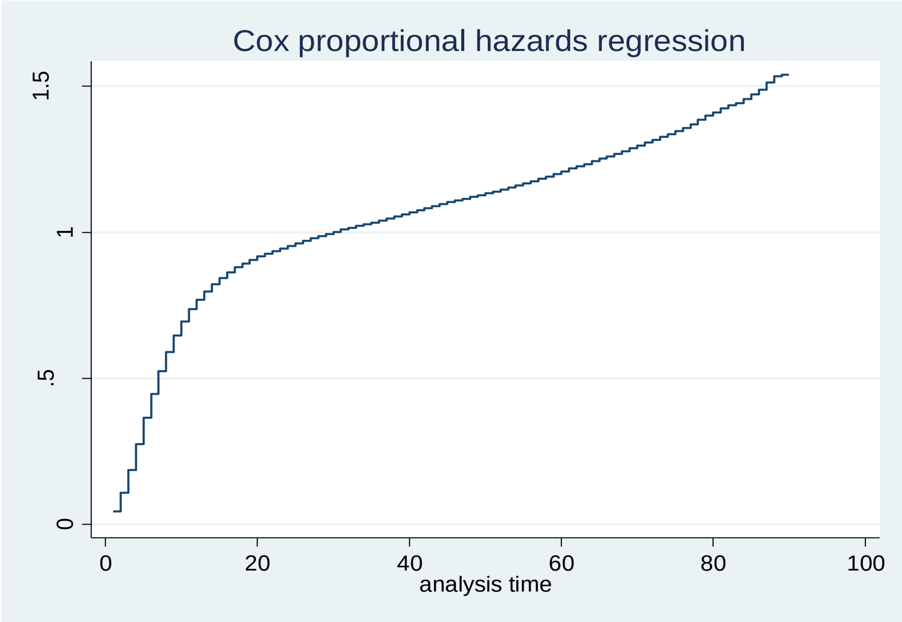

Figure 5.4. The Cumulative Cox Proportional Hazards Regression Graph

After fitting the covariates into a Cox model, the cumulative hazard ratios can be obtained. For illustrative purposes, the same four variables are used in graphical form to compare the cumulative hazard ratios before and after fitting the covariates into the model. These graphs can be seen in Figure 5.5. The previous cumulative graphs for covariates (Figure 5.4) were estimated on raw data and using a non-parametric method.

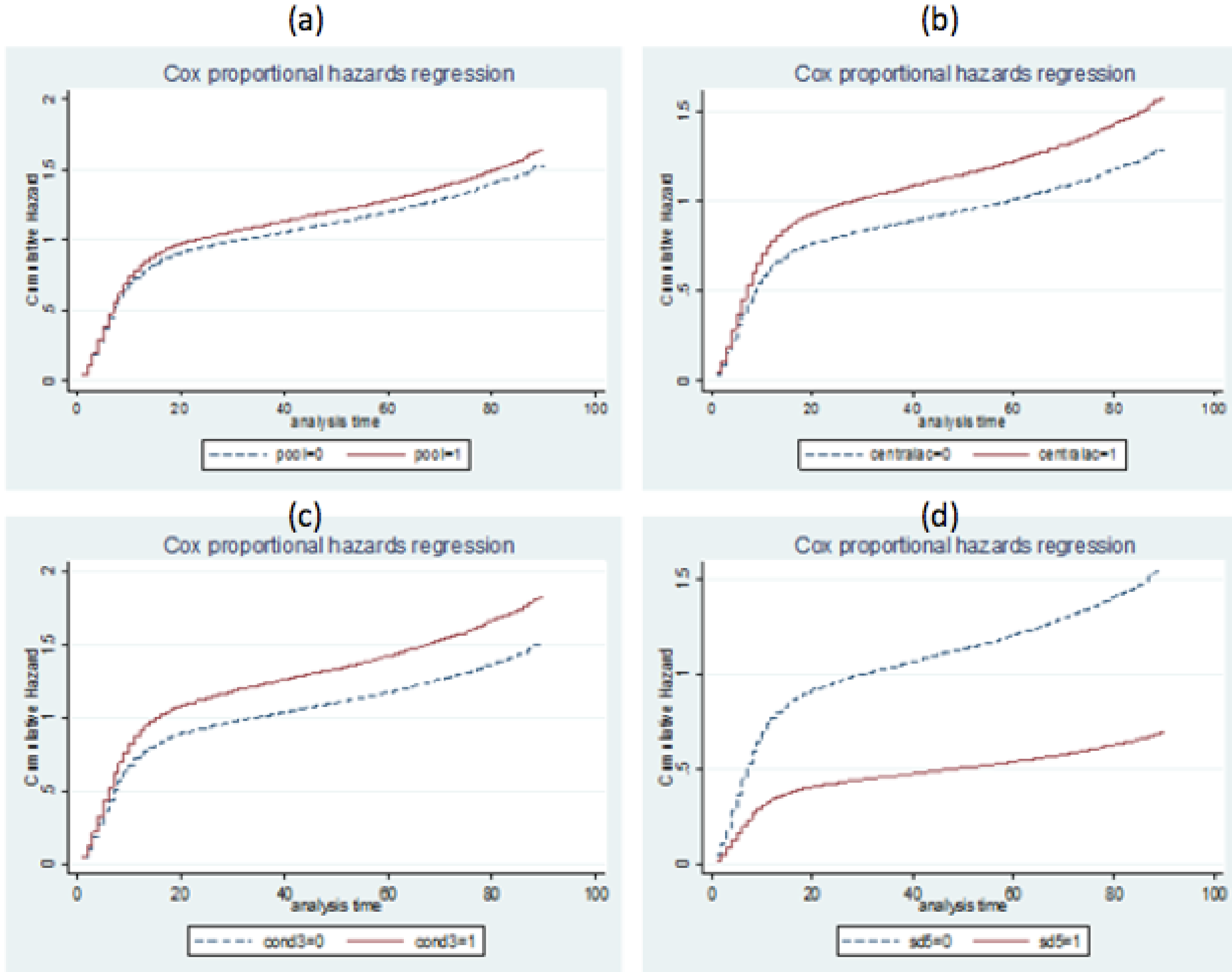

Figure 5.5. Cox Proportional Hazards Cumulative Graphs for Variables Pool, Central Air Conditioning, Condition 3 and School District 5

The cumulative hazard ratio for the variable pool, over time does not have as large a gap as shown in the Nelson- Aalen graph before. Until 10 months, the house without a pool stands about an equal chance of being resolved, however after that time period houses with a pool are likely to resolve faster. Looking at the central air conditioning variable, it is indicated that a house with that variable is preferred over a house without. Cumulatively, this difference starts at an early time period and continues to maintain the gap. A house in fair condition is more likely to be resolved faster. The effect is large, especially when compared with the corresponding Nelson-Aalen graph for the variable. Cumulatively, a house in a fair condition is likely to be resolved faster than any other house condition. It once again starts at an early stage and increases over time. The largest effect can be seen by school district 5, Highland Park. A property located in this school district is less likely to be resolved and the difference between the two curves increases over time.

5.2.4. Assessing the Goodness of Fit for Cox Proportional Methods

A Cox-Snell residual is the difference between the observed data point and the predicted value. The Cox-Snell equals the negative of the log of the survival time, written as: $-\log S(t_i)$. The Kaplan-Meier estimator is computed on the Cox-Snell residuals and from those estimates the integrated hazards are estimated. If the plot of the integrated hazard against the hazard rate estimates falls along a 45˚ line meaning that the distribution of the Cox-Snell is exponential and then the model is a good fit. Figure5.6 shows the Cox-Snell residual for the model used for this study. Looking at the graph, the model seems to be a good fit for most parts where towards the end it moves away from the reference line.

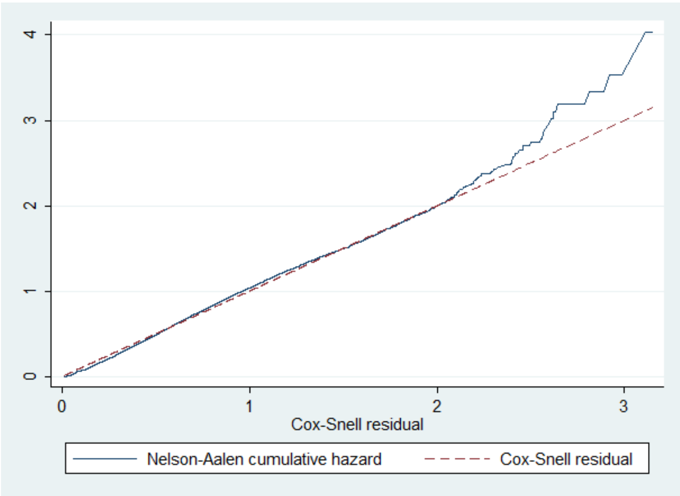

Figure 5.6. Cox-Snell Residual

5.3. Hedonic Pricing Model

5.3.1. Housing Characteristics

Results obtained from the OLS regression method for sales prices show some interesting results. Most housing characteristics are significant with the exception of attached carport, detached carport, lot area, and the number of the stories in the house. The presence of characteristics such as attached garage, pool, central heat, central ac, baths, living area and fire place all are positively significant in determining the sales price of the property. The increase of 1000 square footage in the living area contributes to a 33% increase in the price. The study found that properties with a slab foundation will have a decreased sales price by 10%. Efficiency age of the property increase can also negatively affect the sales price of the foreclosed property. The number of stories in a house is not significant for all three levels tested out in the study. In other words, whether a house has more than one story or not doesn't have any impact on the sales price of a foreclosed property.

