Determinants of Foreclosure: A Chicago Case Study

By: Olga Malkova

Senior Honors Thesis
Summer 2008

Economics Department
University of California, Berkeley

Thesis Advisor: John M. Quigley

I would like to thank Professor Quigley for his time, inspiration and guidance. I would also like to thank Harrison Dekker for his help with data collection, and Paavo Monkkonen for his help with GIS.

ABSTRACT

This paper examines the determinants and the geographic distribution of foreclosure in the Chicago area at the census tract level. Subprime lending has a large positive effect on foreclosures, all else constant. Out of the demographic characteristics, the black share is the most important positive determinant of foreclosure. The share of college graduates, the vacancy rate and the median year built have a positive effect on foreclosures, while the share of population 65 and over has a negative effect. House price appreciation has a negative impact on foreclosures and a positive impact on subprime lending, all else equal.

INTRODUCTION

America's current foreclosure problem is a serious cause for concern. About 0.2 percent of all US households received a foreclosure filing notice in May 2008, which is the highest recorded monthly rate.[1] This rate represents a seven percent increase since April 2008, and an almost fifty percent increase in the past year. The rapid increase is partially due to increases in subprime mortgage foreclosures. The subprime mortgage market provides an opportunity to get a loan for high-risk borrowers with low credit scores who would otherwise be unqualified in the prime market. These loans typically have higher interest rates and involve penalties as well as upfront fees, which makes them more risky. About a quarter of adjustable rate subprime mortgages are now delinquent.[2] This is very troubling and calls into question the benefit of subprime lending. The initial purpose of the subprime sector was to give those groups that were previously unqualified for credit an opportunity for home-ownership. However, some critics argue that the growth of the subprime sector led to extremely loose underwriting standards and irresponsible lending, the effect of which can now be felt in the enormous increase in foreclosures.

As foreclosures keep rising, the national debate about the causes of this phenomenon escalates. Recently, Chairman Bernanke stated that changes in home values might be affecting foreclosures. States with the highest foreclosure rate increases such as California, Florida, Nevada, Michigan, and parts of Arizona and Colorado experienced decreases in house prices between 2006

[1] According to an article by Huffman (2008) , based on RealtyTrac data
[2] According to a speech by Bernanke (2008), based on servicer data from First American LoanPerformance

and 2007. Bernanke (2008) argues: "sharp declines in house prices, and thus in homeowners' equity, reduce both the ability and incentive of homeowners, particularly those under financial stress for other reasons, to retain their homes."

Much of this type of national debate on the causes of foreclosure happens at the aggregate level. However, it is impossible to infer the true causes of foreclosure from aggregate data, since it only presents very general trends. To gain more insight, it is crucial to examine characteristics associated with foreclosure at an individual or neighborhood level. Since it is a borrower's choice to default on a mortgage, the determinants of foreclosure are individual in nature. When individual data is lacking, neighborhood characteristics are reasonable to use in an analysis of determinants of foreclosure. The demographics of an area are fairly good indicators of individuals living in the area. Moreover, the demographics might also point to purely neighborhood effects on foreclosure: the type of neighborhood where an individual resides affects his probability of foreclosure.

This paper examines the neighborhood characteristics of foreclosure in six Chicago area counties in 2006. The analysis explores the relationship between foreclosure rates and the loan level, demographic, economic and housing characteristics of a neighborhood, focusing on the identification of factors that are closely associated with foreclosure. It is important to pinpoint neighborhood characteristics that can identify high foreclosure-risk neighborhoods in advance, as that would help in the design of foreclosure prevention and mitigation programs. Moreover, this paper attempts to elucidate the relationship between foreclosures, house price appreciation and subprime lending. This paper's methods and hypotheses are an extension to a large body of literature that has already explored the topic in a considerable amount of detail.

LITERATURE REVIEW

Determinants of Foreclosure

Many studies have examined the relationship between foreclosures and neighborhood characteristics, identifying factors that are associated with foreclosure.

A study on mortgage foreclosures in Atlanta by Duda et al. (2005) examines the spatial distribution of foreclosures; it also uncovers evidence of abusive lending practices. The authors use a list of initiation of foreclosure proceedings in 12 Atlanta counties from January 2000 through March 2005. Analysis indicates that foreclosures are concentrated in certain "hot spots" that are generally urban, low income and high minority. For instance, in urban neighborhoods whose median incomes were below $35,000 and had more than 80 percent minorities, there were 7.8 percent foreclosure starts on mortgaged homes annually. The foreclosure rate in similar suburban neighborhoods was 4.5 percent, which is also quite high. The authors also calculate the percent of quick foreclosures, foreclosures that happened the same year or the year after the loan origination. If the percent of quick foreclosures is high, it can be an indicator of "a breakdown in oversight and/or underwriting standards." Alarmingly, almost a fifth of foreclosure filings were quick foreclosures. Moreover, quick foreclosures were a larger share of overall foreclosure activity in areas with a high foreclosure rate, a high subprime rate, a high concentration of minorities and a high concentration of low-income households.

Mueller et al. (2006) examine the association between neighborhood characteristics and foreclosure rates in 6 counties in Texas. The authors use foreclosure-reporting data from 2004 to mid-2006 to identify a foreclosure as the actual loss of a home as the result of a legal process. The data include those properties that the mortgage holder acquired at the foreclosure sale. High concentrations of minorities correlated with higher foreclosure rates in all five of the six counties. Analysis indicated trends connecting tracts with higher foreclosure rates to lower income levels and higher volumes of subprime lending.

Grover et al. (2007) inspect variables correlated with foreclosure risk in Hennepin and Ramsey counties in Minneapolis and test how they would perform as predictors of high foreclosure risk neighborhoods in 2002. The credit score was the best predictor of foreclosure, as it identified 36 of the 50 tracts with the highest foreclosure rates. Other variables such as the prime

mortgage denial rate, the household income, the minority population share, and the subprime refinancing identified approximately two thirds of the top quintile of foreclosure rate neighborhoods. So, these variables seem to be good substitutes for credit risk variables to identify tracts that may experience high foreclosure. The authors use a minimum chi-square estimator and find that a high percentage of adults with low credit scores and an increased percentage of minority homeowners are the most associated with high foreclosure rates. The authors hypothesize that a rising economy in the late 90's in the Twin Cities enabled many minority households to become homeowners. However, when the economy slowed down in 2000, many of these minority homeowners experienced income declines and had too little home equity or other household wealth to maintain their mortgage payments. When the authors replace the credit risk variable with the denial rate, the coefficient on the minority share becomes significant. Since a variable that was not significant before became significant when the denial rate replaced the credit risk variable, the denial rate does not explain all the variation in foreclosure rates that the credit risk variable does. The significant coefficient on the minority share indicates that the denial rate may underestimate the credit risk in minority neighborhoods.

Lanzerotti (2006) examines how neighborhood characteristics are associated with foreclosures in three California counties (Alameda, Fresno and Riveside). She uses notices of default from January 2005 through February 2006 to proxy for foreclosure risk and 2004 high-cost loan data. Subprime loans and notices of default are more prevalent in neighborhoods with a higher concentration of blacks and Hispanics, a greater number of recent immigrants, a higher percentage of poverty, a lower percentage of college graduates, a lower median income and lower median housing values. In a negative binomial model controlling for demographic characteristics, Lanzerotti finds a positive relationship between subprime lending and foreclosure.

Subprime Lending and Foreclosure

Previous studies provide evidence that subprime lending and foreclosure are tightly connected, which makes it interesting to study their relationship more closely.

Immergluck et al. (2004) examine the quantitative relationship between subprime lending and foreclosure levels. The study concentrates on the five county Chicago area and includes Cook, DuPage, Lake, Kane and Will counties. The authors use initiations of foreclosure proceedings data from 1995 until 2002, as well as HMDA subprime lending data from 1996 until 2001. Results from an OLS model show, even while controlling for neighborhood demographics, that for every 100 additional subprime loans for owner occupied units made in 1996-2001, there were 9 additional foreclosure starts in a census tract in 2002. This is a pretty big increase, considering that on average there were 11 foreclosures per tract that year. In fact, subprime loans are about 20 times more likely to foreclose than prime loans. Prime lending virtually did not affect the foreclosure level for refinancing, while it reduced the level of expected home-purchase foreclosures. The effect of other neighborhood characteristics, even though sometimes significant, was too small. The authors conclude that subprime lending was the main factor that contributed to the increase of neighborhood foreclosure levels in 1996-2002. This study raises concerns about irresponsible lending, considering how foreclosure-prone subprime loans are.

Since subprime lending appears to be a major factor in driving foreclosure, it is useful to know how often subprime borrowers default. Gerardi et al. (2008) try to characterize sale and default probabilities across a particular ownership experience by using deeds records from January 1987 through August 2007 in Massachusetts. The registry of deeds is a unique dataset that provides information on all residential home sales and mortgage originations, which lets the authors track every mortgage issued on any property over the sample period. The authors use a competing risks, proportional hazard and duration model where they include explanatory variables such as the loan to value ratio, the town level house price appreciation, the unemployment rate, the

median household income, and the percent minority households. Their main finding is that "homeownerships financed by subprime mortgages are five to six times more likely to default than are homeownerships financed by prime mortgages, on average, at any point in the life of the ownership." For the entire 18-year old sample, within 12 years of purchasing a home, the cumulative probability of default for a subprime borrower is approximately 18 percent, compared to 3 percent for prime borrowers."

Schloemer et al. (2006) analyze the performance of more than 6 million subprime mortgages from 1998 through 2004 and project lifetime cumulative subprime foreclosure rates for each annual cohort. Their data covers the entire US and represents over 70 percent of the subprime market. They only analyze first lien, owner-occupied loans. The authors project that 19 percent of subprime mortgages originated during 2005-2006 will end in foreclosure, which is double of what they predict for subprime loans made in 2002. The authors project that 2.2 million borrowers will lose their homes to foreclosure from subprime mortgages originated from 1998 until 2006, which will cost them up to $164 billion of wealth. The authors also find that many features of subprime loans increase the risk of foreclosures by 65 to 123 percent depending on the year of origination, regardless of the borrower credit history. These features include adjustable interest rates, loans with prepayment or balloon payments and low documentation loans, where lenders approve loans based on little or no verification of income and assets. The authors conclude that the current loose underwriting standards and the fact that lenders benefit from granting as many loans as possible are actually very damaging to the borrowers.

House Price Appreciation and Foreclosure

A few studies have discovered that another important factor that predicts foreclosures at a local as well as at a more aggregate level is house price appreciation. Lanzerotti (2006) suggests that low foreclosure rates in California that persisted until the beginning of 2006 were due to the rapidly rising house prices. "It is the general consensus that rising home prices, low interest rates,

exotic mortgage products, enabled more households to avoid foreclosure." She predicted that as soon as house prices stabilized, the foreclosure rates would start going up because of the enormous subprime lending volume. Lanzerotti later observes that the number of foreclosures started going up in the beginning of 2006 and attributes it to the consequences of excessive subprime lending in the previous years. Even though her link between house price changes and subprime lending is only speculative, her predictions for California came true, as foreclosure rates in the state increased dramatically in 2006 and 2007.

Immergluck (2008) uses aggregate MSA level data to explore the relationship between foreclosures and house price appreciation. He suggests that the relationship between subprime lending and foreclosure depends on the strength of the local housing market. It is reasonable to expect lower foreclosure rates in appreciating housing markets, because owners with difficulties in paying off their loans can more easily sell or refinance into more affordable mortgages. Immergluck plots foreclosure rates against subprime lending rates for 100 metropolitan areas and labels each area as either coming from an appreciating housing market or a depreciating housing market. In areas with a depreciating housing market, high levels of subprime lending are associated with much higher foreclosure rates than in areas with an appreciating housing market. Moreover, metropolitan areas that experienced the most rapid appreciation until approximately 2006 had the highest foreclosure growth since. However, areas with modest levels of appreciation in recent years didn't experience a significant change in foreclosures. Even though these results are convincing, it is important to remember that they are at a very general aggregate level and research at a deeper level is needed to test the relationships found.