Table 5.4. Results for Hedonic Regression (OLS)

Natural Log of Sales Price OLS- Abridged			
Variable	**Coefficient**	**95% Confidence Interval**	
Constant	10.2362***	10.02643	10.44596
	(0.1070189)		
Attached Garage	0.0344818***	0.0208283	0.0481353
	(0.0069658)		
Attached Carport	-0.010945	-0.033411	0.0115205
	(0.0114616)		
Detached Carport	0.0261764	-0.007547	0.0599002
	(0.0172054)		
Pool	0.0667078***	0.0505147	0.0829008
	(0.0082615)		
Slab Foundation	-0.1041215***	-0.11866	-0.0895826
	(0.0074175)		

Table 5.4. Results for Hedonic Regression (OLS) (Continued)

Natural Log of Sales Price OLS- Abridged (Continued)			
Variable	**Coefficient**	**95% Confidence Interval**	
Central Heat	0.0751414***	0.0471359	0.103147
	(0.0142881)		
Central Air Conditioning	0.1696838***	0.1432456	0.196122
	(0.0134884)		
Baths	0.0636866***	0.052185	0.0751882
	(0.005868)		
Lot Area	0.0003434	-0.000149	0.0008359
	(0.0002513)		
Living Area	0.3315146***	0.3211982	0.341831
	(0.0052633)		
Fireplace	0.1258583***	0.114269	0.1374476
	(0.0059127)		
1 Story	0.0334408	-0.13121	0.1980913
	(0.0840024)		
1.5 Story	-0.0035174	-0.168089	0.1610544
	(0.0839622)		
2 Story	-0.0157234	-0.179834	0.1483874
	(0.0837271)		
Effective Age	-0.0045981***	-0.004983	-0.0042132
	(0.0001964)		
Community Percentage of School Aged Children	0.2073673***	0.137986	0.2767486
	(0.0353974)		
Community Percentage Over 60	-0.7615106***	-0.835568	-0.6874529
	(0.0377832)		
Community Percentage Below Poverty	-0.526034***	-0.570428	-0.4816404
	(0.022649)		
Community Percentage White	0.6305285***	0.6062176	0.6548395
	(0.0124031)		
Condition Dummy	Yes		
School District Dummy	Yes		
Seasonal Dummy	Yes		
Foreclosure Year Dummy	Yes		
Number of observations	22698		
R-squared	0.7391		
Adj R-squared	0.7386		
F Statistic	1458.65		

5.3.2. Neighborhood Characteristics

All neighborhood characteristics are significant with both positive and negative impacts as seen in Table 5.2. The increase of the white percentage in the neighborhood would lead to a price increase of by 0.63%. In real estate literature there is evidence that racial composition has an impact on the price of a house (Kim, 2008). It is widely believed that neighborhoods composed with high white percentages have higher house prices when compared to house prices from non-white neighborhoods. The result in this study shows the role of racial composition in determining house prices for foreclosed properties as well.

The percentage increase in school aged children in the neighborhood would increase the prices by 0.2%. Intuitively one can expect a negative relationship between sales prices and poverty. As expected the increase of below poverty percentage would lead to a decrease of 0.5%. Also the percentage of people who are above 60 in the community can also have a negative impact on the prices by 0.76%. While these values may seem low, they still indicate that these characteristics have a role to play when determining the prices and it can be a positive or negative relationship.

The house condition dummy variables clearly show a direct relationship with the sale price. All dummy variables for this are significant with positive relationships. A house in an excellent, good, average or poor would have an impact of 38%, 30%, 24% or 13% respectively. When determining the sales price, the condition of the property has a great influence.

Out of the 13 school districts dummies tested out to control for any geographical effects, 7 of the districts are significant. School district of Highland Park has a 78% impact on the sales price. This is the highest impact that is seen among all the other significant school districts. This estimate could be a result of some preferences from the buyers' side capturing some geographical and other local trends. Highland Park is also one of the two school districts that has the exemplary rating.

The location of the property in both Lancaster ISD and Mesquite ISD both lead to a decrease in the sales price by 16%. Lancaster is an academically acceptable school district and Mesquite is a recognized school district according to the rates given to each of the school districts.

In real estate markets, seasonality tends to play an important role. The study finds that a house foreclosed in winter seasons will increase the house price by 0.8%. Most winter listings have a tendency to be sold by spring and summer season. There is a general tendency to look for a house in the spring or summer season than in the fall or winter seasons. This could be the reason why the prices of foreclosed properties have a positive impact with the seasonality. In order to determine whether the year of the foreclosure plays a role on the sale price, the study included five foreclosure year dummy variables and they all are significant. All years from 2006-2010 show evidence of having a negative impact on the sale price. For a house foreclosed in year 2006 the impact is a negative 0.5% and it increases up to a negative 21% in 2008. The houses foreclosed in 2010 the impact is 22%. The higher decrease in the price for years 2008-2010 could reflect the market conditions prevailed at that time.

CHAPTER 6. CONCLUSIONS

The main purpose of this study was to determine whether housing and neighborhood characteristics have an impact on the sale of foreclosed properties, and if it holds true to identify the desirable characteristics that help resolve a property faster. For this Dallas County, Texas data was used from the years 2005- 2010. The data included a list of foreclosures, sale prices, housing characteristics and neighborhood characteristics. The time varying nature of data weakens the results obtained by traditional methods such as OLS. This is where survival methods can help accompany time variant covariates. Cox proportional hazard method allows the inclusion of time varying covariates yet does not impose any assumptions on the shape of the baseline hazard. Therefore, it allows the baseline hazard to take any shape and yet calculate the impact of covariates on the base line hazard rate.

The low R-squared of 0.1189 reported by the OLS regression for the dependent variable of time to resolution, proved the chosen method is not adequate and it only takes uncensored data into account. When the dependent variable time to resolution was used in the Cox model, housing characteristics such as attached garage, pool, central heat, central air conditioning, living area, fair house condition and stories from 1- 2 are all highly significant and help to sell a foreclosed property faster. On the other hand, housing characteristics such as detached carport, houses with a slab foundation and the efficiency age of the house are highly significant factors that hinder the chances of a foreclosed property being sold.

Neighborhood characteristics such as the percentage of school aged children in the neighborhood and the percentage of people who are below the poverty line are all highly significant with two different impacts on the resolution rate; where the school age variable helps to resolve a property faster and below poverty slows down the sale of a foreclosed property. There was no evidence to support the idea of buyers placing value on the racial composition in the neighborhood.

The results from hedonic regression for identifying characteristics that have an impact on the sale price of foreclosed properties show that characteristics such as attached garage, pool, central heat, central air conditioning, the number of bathrooms, living area, and the condition of the house are highly significant and have a positive impact on the sales price. The variable slab is highly significant, but it has a negative impact on the sales price. The results also reveal how neighborhood characteristics have both negative and positive impacts on the sales price of a foreclosed property.

6.1. Limitations

This study sheds light on the speed of a sale of a foreclosed property and the characteristics that impact it. However, there are limitations associated with the study and identifying them may help future studies. First and foremost, the data collected for this study is county level and comes from one of the counties in Texas. Texas is the second largest economy in the United States and during the financial crisis the impact of it on the state was minimal; especially in the housing market. As a result of a more stable economy and a less volatile housing market, there exists a chance of the observed foreclosure activity level in the state could be less than what was observed from other states.

The dependent variable used in the hedonic regression can be considered as a choice variable and it can only be observed when the property is foreclosed. Therefore, both housing and neighborhood characteristics used are the determinants when evaluating the price of a foreclosed house. One of the major limitations of a hedonic model is determining the variables that need to be included in the model. Even though it helps to include a wide variety of variables representing the most functional form of a house, there could be other types of variables that are not included in the model that may or may not add be preferred by buyers in Dallas County.

6.2. Need for Future Study

In the future, it would be beneficial if a study could be done with more county level data and find whether there are similarities across counties. Housing markets are heterogeneous and can have local differences, but a study done on different counties for the same state can help understand whether there are characteristics that are valued in general more than others. Another area that can be explored is comparing housing and neighborhood characteristics for both distressed and non-distressed properties for a single market. This would provide evidence that buyers may have different preferences depending on the status of the house.

REFERENCES

Bailey, M.J., R.F. Muth and H.O. Nourse "A regression method for real estate price index construction." *Journal of the American Statistical Association*, Vol. **58**, (1963) pp. 933-942.

Box-Steffensmeier, J.M. and B.S. Jones *Event history modeling: A guide for social scientists*, Cambridge University Press, (2004).

Campbell, J.Y., S. Giglio and P. Pathak "Forced sales and house prices": National Bureau of Economic Research, (2009).

Dallas County "Dallas County, Texas Public Schools By Independent School Districts", (2014).

Dallas County and D.C. Clerk "Foreclosure Notices", (2009).

Daneshvary, N. and T.M. Clauretie "Toxic Neighbors: Foreclosures and Short- Sales Spillover Effects from the Current Housing- Market Crash." *Economic Inquiry*, Vol. **50**, (2012) pp. 217-231.

Daneshvary, N., T.M. Clauretie and A. Kader "Short-term own-price and spillover effects of distressed residential properties: The case of a housing crash." *Journal of Real Estate Research*, Vol. **33**, (2011) pp. 179-207.

Ellen, I.G., J. Lacoe and C.A. Sharygin "Do foreclosures cause crime?" *Journal of Urban Economics*, Vol. **74**, (2013) pp. 59-70.

Ellen, I.G., A.E. Schwartz, I. Voicu and M.H. Schill "Does federally subsidized rental housing depress neighborhood property values?" *Journal of Policy Analysis and Management*, Vol. **26**, (2007) pp. 257-280.

Farley, R., E.L. Fielding and M. Krysan "The residential preferences of blacks and whites: A four- metropolis analysis." *Housing Policy Debate*, Vol. **8**, (1997) pp. 763-800.

Foreclosurelaw.org "Texas Foreclosure Law", (2014).

Forgey, F.A., R.C. Rutherford and M.L. VanBuskirk "Effect of foreclosure status on residential selling price." *Journal of Real Estate Research*, Vol. **9**, (1994) pp. 313-318.

Gopal, P. "Housing Boom Bigger in Texas as Home Bidding Wars Erupt": Bloomberg, (2013).

Harding, J.P., E. Rosenblatt and V.W. Yao "The contagion effect of foreclosed properties." Vol. **66**, (2009) pp. 164-178.

Harding, J.P., E. Rosenblatt and V.W. Yao "The contagion effect of foreclosed properties." *Journal of Urban Economics*, Vol. **66**, (2009) pp. 164-178.

Immergluck, D. and G. Smith "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values." *Housing Policy Debate*, Vol. **17**, (2006) pp. 57-79.

Immergluck, D. and G. Smith "The impact of single-family mortgage foreclosures on neighborhood crime." *Housing Studies*, Vol. **21**, (2006) pp. 851-866.

Indiana Department of Local Government Finance "The Sales Comparison Approach", (2013).

KeepmyTexashome.org "The Foreclosure Process", (2014).

Kim, S. "Race and home price appreciation in urban neighborhoods: Evidence from Milwaukee, Wisconsin." *The Review of Black Political Economy*, Vol. **28**, (2000) pp. 9-28.

Kluger, B.D. and N.G. Miller "Measuring residential real estate liquidity." *Real Estate Economics*, Vol. **18**, (1990) pp. 145-159.

Kobie, T.F. and S. Lee "The Spatial-Temporal Impact of Residential Foreclosures on Single-Family Residential Property Values." *Urban Affairs Review*, Vol. **47**, (2011) pp. 3-30.

Lee, K.-y. "Foreclosure's price-depressing spillover effects on local properties: A literature review." *Federal Reserve Bank of Boston Community Affairs Discussion Paper*, Vol. **1**, (2008) pp. 1-11.

Leonard, T. and J.C. Murdoch "The neighborhood effects of foreclosure." *Journal of Geographical Systems*, Vol. **11**, (2009) pp. 317-332.

Lin, Z., E. Rosenblatt and V. Yao "Spillover Effects of Foreclosures on Neighborhood Property Values." *The Journal of Real Estate Finance and Economics*, Vol. **38**, (2009) pp. 387-407.

McLeod, D. "Describing the distribution of failure times", (2013).

Moreno, A. "The cost-effectiveness of mortgage foreclosure prevention." *Minneapolis: Family Housing Fund*, (1995).

National Association of Realtors "Housing Affordability Index", (2014).

RealtyTrac "Market Summary", (2014).

Rohe, W. and H. Watson. *Chasing the American Dream: New Perspectives on Affordable Homeownership*. Cornell University Press. (2007)

Rogers, W.H. and W. Winter "The Impact of Foreclosures on Neighboring Housing Sales." *Journal of Real Estate Research*, Vol. **31**, (2009) pp. 455-480.

Sill, M. "Chapter 5: Cox Proportional Hazards Model", (Roswell Park Cancer Institute: GOG Statistical and Data Center, 2004).

Skogan, W.G. *Disorder and decline: Crime and the spiral of decay in American neighborhoods*, Univ of California Press, (1992).

Swiss Federal Institute of Technology "Cox Proportional Hazard Model andits Characteristics": Departement Mathematik,(2011).

Texas A&M Real Estate Center. "MARKET DATA SOURCES", (2014).

Texas Education Agency "2011 Accountability Rating System", (2011).

Towe, C. and C. Lawley "The Contagion Effect of Neighboring Foreclosures." *American Economic Journal: Economic Policy*, Vol. **5**, (2013) pp. 313-335.

United States Census Bureau. "Census Data", (2007).

United States Census Bureau. "Census Data", (2010).

United States Department of Housing and Urban Development, (2014).

United States Securities and Exchange Commission "Mortgage-Backed Securities", (2014).

Zhang, L. and T. Leonard "Neighborhood impact of foreclosure: A quantile regression approach." *Regional Science and Urban Economics*, Vol. **48**, (2014) pp. 133-143.

APPENDIX

Table A.1. Summary Statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
cond1	Unsound	31452	0.0153249	0.1228437	0	1
cond2	Poor	31452	0.0469922	0.2116256	0	1
cond3	Fair	31452	0.1375747	0.3444585	0	1
cond4	Average	31452	0.243196	0.4290193	0	1
cond5	Good	31452	0.2386176	0.4262452	0	1
cond6	Very good	31452	0.1715312	0.3769785	0	1
cond7	Excellent	31452	0.1467633	0.3538754	0	1
sd1	Carrolton Farmers branch ISD	31452	0.0257853	0.1584968	0	1
sd2	Dallas ISD	31452	0.337403	0.4728312	0	1
sd3	Cedar Hill ISD	31452	0.0574208	0.2326487	0	1
sd4	Garland ISD	31452	0.1581457	0.3648834	0	1
sd5	Highland Park ISD	31452	0.0017487	0.0417815	0	1
sd6	Irvin ISD	31452	0.0411421	0.1986219	0	1
sd7	Lancaster ISD	31452	0.0440989	0.2053182	0	1
sd8	Mesquite ISD	31452	0.1142694	0.3181432	0	1
sd9	Coppell ISD	31452	0.0077896	0.087916	0	1
sd10	Grand Prairi ISD	31452	0.0601552	0.2377779	0	1
sd11	Richardson ISD	31452	0.0427	0.2021831	0	1
sd12	Desoto ISD	31452	0.0606321	0.238658	0	1
sd13	Duncanville	31452	0.0474692	0.2126435	0	1
y06	House foreclosed in 2006	31452	0.2401755	0.4271967	0	1
y07	House foreclosed in 2007	31452	0.20978	0.407158	0	1
y08	House foreclosed in 2008	31452	0.1883505	0.390998	0	1
y09	House foreclosed in 2009	31452	0.124698	0.3303814	0	1
y10	House foreclosed in 2010	31452	0.0347514	0.1831523	0	1