Gerardi et al. (2008) take advantage of the 20-year Massachusetts deeds records and town-level house appreciation data to explore the relationship between appreciation and foreclosure rates. The authors find that the probability of default for subprime borrowers, as well as prime borrowers, increases significantly in periods with low or negative house price appreciation. A one

standard deviation increase in house price appreciation increases the default hazard by almost 240 percent, all else constant. In fact, the authors attribute the rise in foreclosures in Massachusetts in 2006 and 2007 to the decline in house prices that began in the summer of 2005.

Schloemmer et al. (2006) conclude that many markets that recently experienced extraordinary house price appreciation (such as California, New York, Maryland and Virginia) are likely to experience increases in subprime foreclosure rates. The authors model foreclosures as a function of house price appreciation in different MSAs and find that foreclosures are more likely in markets with less house price appreciation. They also find that more borrowers refinance when they are delinquent in strong housing markets. For instance, for loans originated in 2001, one percentage point increase in annual housing price appreciation resulted in a 7.23 percent decrease in the odds of foreclosure and in a 2.84 percent increase in the odds of a distressed prepayment. This finding suggests that distressed prepayments and foreclosures are substitute outcomes that respond in opposite directions to a given change in housing prices. While housing price growth led to an increase in equity for many borrowers, it allowed them to refinance their homes even though they were behind in their monthly payments. When the authors add the distressed prepayments to their measure of the foreclosure rate, the total "failure rate" for subprime loans within 5 years of origination rises to a stunning 25 percent from 13 percent.

Determinants and Distribution of Subprime Lending

Since neighborhoods with higher subprime lending are more likely to default, it is important to know what factors are associated with subprime lending. Mayer et al. (2007) examine United States zip code level data to investigate what parts of the country and what types of neighborhoods had more subprime originations in 2005. The researchers limit their sample to first-lien, 1 to 4 family mortgages. Results from zip code level regressions indicate that subprime mortgages are concentrated in neighborhoods with a high share of black and Hispanic population, even while controlling for income and credit scores. Subprime loans are concentrated in zip codes

with more mid-level credit scores, and in counties with higher unemployment rates. The authors hypothesize that subprime loans might serve as an additional source of credit when economic conditions worsen. Subprime loans are also positively associated with new construction and with house price appreciation. However, spatial analysis indicates that there is only partial association between subprime lending and house price appreciation. Locations such as New York and Boston experienced high appreciation, but somewhat few subprime mortgages. However, locations such as Las Vegas and Miami experienced high subprime lending as well as rising new construction and house price appreciation. It is also unclear whether high subprime lending influences the house price appreciation or whether it is the other way around.

Calem et al. (2004) conduct an analysis of factors that contribute to the subprime share of home purchase and refinancing loans in 1999 in Chicago and Philadelphia. Tract-level regressions show a strong association between borrower and neighborhood risk with the share of subprime lending. Credit risk measures account for about half of the overall association of the demographic variables and the subprime share of refinance loans. Moreover, blacks are more likely to receive subprime loans, all else constant. The findings "suggest that concerns about potential disparate access to prime loans among African-American borrowers cannot be dismissed."

Keeping the determinants of subprime lending in mind, it is important to examine the experiences of subprime borrowers. Couchrane et al. (2004) analyze a 2001 Freddie Mac survey of subprime borrowers' decisions to enter the subprime market as well as to stay in it. When they model the decision to receive a subprime loan they find that credit risk (FICO score, front-end ratio, self assessed credit and the financial safety net) is the most important in explaining what type of loan borrowers get. In addition, variables such as Hispanic ethnicity and age 65 and over have a positive effect on a borrower's decision to receive a subprime loan. This analysis suggests that subprime borrowers may be subjected to higher costs for reasons beyond credit risk. Moreover, from their model the authors conclude that previous mortgage segment is an important determinant

of current market segment even after controlling for credit risk and demographic variables. In fact, 60 percent of people with subprime loans and only 13 percent of people with prime loans refinanced with a subprime loans. There might be some truth to the claim that subprime borrowers are getting "stuck" in the subprime segment.

Effects of Foreclosure on Municipalities

There is a general misconception that borrowers and lenders fully bear the cost of foreclosures. If this were the case, a rising foreclosure problem would not be such a huge concern for the affected neighborhoods. Unfortunately, residents living in areas near a foreclosed property and the municipalities bear a large part of foreclosure-related costs. Some costs to the municipality include policing, legal expenses, demand for city social service programs, loss of tax revenues, and a negative reputation for a municipality. The direst costs for municipalities can be gang activity, drug dealing, prostitution, and arson in vacant units. Moreover, the businesses in an affected neighborhood might suffer due to the precarious conditions around it.

A study by Apgar et al. (2005) uses the Budgets of the City of Chicago and Cook County from 2003 and 2004 to estimate the direct costs of foreclosures to a municipality under different foreclosure scenarios. The authors describe 26 separate costs for the provision of 'foreclosure related services,' and 15 separate government agencies that need to take action to incur these costs. The cheapest foreclosure that would cost the municipality $27 is when the loan does not cure, the property is never vacant, and the home is sold at a foreclosure sale. However, more complicated cases are quite a bit more expensive. If a unit becomes vacant, the city has to inspect the building and decide whether it wants to demolish or conserve it. When a demolition case goes to court, then the price jumps to around $13,000. Since in Illinois the foreclosure process typically lasts 13 months, the building may remain vacant for a while, which will force the municipality to spend resources to fight crime, accidental or purposeful arson, remove trash or lawn-mow. The highest that a municipality would have to spend was estimated at $34,199, if arson were involved.

A study by Immergluck et. al examines the effect of 1997 and 1998 foreclosures on house prices for over 9,600 properties that were sold in Chicago in 1999. The authors estimate foreclosure effects on single-family properties, excluding condos, multifamily rental properties and commercial buildings. Other things constant under the less conservative assumption, for each additional conventional foreclosure within an eighth of a mile of a house, property value is expected to decrease by 1.136 percent. Since the average sales price is $164,599 for homes in the city, this amounts to a decrease in value of approximately $1,870 per property because of a single foreclosure within an eighth of a mile. The magnitude of the impact for Chicago is between $598 million and $1.39 billion. For the entire city of Chicago, 3750 foreclosures in 1997 and 1998 reduced nearby property values by an average of $159,000 per foreclosure in the most conservative estimate, and by an average of $371,000 under the less conservative estimate.

In a similar study of foreclosures and crime rates in 2001 in Chicago, Immergluck et. al (2006) conclude that higher neighborhood foreclosure rates lead to more neighborhood crime. While the effect on property crime is not statistically significant, a one standard deviation increase of the foreclosure rate, .028, is expected to increase violent crime by 6.68 percent.

METHODS AND RESEARCH HYPOTHESES

This paper provides new evidence on the linkage between the economic and demographic characteristics of neighborhoods and the default incidences of those neighborhoods. This study contributes to the previous research by analyzing a more exhaustive list of neighborhood characteristics, drawing inspiration from variables used in Perkins (2008). Table 1 provides a listing of the variables used in this analysis along with some detailed descriptions. The subprime share of a neighborhood deserves particular attention, since numerous studies indicated that this variable is one of the most important determinants of foreclosure. In an OLS model of subprime

Table 1: Variable descriptions

Variables	Details
Foreclosure level data	
Foreclosure number	Number of pre-foreclosure and in foreclosure properties in 2006 in the Chicago area
Foreclosures per 100 00 units	Foreclosure number per 100 owner-occupied units
Quick foreclosure number	Foreclosure number originated in 2005 or 2006
Quick foreclosures %	Quick foreclosure number divided by foreclosure number
Loan level data	
Denial rate	Percent of prime home-purchase applications that were denied
Subprime owner-occupied #	Number of Owner-Occupied Subprime 1-4 family home purchase and refinancing loan originations in 2006
Subprime not owner-occupied #	Number of Non-Owner Occupied Subprime 1-4 family home purchase and refinancing loan originations in 2006
Prime owner-occupied #	Number of Owner-Occupied Prime 1-4 family home purchase and refinancing loan originations in 2006
Prime not owner-occupied #	Number of Non-Owner Occupied Prime 1-4 family home purchase and refinancing loan originations in 2006
Demographic tract level data	
Population total	
Recent immigrants	Number of immigrants who entered between 1995 and 2000
Bachelors %	Percent of 25 and over males and females who hold at least a bachelors degree
Public assistance %	Percent of households on public assistance income in 1999
65 and older %	Percent of population that is 65 years and older
Single %	Percent of families 15 years and over who are headed by a male householder without a wive or a female householder without a husband
Foreign %	Percent of population 5 years and over who is foreign born
Speak English poorly %	Percent of population over 5 years that speaks english not well or not at all
White population %	Percent of non-hispanic white population
Black population %	Percent of non-hispanic black populatin
Hispanic population %	Percent of hispanic population
Asian population %	Percent of non-hispanic Asian population
Economic tract level data	
Median household income	
Time to work aggregate	Aggregate travel time to work in minutes for workers 16 and over
Renter Burden	Median gross rent as a percentage of household income in 1999\
Owner Burden	Median selected monthly owner costs as a percentage of household income in 1999
Unemployment rate	
Vacancy rate	Proportion of housing that is "for sale only vacant" or "for rent only vacant"(available units) Divide the number of available units by the sum of occupied units and available units.
Management %	Percent of population in Management, professional and related occupations
Construction %	Percent of population with Construction/Extraction/Maintenance related occupations
Service %	Percent of population with Service occupations
Office/sales %	Percent of population with Office/sales occupations
Housing tract level data	
Average household size	
Median number of rooms	
Median value owner occupied	Median value of all owner occupied housing units
Owner occupied	Number of owner-occupied housing units
Renter occupied	Number of renter-occupied housing units
Median year structure built	
Homeownership %	Number of owner occupied housing units divided by the sum of the number of owner occupied and renter occupied housing units
Median Housing Sale Price Change municipality level data	
Median Price Change	Percent change in median house sale price during a specified period

lending as a function of neighborhood characteristics, this paper observes which variables are the most important in determining the subprime rate. Due to the high degree of association between foreclosures and the subprime share, it is reasonable to expect similar types of variables to affect both of these variables.

At an exploratory level, this study examines spatial concentrations of foreclosures and of a few important neighborhood characteristics to identify foreclosure "hot spots" and then to examine their demographic composition. To gain a more quantitative perspective, I estimate OLS and negative binomial regressions to model foreclosure levels and rates as a function of neighborhood characteristics. The purpose of the model is to uncover which variables are significant in explaining variation in foreclosures, as well as have the highest relative effect on foreclosures. This paper compares coefficients across models to test the robustness of the results.

I expect variables such as the share of blacks, Hispanics, single householders, those who speak English poorly, those on public assistance, those with a management/professional occupation; rates of vacancy, and unemployment; the owner-burden and the renter-burden to have a positive impact on foreclosures. Some of these relationships are speculative, as they have not received considerable attention in previous research. It seems reasonable to expect areas with high concentrations of single householders to be more prone to financial crisis, and thus more vulnerable to foreclosure. However, summary statistics on Contra Costa County in California from Perkins (2008) contradict this expectation. A high neighborhood vacancy rate may indicate the undesirability of an area potentially due to foreclosure-related risk factors. Areas with a high concentration of the share who speak English poorly probably have lower incomes and are more financially insecure, since it is hard to get a decent job without knowing the national language. The owner burden and the renter burden, which represent the median percent of income spent on housing costs, according to Perkins (2008) should have a positive impact on foreclosures. Perkins

(2008) summary statistics indicate that this is true, but she does not statistically test this relationship.