Table A.2. Time to Resolution Estimates

Time to Resolution OLS			
Variable	**Coefficient**	**95% Confidence Interval**	
Constant	22.87205***	13.58362	32.16048
	(4.73883)		
Attached Garage	-0.195586	-0.8001651	0.4089931
	(0.3084479)		
Attached Carport	0.1878966	-0.8068828	1.182676
	(0.5075227)		
Detatched Carport	0.2415704	-1.25173	1.734871
	(0.7618613)		
Pool	-0.2184126	-0.935446	0.4986208
	(0.3658205)		
Slab Foundation	0.8380121**	0.1942272	1.481797
	(0.3284502)		
Central Heat	1.08114*	-0.1589551	2.321235
	(0.6326794)		
Central Air Conditioning	-0.9177252	-2.088417	0.2529668
	(0.5972709)		
Baths	0.4033117	-0.1059821	0.9126055
	(0.2598347)		
Lot Area	-0.0133282	-0.0351357	0.0084794
	(0.0111259)		
Living Area	0.5028554**	0.046044	0.9596668
	(0.2330589)		
Fireplace	0.9281334***	0.4149545	1.441312
	(0.2618168)		
Condition 7-	0.2541659	-1.462183	1.970515
	(0.8756577)		
Condition 6-	0.053073	-1.63739	1.743536
	(0.8624511)		
Condition 5-	0.3019233	-1.367375	1.971222
	(0.8516532)		
Condition 4-	-0.3711085	-2.029778	1.28756
	(0.84623)		
Condition 3-	-1.382133	-3.061999	0.2977341
	(0.8570448)		
Condition 2-	-0.687402	-2.504818	1.130014
	(0.9272208)		

Table A.2. Time to Resolution Estimates (Continued)

Time to Resolution OLS			
Variable	**Coefficient**	**95% Confidence Interval**	
School District 1	1.514121 (2.667591)	-3.714541	6.742783
School District 2	2.656278 (2.61318)	-2.465733	7.77829
School District 3	2.758352 (2.625281)	-2.387379	7.904083
School District 4	1.617617 (2.610638)	-3.499414	6.734647
School District 5	9.179798** (3.644931)	2.035482	16.32411
School District 6	2.994477 (2.647946)	-2.195679	8.184633
School District 7	3.161479 (2.64884)	-2.030428	8.353387
School District 8	1.569744 (2.61787)	-3.561461	6.700948
School District 9	1.152906 (2.819265)	-4.373047	6.678859
School District 10	1.139234 (2.63096)	-4.017629	6.296096
School District 11	1.903445 (2.644834)	-3.280611	7.087501
School District 12	2.878851 (2.627361)	-2.270956	8.028659
School District 13	2.280766 (2.63762)	-2.889151	7.450683
1 Story	-10.22056*** (3.719653)	-17.51133	-2.929782
1.5 Story	-10.04148*** (3.717873)	-17.32876	-2.754192
2 Story	-9.723288*** (3.707461)	-16.99017	-2.456411
Effective Age	0.0385671*** (0.0086951)	0.0215241	0.0556101

Table A.2. Time to Resolution Estimates (Continued)

Time to Resolution OLS			
Variable	**Coefficient**	**95% Confidence Interval**	
Community Percentage of School Aged Children	-2.669003*	-5.741228	0.4032218
	(1.567407)		
Community Percentage Over 60	-0.8624788	-4.141775	2.416817
	(1.673052)		
Community Percentage Below Poverty	1.094134	-0.8716297	3.059898
	(1.002906)		
Community Percentage White	2.043418***	0.9669228	3.119913
	(0.5492129)		
Year- 2006	-4.238269***	-4.866779	-3.60976
	(0.320657)		
Year- 2007	-9.669744***	-10.29048	-9.049008
	(0.3166909)		
Year- 2008	-12.93591***	-13.55943	-12.31239
	(0.3181111)		
Year- 2009	-14.40998***	-15.11611	-13.70385
	(0.3602562)		
Year- 2010	-14.32407***	-15.46611	-13.18203
	(0.5826547)		
Season- Winter	1.031417***	0.5882003	1.474633
	(0.2261228)		
Number of Observations	22698		
R-squared	0.1189		
Adj R-squared	0.1172		
F-Statistic	69.47		

Table A.3. Cox Proportional Hazard Estimates and Ratios

Cox Proportional Hazard Model				
Variable	**Coefficient**	**Hazard Ratio**	**95% Confidence Interval**	
Attached Garage	0.092386***	1.096788	1.051254	1.144295
	(0.0237284)			
Attached Carport	-0.070677**	0.9317628	0.8693409	0.9986668
	(0.0329655)			
Detatched Carport	-0.3062342***	0.7362142	0.6625904	0.8180187
	(0.0395776)			
Pool	0.0667578***	1.069037	1.016555	1.124227
	(0.0274563)			
Slab Foundation	-0.0974065***	0.9071872	0.8672384	0.9489761
	(0.0208448)			
Central Heat	0.1728756***	1.188718	1.089382	1.297112
	(0.0529261)			
Central Air Conditioning	0.193649***	1.21367	1.117157	1.318522
	(0.0513108)			
Baths	-0.0192058	0.9809774	0.9473483	1.0158
	(0.017459)			
Lot Area	-0.0019085**	0.9980934	0.9965213	0.9996679
	(0.0008027)			
Living Area	0.1141445***	1.120914	1.087418	1.155442
	(0.0173508)			
Fireplace	0.0119975	1.01207	0.9761918	1.049266
	(0.0186378)			
Condition 7-	0.0521213	1.053504	0.9353784	1.186546
	(0.0639238)			
Condition 6-	0.1168896**	1.123995	1.000047	1.263306
	(0.0670064)			
Condition 5-	0.0186648	1.01884	0.9079143	1.143318
	(0.0599204)			
Condition 4-	0.0242434	1.02454	0.9135087	1.149066
	(0.0599606)			
Condition 3-	0.1932727***	1.213214	1.079948	1.362924
	(0.0720265)			
Condition 2-	0.142168**	1.15277	1.016323	1.307537
	(0.0740945)			

Table A.3. Cox Proportional Hazard Estimates and Ratios (Continued)

Cox Proportional Hazard Model				
Variable	**Coefficient**	**Hazard Ratio**	**95% Confidence Interval**	
School District 1	-0.0957789	0.9086649	0.634183	1.301946
	(0.1667332)			
School District 2	-0.1920187	0.8252914	0.5804401	1.17343
	(0.1481972)			
School District 3	-0.2526254	0.7767588	0.545424	1.106211
	(0.1401229)			
School District 4	-0.0995105	0.9052805	0.6369098	1.286733
	(0.1624069)			
School District 5	-0.8163302***	0.4420509	0.2700977	0.7234753
	(0.1111105)			
School District 6	-0.1958596	0.8221276	0.5754886	1.174469
	(0.1496116)			
School District 7	-0.3790782**	0.6844921	0.4790344	0.9780705
	(0.1246444)			
School District 8	-0.1834367	0.8324046	0.585058	1.184323
	(0.1497539)			
School District 9	-0.2646126	0.7675032	0.5239813	1.124203
	(0.1494648)			
School District 10	-0.1898964	0.8270448	0.5801497	1.179011
	(0.1496188)			
School District 11	-0.2943624	0.7450065	0.5216526	1.063993
	(0.1354687)			
School District 12	-0.2247423	0.798722	0.5606764	1.137834
	(0.1442082)			
School District 13	-0.2733504	0.7608261	0.5332649	1.085495
	(0.1379553)			
1 Story	1.013217***	2.754447	1.678482	4.520141
	(0.6961101)			
1.5 Story	0.9126388***	2.490887	1.517327	4.08911
	(0.6299626)			
2 Story	0.8952382***	2.447919	1.493019	4.01355
	(0.6175339)			
Effective Age	-0.0029136***	0.9970907	0.9958647	0.9983182
	(0.0006259)			

Table A.3. Cox Proportional Hazard Estimates and Ratios (Continued)

Cox Proportional Hazard Model				
Variable	**Coefficient**	**Hazard Ratio**	**95% Confidence Interval**	
Neighborhood Percentage of School Aged Children	0.4175565***	1.518247	1.225541	1.880862
	(0.1659056)			
Neighborhood Percentage Over 60	-0.1956522*	0.8222982	0.6518004	1.037395
	(0.097488)			
Neighborhood Percentage Below Poverty	-0.4019829***	0.6689922	0.5830795	0.7675635
	(0.0469153)			
Neighborhood Percentage White	0.0405178	1.04135	0.9662476	1.122289
	(0.0397702)			
Year- 2006	0.1663004***	1.180928	1.130167	1.233968
	(0.0264717)			
Year- 2007	0.7335215***	2.082401	1.993357	2.175422
	(0.0464312)			
Year- 2008	0.9385969***	2.556392	2.444385	2.673532
	(0.0584374)			
Year- 2009	0.815602***	2.260536	2.147193	2.379863
	(0.0593295)			
Year- 2010	0.7640252***	2.146901	1.978423	2.329725
	(0.0895199)			
Season- Winter	0.0264968*	1.026851	0.9954982	1.059191
	(0.0162459)			
Number of Observations	30581			
Number of Failures	21828			
Log-Likelihood	-212569.35			