In line with the results of other studies, I expect variables such as the share of college graduates, the median income, and the median value of houses to have a negative impact on foreclosures. I failed even to speculate on the direction of the relationship of some variables with foreclosures. These variables include the foreign share, the number of recent immigrants (who arrived during 1995-2000), as well as the share of population that is 65 and older. It is unclear if areas with high concentrations of foreigners are more or less wealthy, as this depends on the nature of the immigrants. Similarly, it is unclear if areas with more recent immigrants are more desirable or less desirable because of it. Perkins (2008) speculates that mobility, and hence the number of recent immigrants, is associated with a higher share of foreclosures. However, her hypothesis is based the nature of Contra Costa County, which makes its applicability to the Chicago area doubtful. There has been some evidence from Calem et al. (2004) that areas with a higher concentration of population 65 and older are more likely to receive more subprime loans. Since the relationship between subprime lending and foreclosure is so stark, one can assume that the percent 65 and older would have a similar effect on foreclosures as on the subprime rates. However, this relationship has not been widely tested for different regions, so I am doubtful to speculate on its direction based on one study.

Inspired by a growing body of evidence on a positive relationship between house price appreciation and foreclosure, I examine this relationship at a neighborhood level. The idea is that when house prices are appreciating, then borrowers experiencing financial woes will still have positive equity due to the fact that a large part of their wealth, their house, has appreciated in price. Thus, when a borrower genuinely cannot pay his loans he can sell the house at a higher price and payoff his debt, or he can refinance into a better mortgage with his lender because his house value has appreciated. This would result in fewer foreclosures. Even though some evidence of this

relationship has been at the individual loan level or the neighborhood level, most of it is based on aggregate data. This is why it is important to test this hypothesis for a particular area, in this case the Chicago area, at a neighborhood level. I test the hypothesis by including the house price change variable in the model of foreclosures as a function of neighborhood characteristics. A negative coefficient on this variable provides further evidence for the appreciation hypothesis.

Similarly to the case of foreclosures, I explore the association between subprime lending and house price appreciation. Mayer et al. (2008) have found a positive relationship between subprime lending and appreciation. A reasonable explanation for this finding may be that when house prices are appreciating, subprime lenders don't have to worry as much about a lender's capacity to pay, and thus they grant more subprime loans. As Mayer noted, even though the overall relationship was positive, some areas in the US did not have this relationship. Thus, it makes it interesting to test this hypothesis on the Chicago area. Once again, I test the hypothesis by examining the coefficient on the house price change variable in the model of subprime lending as a function of neighborhood characteristics.

DATA

This report uses data from four different sources: foreclosure start listings from Record Information Services Inc at www.public-record.com, Home Mortgage Disclosure Act (HMDA), US Census 2000, and median housing sale price data from the Greater Chicago Housing and Community Development Website at http://www.chicagoareahousing.org/HousingHomePage.asp. Since this analysis uses census tracts as a proxy for neighborhoods, the data from all sources has been aggregated to census tracts. The data covers nine Chicago area counties (Cook, Will, Lake, Dupage, Kane, McHenry, Kendall, Winnebago and Dekalb), except for the house price data, which only covers six counties (and excludes Kendall, Winnebago, and Dekalb).

The Record Information Services data represents a complete listing of properties that are in pre-foreclosure or are scheduled for an auction in 2006. Not all properties in pre-foreclosure

actually make it to foreclosure, because the lender and the individual might work out a resolution enabling the borrower to pay off the debt. However, this study uses both the pre-foreclosure and in foreclosure listings as a proxy for the number of foreclosures. Listings where a loan is delinquent enough to enter the process of foreclosure are an important indicator of neighborhood distress, which is what the foreclosure variable is meant to measure in this study. In order to prepare this data for analysis I geocoded it to census tracts, removed all duplicate observations, and kept only single family homes, town homes and condominiums.[3] In the end, I was left with 33,053 foreclosure filings. Following several other studies, I used the number of owner-occupied units in 2000 to calculate the foreclosure rate. Unfortunately the number of owner-occupied units is not available for 2006, so this paper's measure of foreclosure rate is imperfect. In fact, all the Census 2000 neighborhood characteristics are measured in 2000, which limits the results of this study.

The HMDA data includes 2006 Loan Application Records data that financial institutions collect after applications for and originations of home-purchase, home-improvement and refinancing loans. This study aggregates this data mainly to estimate the subprime-lending rate for 1-4 family originations of home purchase and refinancing loans, as well as to calculate the prime home-purchase denial rate for each census tract. Even though this data lacks information on whether a loan is subprime, the rate spread variable is used to identify these types of loans. The rate spread indicates the amount by which a loan's APR exceeds a certain threshold depending on its lien status. According to the 2005 Federal Reserve Bulletin, "the Federal Reserve sought to select thresholds that would exclude the vast majority of prime rate loans and include the vast majority of sub prime loans."

The data from the Greater Chicago Housing and Community Development Website includes median quarterly housing sale price from the first quarter in 2003 through the first quarter

[3] A more detailed description of the data cleaning process may be found in the Appendix

in 2006 from 237 municipalities and 77 Chicago community areas. I used the percent change in house prices in order to estimate house price appreciation during a certain period.

EXPLORATORY DATA ANALYSIS

Summary statistics help to understand the variation in neighborhood characteristics in the Chicago area. I focus on six counties for which I have data on house prices, for a total of 1722 tracts. In order to examine characteristics of an average tract, I report the mean, the standard deviation, the minimum and the maximum values of all the explanatory variables in Table 2. The average number of foreclosures per tract is about 18 with a standard deviation of about 24. The spread of foreclosures appears to be pretty large with 118 tracts (6.85 percent) having no foreclosures, and 46 tracts (2.7 percent) having more than eighty foreclosures. The tract with the largest amount of foreclosures had 278 of them. The average number of foreclosures per 100 owner occupied housing units is 2.4 with a standard deviation of 4.3, which indicates a fairly large spread. A little more than a third of the loans originated were subprime, which indicates that the subprime segment is a major player in this housing market. The average house price change during 2003-2006 is about 27 percent, which means that on average houses have been appreciating in the area. In fact, only 140 tracts (8.1 percent) experienced zero or a negative house price change. The population is fairly diverse: about 51% white, 26% black, 4.1% Asian and 17.1% Hispanic. The rest of the demographic statistics indicate that the area is quite diverse, which makes it interesting to examine.

Next, I examine the types of neighborhood characteristics associated with high and low foreclosure tracts. So, I divide tracts into four quartiles based on foreclosure rate. Tracts are in the 1st quartile if they are in the bottom 25 percent of the foreclosure rate distribution and in the 4th quartile if they are in the top 25 percent of the foreclosure rate distribution. Table 3 shows the mean of each variable within the tracts in each quartile. The last column shows the percent change in the mean of each variable from the 1st quartile to the 4th quartile. The average number of

foreclosures per tract grows an enormous 1136 percent from 2.87 to 35.52, and similarly, the average number of foreclosures per 100 owner occupied units grows by 2619 percent from 0.24 to

Table 2: Summary of tract-level foreclosure data and explanatory variables

Variables	Mean	SD	Min	Max
Foreclosure level data				
Foreclosure number	17.7	23.6	0	278
Foreclosures per 100 OO units	2.4	4.3	0	100
Quick foreclosure number	5.7	7.4	0	85
Quick foreclosures %	31.9	22.7	0	100
Loan level data				
Subprime total %	36.1	19.9	0	100
Subprime owner-occupied %	35.9	20.6	0	100
Subprime not owner-occupied %	39.6	23.2	0	100
Denial rate	19.8	12.9	0	100
Subprime owner-occupied number	67	71	0	1,046
Subprime not owner-occupied number	9	11	0	122
Prime owner-occupied number	137	167	0	3,715
Prime not owner-occupied number	12	19	0	454
Demographic tract level data				
Population total	4,408	2,656	14	34,055
Recent immigrants number	195	307	0	2,755
Bachelors %	17.3	12.2	0	61.4
Poverty %	13.7	14.2	0	92.7
Public assistance %	5.4	7.5	0	76.4
65 and older %	11.1	6.5	0	100
Single %	30.6	20.3	0	100
Foreign %	17.2	16.0	0	76.3
Speak english poorly %	6.4	8.6	0	48.3
White population %	51.0	35.5	0	100
Black population %	26.0	37.4	0	100
Asian population %	4.1	7.0	0	84.5
Hispanic population %	17.1	23.6	0	100
Economic tract level data				
Median household income	$51,232	$25,072	$4,602	$200,001
Aggregate time to work	62,266	43,181	0	570,780
Renter burden	25.8	6.4	0	50.1
Owner burden	21.0	6.1	0	50.1
Unemployment rate	8.7	8.9	0	81.4
Vacancy rate	3.7	4.2	0	35.3
Management %	33.9	16.9	0	100
Construction %	22.9	12.8	0	100
Service %	14.6	7.4	0	48.1
Office/sales %	28.6	6.5	0	100

Housing tract level data				
Average household size	2.81	0.57	1.29	4.67
Median number of rooms	5.39	1.12	1.7	9.1
Median value owner-occupied (2000)	$177,501	$100,343	$9,999	$976,400
Owner occupied number	1,015	795	4	9,880
Renter occupied number	576	596	0	4,815
Median year built	1957	15	1939	1997
Homeownership %	60.2	26.3	0.6	100
Median Sale Price Change				
Change from 2003-2006	26.9	19.1	-49.3	128.2
Change from 2005-2006	9.4	11.8	-50.2	161.4
Change from 2004-2006	20.2	15.0	-27.5	109.0
Change from 2003-2004	5.9	12.1	-69.5	99.1
Change from 2003-2005	16.2	14.5	-76.0	105.5
Change from 2004-2005	9.8	11.9	-79.1	104.8
Median housing sale price in 2006	$280,416	$132,217	$95,500	$2,875,000
Number of tracts/observations: 1722				

Sources: www.public-record.com, HMDA 2006, Chicago Housing and Community Development Website, US Census 2000

6.47. The mean foreclosure rate is a lot higher in the 4th quartile than in any other quartiles, which indicates that tracts with high foreclosures experience severe problems, and thus need more attention. The mean percent subprime grows from 22–26 percent to 57 percent from the lowest to the highest quartile for owner-occupied and non owner-occupied loans. The fact that more than half of loans are subprime in 2006 in tracts with the highest foreclosure rates in 2006 raises concerns about the benefits of subprime lending. It is expected that subprime loans are more risky, but how much risk is too much risk? The denial rate on prime home-purchase loans grows from 13% in the 1st quartile to 31% in the 4th quartile, which is as expected since the denial rate partially proxies for credit risk.

The share of population in poverty, on public assistance, the share of single householders, the unemployment rate, and the vacancy rate all grow by approximately 100 percent from the lowest to the highest foreclosure tracts. Tracts with high concentrations of singles appear highly vulnerable to foreclosure, since one half of the householders in the highest foreclosure quartile are

singles. The share of college graduates, the share of foreigners, the share of those who speak English poorly, the share of Asians, and the share of whites drop 60-84 percent from the lowest