Table A.4. Kaplan-Meier Survival Statistics

Kaplan-Meier Survivor Function							
Time	**Beg. Total**	**Fail**	**Net Lost**	**Survivor Function**	**Std. Error**	**[95% Conf. Int.]**	
1	30582	1474	0	0.9518	0.0012	0.9493	0.9541
2	29108	2007	0	0.8862	0.0018	0.8826	0.8897
3	27101	2252	0	0.8125	0.0022	0.8081	0.8169
4	24849	2315	0	0.7368	0.0025	0.7319	0.7417
5	22534	2049	0	0.6698	0.0027	0.6645	0.6751
6	20485	1700	0	0.6143	0.0028	0.6088	0.6197
7	18785	1432	0	0.5674	0.0028	0.5619	0.5730
8	17353	1130	0	0.5305	0.0029	0.5249	0.5361
9	16223	895	0	0.5012	0.0029	0.4956	0.5068
10	15328	700	0	0.4783	0.0029	0.4727	0.4839
11	14628	571	0	0.4596	0.0028	0.4541	0.4652
12	14057	441	0	0.4452	0.0028	0.4397	0.4508
13	13616	367	0	0.4332	0.0028	0.4277	0.4388
14	13249	298	0	0.4235	0.0028	0.4179	0.4290
15	12951	265	0	0.4148	0.0028	0.4093	0.4203
16	12686	221	0	0.4076	0.0028	0.4021	0.4131
17	12465	204	0	0.4009	0.0028	0.3954	0.4064
18	12261	158	0	0.3958	0.0028	0.3903	0.4012
19	12103	138	0	0.3912	0.0028	0.3858	0.3967
20	11965	128	0	0.3871	0.0028	0.3816	0.3925
21	11837	101	0	0.3838	0.0028	0.3783	0.3892
22	11736	86	0	0.3809	0.0028	0.3755	0.3864
23	11650	101	0	0.3776	0.0028	0.3722	0.3831
24	11549	91	0	0.3747	0.0028	0.3692	0.3801
25	11458	102	0	0.3713	0.0028	0.3659	0.3767
26	11356	90	0	0.3684	0.0028	0.3630	0.3738
27	11266	88	0	0.3655	0.0028	0.3601	0.3709
28	11178	77	88	0.363	0.0027	0.3576	0.3684
29	11013	76	109	0.3605	0.0027	0.3551	0.3659
30	10828	76	93	0.358	0.0027	0.3526	0.3633
31	10659	73	114	0.3555	0.0027	0.3501	0.3609
32	10472	61	75	0.3534	0.0027	0.3481	0.3588
33	10336	57	90	0.3515	0.0027	0.3461	0.3568
34	10189	51	181	0.3497	0.0027	0.3444	0.3551
35	9957	54	21	0.3478	0.0027	0.3425	0.3532
36	9882	56	137	0.3459	0.0027	0.3405	0.3512
37	9689	62	80	0.3436	0.0027	0.3383	0.3490
38	9547	63	12	0.3414	0.0027	0.3361	0.3467
39	9472	57	94	0.3393	0.0027	0.3340	0.3446
40	9321	57	111	0.3372	0.0027	0.3319	0.3426
41	9153	56	19	0.3352	0.0027	0.3299	0.3405
42	9078	56	79	0.3331	0.0027	0.3278	0.3384
43	8943	56	101	0.331	0.0027	0.3257	0.3363
44	8786	51	112	0.3291	0.0027	0.3238	0.3344
45	8623	52	61	0.3271	0.0027	0.3218	0.3324

Table A.4. Kaplan-Meier Survival Statistics (Continued)

Kaplan-Meier Survivor Function							
Time	**Beg. Total**	**Fail**	**Net Lost**	**Survivor Function**	**Std. Error**	**[95% Conf. Int.]**	
46	8510	34	49	0.3258	0.0027	0.3205	0.3311
47	8427	41	85	0.3242	0.0027	0.3190	0.3295
48	8301	43	53	0.3226	0.0027	0.3173	0.3278
49	8205	44	91	0.3208	0.0027	0.3156	0.3261
50	8070	45	126	0.319	0.0027	0.3138	0.3243
51	7899	35	118	0.3176	0.0027	0.3124	0.3229
52	7746	37	61	0.3161	0.0027	0.3108	0.3214
53	7648	42	111	0.3144	0.0027	0.3091	0.3196
54	7495	48	105	0.3124	0.0027	0.3071	0.3176
55	7342	31	102	0.311	0.0027	0.3058	0.3163
56	7209	45	78	0.3091	0.0027	0.3038	0.3144
57	7086	48	72	0.307	0.0027	0.3018	0.3123
58	6966	33	151	0.3055	0.0027	0.3003	0.3108
59	6782	45	85	0.3035	0.0027	0.2983	0.3088
60	6652	46	103	0.3014	0.0027	0.2962	0.3067
61	6503	45	82	0.2993	0.0027	0.2941	0.3046
62	6376	33	176	0.2978	0.0027	0.2925	0.3030
63	6167	28	108	0.2964	0.0027	0.2912	0.3017
64	6031	46	221	0.2942	0.0027	0.2889	0.2994
65	5764	33	41	0.2925	0.0027	0.2872	0.2977
66	5690	26	55	0.2912	0.0027	0.2859	0.2964
67	5609	37	99	0.2892	0.0027	0.2840	0.2945
68	5473	28	88	0.2878	0.0027	0.2825	0.2930
69	5357	40	12	0.2856	0.0027	0.2804	0.2909
70	5305	29	50	0.284	0.0027	0.2788	0.2893
71	5226	37	62	0.282	0.0027	0.2768	0.2873
72	5127	27	158	0.2805	0.0027	0.2753	0.2858
73	4942	34	491	0.2786	0.0027	0.2734	0.2839
74	4417	25	285	0.277	0.0027	0.2718	0.2823
75	4107	32	245	0.2749	0.0027	0.2696	0.2802
76	3830	24	197	0.2732	0.0027	0.2679	0.2785
77	3609	27	344	0.2711	0.0027	0.2658	0.2764
78	3238	31	605	0.2685	0.0027	0.2632	0.2739
79	2602	22	23	0.2662	0.0027	0.2609	0.2716
80	2557	18	273	0.2644	0.0028	0.2590	0.2698
81	2266	17	321	0.2624	0.0028	0.2570	0.2679
82	1928	13	2	0.2606	0.0028	0.2551	0.2661
83	1913	9	294	0.2594	0.0028	0.2539	0.2649
84	1610	12	214	0.2575	0.0029	0.2519	0.2631
85	1384	14	219	0.2549	0.0029	0.2492	0.2606
86	1151	10	329	0.2526	0.003	0.2468	0.2585
87	812	12	228	0.2489	0.0031	0.2428	0.2550
88	572	7	275	0.2459	0.0033	0.2395	0.2523
89	290	1	282	0.245	0.0034	0.2384	0.2517
90	7	0	7	0.245	0.0034	0.2384	0.2517