Table 3: Variables by foreclosure rate quartiles

Variables	1st quartile	2nd quartile	3rd quartile	4th quartile	% change
Foreclosure level data					
Foreclosure number	2.87	10.12	21.05	35.52	1136
Foreclosures per 100 OO units	0.24	0.80	1.80	6.47	2619
Quick foreclosure number	1.01	3.63	6.86	10.82	972
Quick foreclosures %	25.8	35.5	33.5	32.6	26
Loan level data					
Subprime total %	22.3	25.9	37.7	57.4	158
Subprime owner-occupied %	21.7	25.5	37.5	57.9	167
Subprime non owner-occupied %	26.1	32.5	41.4	57.3	120
Denial rate	13.4	14.6	19.2	31.4	135
Subprime owner-occupied #	26	54	90	96	277
Subprime non owner-occupied #	3	5	8	18	468
Prime owner-occupied number	109	165	184	90	-17
Prime not owner-occupied #	8	12	12	14	69
Demographic tract level data					
Population total	3,793	4,993	5,121	3,706	-2
Recent immigrants number	180	257	268	76	-58
Bachelors %	24.5	21.3	14.9	9.0	-63
Poverty %	11.9	8.1	11.7	22.9	93
Public assistance %	4.0	2.2	4.2	11.0	179
65 and older %	12.0	11.9	10.0	10.5	-12
Single %	23.0	20.9	27.7	50.0	117
Foreign %	20.3	21.6	20.4	7.0	-66
Speak english badly %	7.4	7.3	8.3	3.0	-60
White population %	64.0	69.3	54.4	17.8	-72
Black population %	9.8	5.1	16.5	70.7	618
Asian population %	6.2	5.6	4.0	1.0	-84
Hispanic population %	18.1	18.1	23.4	9.2	-49
Economic tract level data					
Median household income	$62,188	$57,999	$48,974	$36,614	-41
Aggregate time to work	53,880	72,465	72,635	49,863	-7
Renter burden	24.7	24.2	25.6	28.7	16
Owner burden	19.8	20.5	21.2	22.6	15
Unemployment rate	6.7	4.8	7.6	15.4	132
Vacancy rate	2.9	2.3	3.1	6.2	118
Management %	43.0	38.4	29.6	25.3	-41
Construction %	17.4	21.1	27.7	24.9	43
Service %	12.1	12.2	14.6	19.3	60
Office/sales %	27.4	28.3	28.2	30.5	11

Housing tract level data					
Average household size	2.65	2.69	2.88	3.01	14
Median number of rooms	5.51	5.53	5.29	5.24	-5
Median value owner-occupied	$251,664	$203,292	$147,435	$112,676	-55
Owner occupied number	898	1263	1182	714	-21
Renter occupied number	532	623	622	526	-1
Median year built	1954	1960	1960	1955	0
Homeownership %	58.8	66.5	63.5	51.9	-12
Median Sale Price Change					
Change from 2003-2006	24.8	23.3	27.0	32.2	30
Change from 2005-2006	8.4	9.1	9.3	10.7	28
Change from 2004-2006	17.4	18.0	20.6	24.4	40
Change from 2003-2004	6.9	4.8	5.5	6.4	-8
Change from 2003-2005	15.7	13.2	16.4	19.6	25
Change from 2004-2005	8.4	7.9	10.4	12.5	50
Median 2006 housing sale price	369,981	307,413	252,424	198,728	-46
Observations: 1722					

Sources: www.public-record.com, HMDA 2006, Chicago Housing and Community Development Website, US Census 2000

foreclosure tracts to the highest foreclosure tracts. The fact that tracts with more foreclosures have a lower share of foreigners and a lower share of those who speak English poorly indicates that it is the natives who are affected by foreclosure. There doesn't appear to be a clear relationship between the number of recent immigrants and foreclosures. This number grows from 180 in the first quartile to around 260 in the 2nd and 3rd quartiles but then drops dramatically to 76 in the 4th quartile. This relationship is similar to that of the share of Hispanics, suggesting that neighborhoods with more immigrants or Hispanics may experience more foreclosures, but they are not the ones experiencing the highest rates of foreclosure. On the other hand, the black population share increases by an incredible 618% from around 10% in the 1st quartile to 70% in the 4th quartile, indicating that neighborhoods with exceptionally high concentrations of blacks are the most affected by foreclosure.

The median household income, the median value owner-occupied, the median sale price in 2006 and the share with management occupations decrease by about 40-50 percent from the lowest to the highest foreclosure rate tracts. It makes sense that neighborhoods with lower incomes, lower house prices and a lower number of professionals are riskier and therefore are more affected by

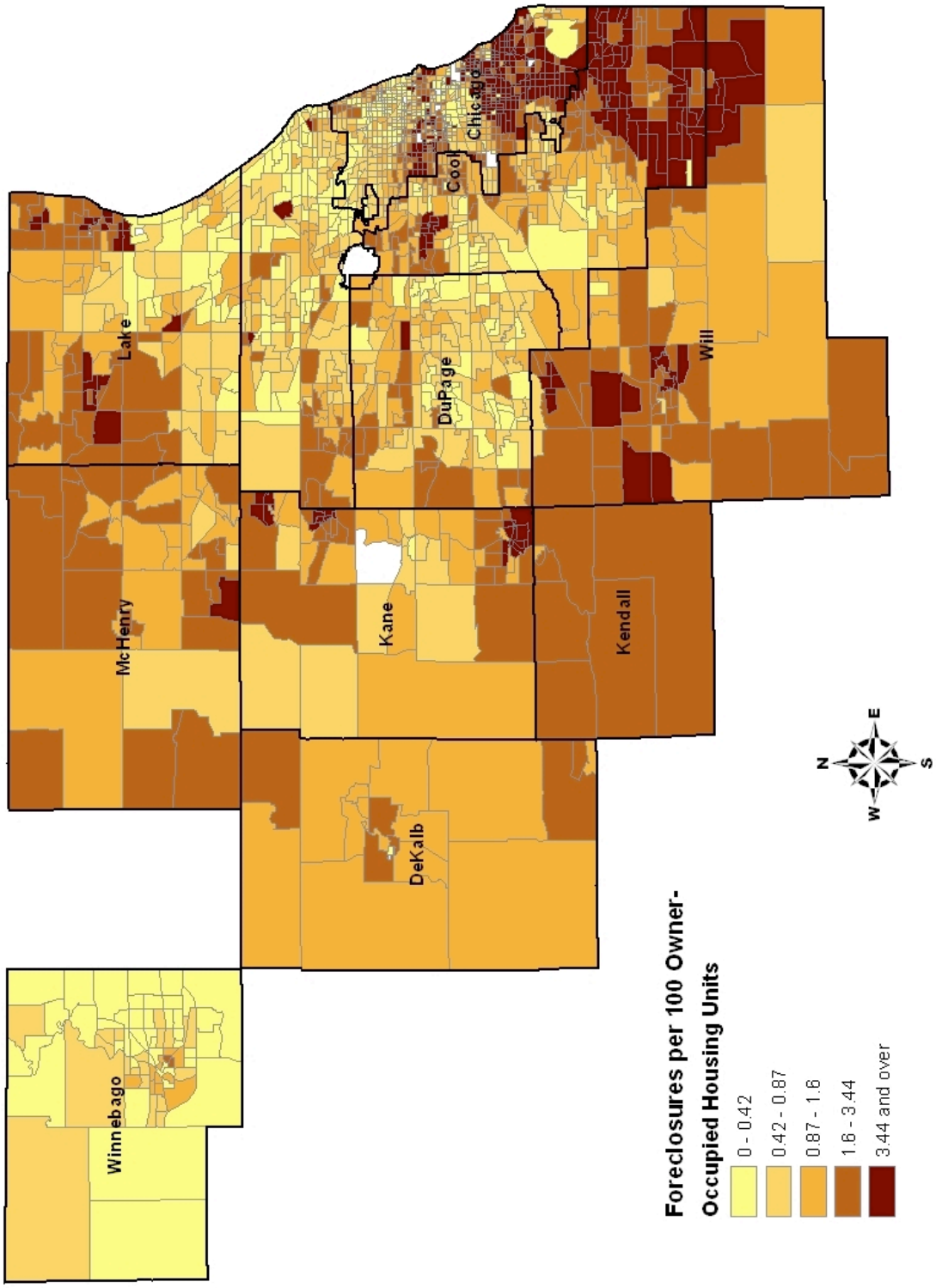
foreclosure. There doesn't appear to be a clear-cut relationship between the median-year built, the percent homeownership and foreclosures. The median housing sale price seems to have mostly grown by approximately 30 percent from the lowest foreclosure tracts to the highest foreclosure tracts. This is pretty surprising considering that I would expect tracts that experience high foreclosure to have lower appreciation rates.