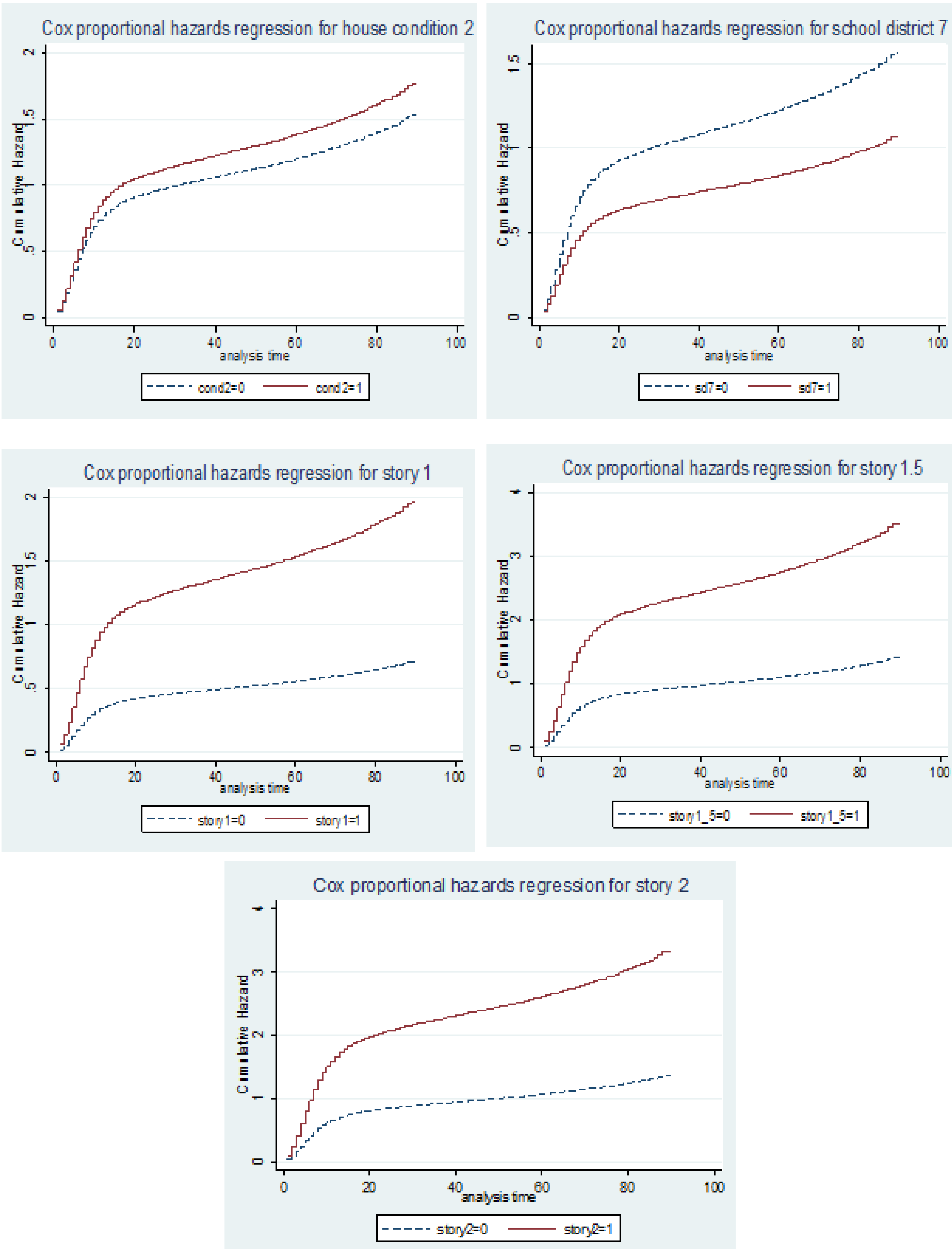

Figure A.1. Hazard Graphs

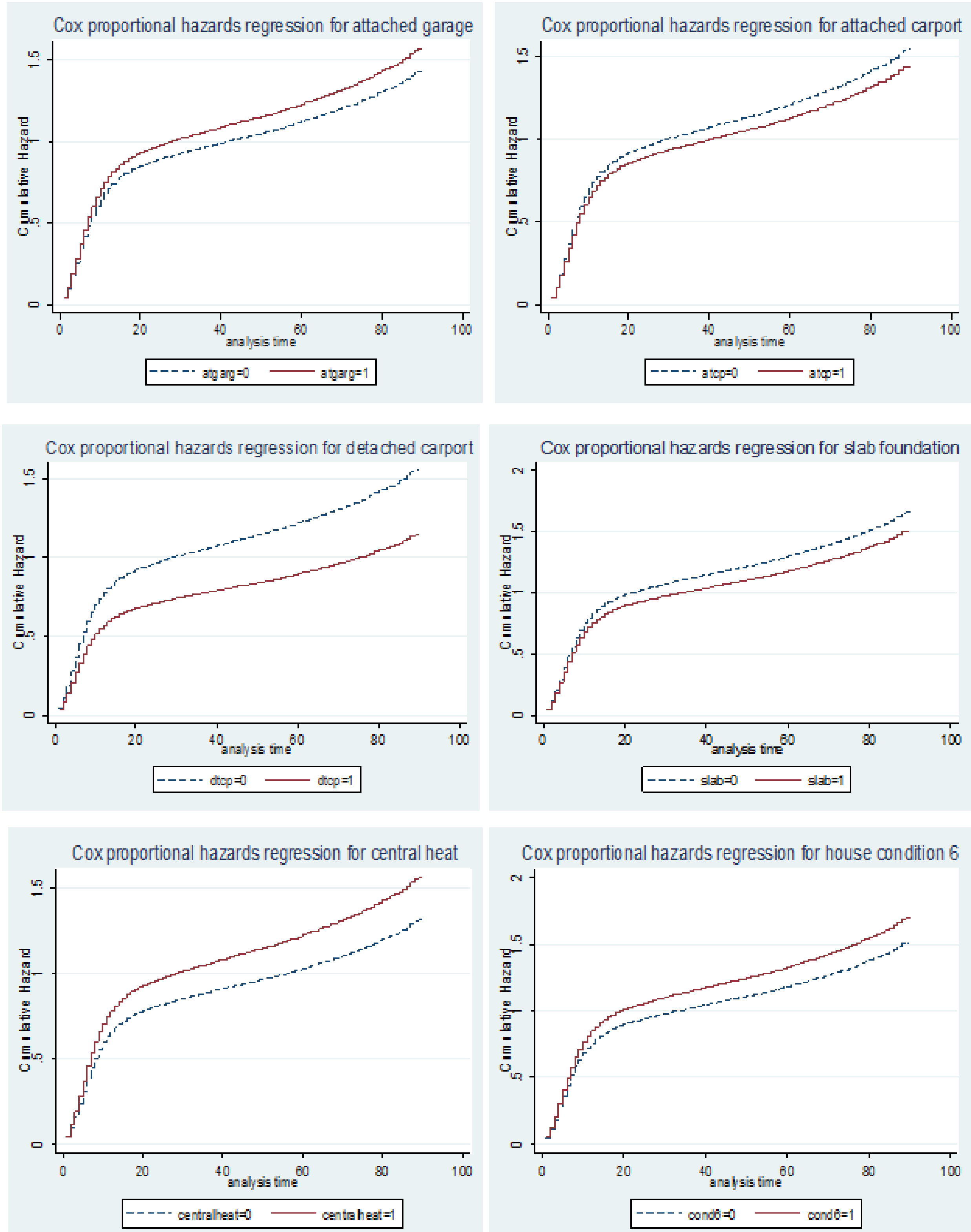

Figure A.1. Hazard Graphs (Continued)

Table A.5. Hedonic Model Estimates

Natural Log of Sales Price OLS			
Variable	**Coefficient**	**95% Confidence Interval**	
Constant	10.2362***	10.02643	10.44596
	(0.1070189)		
Attached Garage	0.0344818***	0.0208283	0.0481353
	(0.0069658)		
Attached Carport	-0.010945	-0.0334105	0.0115205
	(0.0114616)		
Detatched Carport	0.0261764	-0.0075474	0.0599002
	(0.0172054)		
Pool	0.0667078***	0.0505147	0.0829008
	(0.0082615)		
Slab Foundation	-0.1041215***	-0.1186603	-0.0895826
	(0.0074175)		
Central Heat	0.0751414***	0.0471359	0.103147
	(0.0142881)		
Central Air Conditioning	0.1696838***	0.1432456	0.196122
	(0.0134884)		
Baths	0.0636866***	0.052185	0.0751882
	(0.005868)		
Lot Area	0.0003434	-0.000149	0.0008359
	(0.0002513)		
Living Area	0.3315146***	0.3211982	0.341831
	(0.0052633)		
Fireplace	0.1258583***	0.114269	0.1374476
	(0.0059127)		
Condition 7-	0.3806472***	0.3418862	0.4194082
	(0.0197753)		
Condition 6-	0.3609221***	0.3227457	0.3990986
	(0.0194771)		
Condition 5-	0.3020336***	0.2643351	0.339732
	(0.0192332)		
Condition 4-	0.2482623***	0.2108039	0.2857207
	(0.0191108)		
Condition 3-	0.2008731***	0.162936	0.2388102
	(0.019355)		
Condition 2-	0.127279***	0.0862355	0.1683224
	(0.0209398)		

Table A.5. Hedonic Model Estimates (Continued)

Natural Log of Sales Price OLS			
Variable	**Coefficient**	**95% Confidence Interval**	
School District 1	0.1527971**	0.0347161	0.2708781
	(0.0602433)		
School District 2	0.0190371	-0.0966354	0.1347095
	(0.0590145)		
School District 3	-0.0841696	-0.2003777	0.0320385
	(0.0592878)		
School District 4	-0.0423482	-0.1579082	0.0732117
	(0.0589571)		
School District 5	0.7812978***	0.6199548	0.9426407
	(0.082315)		
School District 6	0.1119157*	-0.0052957	0.2291271
	(0.0597996)		
School District 7	-0.1670927***	-0.2843436	-0.0498418
	(0.0598198)		
School District 8	-0.1613277***	-0.2772078	-0.0454477
	(0.0591204)		
School District 9	0.1776445***	0.0528497	0.3024393
	(0.0636686)		
School District 10	-0.0625453	-0.1790047	0.0539142
	(0.059416)		
School District 11	0.101973*	-0.0151006	0.2190467
	(0.0597293)		
School District 12	-0.069103	-0.1854032	0.0471972
	(0.0593347)		
School District 13	-0.0655489	-0.1823032	0.0512054
	(0.0595664)		
1 Story	0.0334408	-0.1312097	0.1980913
	(0.0840024)		
1.5 Story	-0.0035174	-0.1680891	0.1610544
	(0.0839622)		
2 Story	-0.0157234	-0.1798342	0.1483874
	(0.0837271)		
Effective Age	-0.0045981***	-0.0049829	-0.0042132
	(0.0001964)		

Table A.5. Hedonic Model Estimates (Continued)

Natural Log of Sales Price OLS			
Variable	**Coefficient**	**95% Confidence Interval**	
Neighborhood Percentage of School Aged Children	0.2073673***	0.137986	0.2767486
	(0.0353974)		
Neighborhood Percentage Over 60	-0.7615106***	-0.8355682	-0.6874529
	(0.0377832)		
Neighborhood Percentage Below Poverty	-0.526034***	-0.5704277	-0.4816404
	(0.022649)		
Neighborhood Percentage White	0.6305285***	0.6062176	0.6548395
	(0.0124031)		
Year- 2006	-0.0056909	-0.0198848	0.0085029
	(0.0072415)		
Year- 2007	-0.055953***	-0.0699713	-0.0419347
	(0.007152)		
Year- 2008	-0.2103005***	-0.2243817	-0.1962193
	(0.007184)		
Year- 2009	-0.209571***	-0.2255177	-0.1936242
	(0.0081358)		
Year- 2010	-0.2263275***	-0.2521187	-0.2005363
	(0.0131583)		
Season- Winter	0.0086093*	-0.0014001	0.0186186
	(0.0051066)		
Number of observations	22698		
R-squared	0.7391		
Adj R-squared	0.7386		
F Statistic	1458.65		

Table A.6. Correlation Matrix

	atgarg	atcp	pool	dtcp	slab	centra~t	centra~c	baths	lotarea
atgarg	1								
atcp	-0.2811	1							
pool	0.0421	-0.0036	1						
dtcp	-0.2233	-0.0272	-0.0102	1					
slab	0.4283	-0.102	0.0567	-0.0967	1				
centralheat	0.3378	-0.0733	0.0787	-0.0548	0.4022	1			
centralac	0.3471	-0.0772	0.0877	-0.0488	0.4219	0.7997	1		
baths	0.3224	-0.0949	0.2476	-0.036	0.3249	0.3537	0.3845	1	
lotarea	-0.0309	0.0262	0.127	0.021	-0.0871	-0.0267	-0.0181	0.1716	1
livarea	0.2324	-0.0748	0.2624	-0.0296	0.1981	0.2452	0.27	0.8016	0.2501
fireplace	0.3089	-0.1012	0.1705	-0.0411	0.3664	0.3082	0.3529	0.4767	0.0527
cond7	0.1634	-0.0626	-0.027	-0.047	0.172	0.1136	0.1265	0.2231	0.019
cond6	0.0852	-0.0306	0.0176	-0.0144	0.0939	0.0986	0.1092	0.0769	-0.0058
cond5	0.0471	-0.0111	0.0252	-0.012	0.0524	0.0804	0.088	0.0355	-0.0145
cond4	-0.1068	0.041	-0.0081	0.038	-0.1331	-0.1118	-0.127	-0.1505	-0.0012
cond3	-0.1126	0.0342	-0.0028	0.0305	-0.1079	-0.1002	-0.1128	-0.1021	-0.0012
cond2	-0.0819	0.0254	-0.0055	-0.0017	-0.0845	-0.0826	-0.0878	-0.082	0.0003
sd1	0.0168	-0.0038	0.0285	0.0083	0.0253	0.0318	0.0268	0.0404	-0.0155
sd2	-0.2674	0.057	-0.0942	0.0617	-0.4399	-0.2431	-0.2911	-0.2376	-0.005
sd3	0.114	-0.0502	-0.0248	-0.0306	0.1214	0.071	0.0804	0.1394	0.0768
sd4	0.0797	-0.0107	0.0114	-0.0372	0.1378	0.0577	0.0895	0.0207	-0.0291
sd5	-0.0295	-0.0057	0.0212	0.0144	-0.0457	0.0128	0.0139	0.1143	-0.0002
sd6	-0.0361	0.0344	0.0276	0.0277	0.0036	0.0393	0.0461	-0.0119	-0.0069
sd7	0.0635	-0.0205	-0.0323	-0.0059	0.0619	0.0291	0.0375	0.0093	0.015
sd8	0.0152	0.0259	-0.0167	-0.0213	0.1245	0.0447	0.051	-0.0524	-0.0573
sd9	0.0398	-0.0162	0.0815	-0.0108	0.0483	0.0271	0.0295	0.0691	-0.004
sd10	0.0247	-0.0092	-0.0174	0.0102	0.0682	0.0056	0.0147	-0.0137	-0.0372
sd11	0.0379	-0.0156	0.1624	0.01	0.0112	0.0606	0.0541	0.1208	0.0026
sd12	0.1156	-0.0367	0.0248	-0.0336	0.1266	0.0681	0.0771	0.1605	0.0635

Table A.6. Correlation Matrix (Continued)

	livarea	fireplace	cond7	cond6	cond5
livarea	1				
fireplace	0.4631	1			
cond7	0.2284	0.1072	1		
cond6	0.059	0.1144	-0.1887	1	
cond5	0.0134	0.0727	-0.2322	-0.2547	1
cond4	-0.135	-0.1357	-0.2351	-0.2579	-0.3173
cond3	-0.0962	-0.0844	-0.1657	-0.1817	-0.2236
cond2	-0.0654	-0.081	-0.0921	-0.101	-0.1243
sd1	0.0191	0.0011	-0.0233	-0.0011	0.0209
sd2	-0.2217	-0.3363	-0.0385	-0.0864	-0.0857
sd3	0.1509	0.1273	0.0328	0.0516	0.0237
sd4	0.0057	0.0969	0.0025	0.0104	0.0039
sd5	0.1167	0.0294	0.0257	-0.0069	-0.0056
sd6	-0.0289	-0.0633	-0.0429	0.0004	0.0132
sd7	0.023	0.0806	0.0895	0.0062	0.0051
sd8	-0.0717	0.0248	-0.0241	0.0184	0.0245
sd9	0.0621	0.0642	-0.0122	0.0412	0.0234
sd10	-0.0051	0.0081	-0.012	0.0371	0.0268
sd11	0.0972	0.0734	-0.0569	-0.0072	0.0341
sd12	0.2082	0.1688	0.0765	0.0057	0.0019
sd13	0.0361	0.0817	0.0308	0.0261	0.016
story1	-0.5616	-0.266	-0.19	-0.0522	-0.0002
story1_5	0.2517	0.1731	0.0187	0.0388	0.0163
story2	0.4709	0.1801	0.212	0.0319	-0.0124
eff_age	-0.4373	-0.4577	-0.4082	-0.1436	-0.012
school_age	-0.0863	-0.0401	0.0479	0.0215	-0.0111
commute_gt60	-0.0926	-0.0514	0.0313	0.0037	-0.0411
below_poverty	-0.3587	-0.3913	-0.0522	-0.0914	-0.0703
white	0.2972	0.2638	-0.0254	0.0496	0.0515
y06	0.0142	0.0426	-0.0473	-0.0122	-0.0123
y07	0.0161	-0.0022	-0.0533	-0.083	-0.0611
y08	-0.005	-0.0307	0.0737	0.0311	0.0165
y09	-0.0164	-0.028	0.0588	0.0322	0.024
y10	-0.0098	-0.0165	0.0292	0.0182	0.0086
winter	-0.0049	-0.018	0.0097	0.0096	-0.0015

Table A.6. Correlation Matrix (Continued)

	atgarg	atcp	pool	dtcp	slab	centralheat	centralac	baths	lotarea
sd13	0.0635	-0.0275	0.0224	-0.0152	0.0981	0.0527	0.0589	0.051	0.0184
story1	-0.1536	0.0592	-0.0754	-0.0029	-0.1839	-0.1409	-0.1579	-0.5143	-0.0421
story1_5	0.0596	-0.0153	0.0834	0.0255	0.0728	0.0774	0.0862	0.2604	0.0407
story2	0.1384	-0.0592	0.0219	-0.0181	0.164	0.1063	0.1195	0.4023	0.018
eff_age	-0.5262	0.1736	-0.0225	0.1175	-0.6436	-0.4343	-0.4589	-0.5078	0.0409
school_age	0.0664	-0.0221	-0.0994	-0.0217	0.1356	0.0287	0.033	-0.0549	-0.049
commute_gt60	0.0211	0.0016	-0.0916	-0.0171	0.0438	-0.0278	-0.0416	-0.0965	-0.0119
below_poverty	-0.2323	0.0664	-0.1427	0.0545	-0.2285	-0.252	-0.2631	-0.3212	-0.0518
white	0.0849	-0.0124	0.237	-0.0026	0.0927	0.187	0.2072	0.2975	0.0897
y06	-0.0044	0.0018	0.0073	-0.0009	0.003	0.0142	0.0146	0.0117	0.0131
y07	0.0078	0.0052	-0.0112	-0.0046	0.0161	0.0075	0.0108	0.0089	-0.0067
y08	0.0063	-0.0028	-0.0086	0.0057	0.005	-0.025	-0.0263	-0.0088	0.0002
y09	0.0064	-0.0053	-0.007	0.0082	0.0004	-0.0062	-0.0116	-0.0104	-0.0036
y10	0.0005	0.0089	0.0057	-0.0025	0.0045	-0.0041	-0.0004	-0.0046	-0.0098
winter	-0.0032	-0.0085	0.007	-0.0015	-0.0109	-0.0028	-0.0081	-0.0059	0.0023

Table A.6. Correlation Matrix (Continued)

	cond4	cond3	cond2	sd1	sd2	sd3	sd4	sd5	sd6
cond4	1								
cond3	-0.2264	1							
cond2	-0.1259	-0.0887	1						
sd1	0.013	-0.0114	-0.002	1					
sd2	0.0883	0.064	0.0687	-0.1161	1				
sd3	-0.0392	-0.0379	-0.0406	-0.0402	-0.1761	1			
sd4	0.0009	-0.0031	-0.0139	-0.0705	-0.3093	-0.107	1		
sd5	-0.0042	-0.0035	-0.0021	-0.0068	-0.0299	-0.0103	-0.0181	1	
sd6	0.0214	0.0018	0.0039	-0.0337	-0.1478	-0.0511	-0.0898	-0.0087	1
sd7	-0.0503	-0.0166	-0.0353	-0.0349	-0.1533	-0.053	-0.0931	-0.009	-0.0445
sd8	-0.0252	0.0158	-0.007	-0.0584	-0.2563	-0.0887	-0.1557	-0.015	-0.0744
sd9	-0.0207	-0.0249	-0.0128	-0.0144	-0.0632	-0.0219	-0.0384	-0.0037	-0.0184
sd10	-0.0159	-0.0242	-0.0132	-0.0412	-0.1805	-0.0624	-0.1097	-0.0106	-0.0524
sd11	0.017	0.0024	0.0103	-0.0344	-0.1507	-0.0521	-0.0915	-0.0088	-0.0437
sd12	-0.0539	-0.0326	-0.0004	-0.0413	-0.1813	-0.0627	-0.1101	-0.0106	-0.0526
sd13	-0.0248	-0.0227	-0.0298	-0.0363	-0.1593	-0.0551	-0.0968	-0.0093	-0.0462
story1	0.1055	0.0837	0.046	-0.035	0.1546	-0.1268	-0.0146	-0.0593	0.0285
story1_5	-0.0354	-0.021	-0.0236	0.0247	-0.0786	0.0281	0.0275	0.0016	-0.0117
story2	-0.098	-0.084	-0.0358	0.0224	-0.1262	0.1328	-0.0048	0.0632	-0.0245
eff_age	0.2423	0.1843	0.1296	0.0243	0.3401	-0.2068	-0.07	-0.003	0.0901
school_age	-0.0352	-0.0102	-0.0117	-0.0257	-0.1132	0.0648	0.0153	0.0251	-0.0246
commute_gt60	0.001	-0.0018	0.0092	-0.1445	0.055	0.0824	0.0473	-0.0464	-0.1291
below_poverty	0.0837	0.0626	0.0759	-0.0749	0.4268	-0.1147	-0.1308	-0.0315	0.0056
white	-0.0212	-0.0228	-0.0367	0.0746	-0.3125	-0.0315	0.2271	0.1021	0.0463
y06	0.0023	0.0665	0.0144	0.001	-0.0056	0.0062	-0.0001	-0.0128	-0.0055
y07	-0.0259	0.1259	0.1479	-0.0055	-0.0118	0.0101	0.0085	-0.0085	-0.0021
y08	-0.0021	-0.0833	-0.0562	-0.004	0.0035	0.0056	-0.0082	-0.0124	0.0005
y09	0.0061	-0.0834	-0.0584	0.0017	0.0073	-0.0162	-0.0043	0.0072	0.0061
y10	0.0005	-0.037	-0.0307	0.0053	-0.0058	-0.0125	0.0024	0.0128	0.0201
winter	0.0084	-0.0199	-0.0097	0.0051	-0.0006	-0.0029	-0.0037	-0.0016	0.016

Table A.6. Correlation Matrix (Continued)

	sd7	sd8	sd9	sd10	sd11	sd12	sd13	story1	story1_5
sd7	1								
sd8	-0.0772	1							
sd9	-0.019	-0.0318	1						
sd10	-0.0543	-0.0909	-0.0224	1					
sd11	-0.0454	-0.0759	-0.0187	-0.0534	1				
sd12	-0.0546	-0.0912	-0.0225	-0.0643	-0.0536	1			
sd13	-0.0479	-0.0802	-0.0198	-0.0565	-0.0471	-0.0567	1		
story1	0.013	-0.0113	-0.0548	-0.014	-0.0036	-0.1018	-0.0026	1	
story1_5	-0.0374	0.0094	0.0523	0.0084	0.0473	0.0266	0.0004	-0.5937	1
story2	0.017	0.0069	0.0227	0.0108	-0.0359	0.1036	0.0036	-0.7164	-0.131
eff_age	-0.1304	-0.0557	-0.0511	-0.0472	0.0617	-0.1718	-0.0895	0.361	-0.1108
school_age	0.0281	0.0499	0.03	0.0724	-0.1254	0.0538	0.0566	0.0062	-0.0326
commute_gt60	-0.0659	0.0912	-0.0737	0.0102	-0.154	0.0741	-0.0466	0.0103	-0.0275
below_poverty	-0.0905	-0.0708	-0.0774	-0.0212	-0.075	-0.1445	-0.056	0.174	-0.0923
white	-0.1676	0.1745	0.1278	0.0137	0.2372	-0.1172	-0.0482	-0.1048	0.1081
y06	0.0043	-0.0085	0.0103	-0.0123	0.0141	0.0154	-0.002	-0.0143	0.0274
y07	0.0076	-0.0142	-0.0163	-0.0019	-0.0014	0.0249	0.0043	-0.004	0.0013
y08	0.009	0.0136	-0.0177	0.0071	-0.0237	-0.001	-0.0001	-0.0018	-0.02
y09	-0.0037	0.0069	0.0049	0.0239	-0.0045	-0.0241	-0.0023	0.0101	-0.0159
y10	-0.0036	-0.0027	0.0029	0.0111	-0.0023	-0.0096	0.005	0.0071	-0.0111
winter	-0.0014	-0.0126	-0.0068	0.0031	0.0032	0.0085	0.0008	0.003	-0.0028

Table A.6. Correlation Matrix (Continued)

	story2	eff_age	school~e	comm~t60	below_~y	white	y06	y07
story2	1							
eff_age	-0.3461	1						
school_age	0.024	-0.1548	1					
commute_gt60	0.0129	-0.0909	0.1807	1				
below_pove~y	-0.1363	0.2575	0.1912	0.1549	1			
white	0.0331	-0.0356	-0.2924	-0.1758	-0.4284	1		
y06	-0.0045	0.0018	-0.0094	-0.0019	-0.0223	0.0137	1	
y07	0.0038	-0.0154	0.0167	0.0067	0.0163	-0.019	-0.2897	1
y08	0.0208	-0.0173	0.0119	0.0093	0.0235	-0.0288	-0.2708	-0.2482
y09	-0.0015	-0.0029	-0.0023	0.0018	0.0063	0.0029	-0.2122	-0.1945
y10	-0.0001	0.0005	-0.0011	-0.019	0.0019	0.0067	-0.1067	-0.0978
winter	-0.0003	0.0085	0.0005	-0.0081	0.0066	-0.0061	0.1542	-0.1012

Table A.6 Correlation Matrix (Continued)

	y08	y09	y10	winter
y08	1			
y09	-0.1818	1		
y10	-0.0914	-0.0716	1	
winter	0.0268	-0.0262	0.194	1