Maps

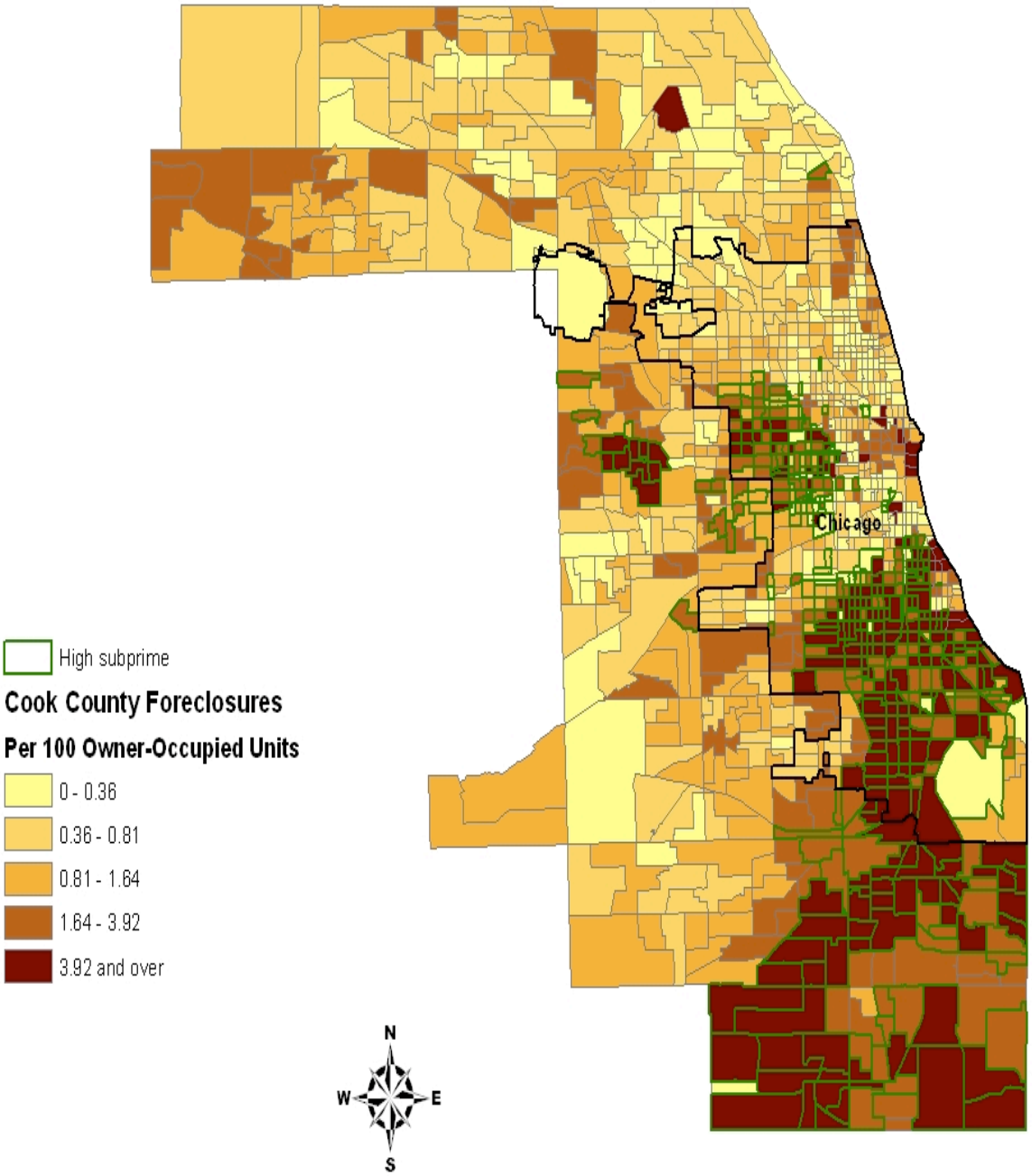
It is useful to examine the geographic distributions of foreclosure and demographics to identify variables that can predict high foreclosure areas. I draw maps of some of the demographics for the Chicago 9 county study area, which consists of 1912 tracts. Map 1 represents the distribution of foreclosures based on five quintiles. Tracts with the highest foreclosure rates are the darkest, while tracts with the lowest foreclosure rates are the lightest. Most of the highest foreclosure tracts, with more than 3.44 foreclosures per 100 owner-occupied units, are concentrated in south Chicago, south Cook county right below Chicago and the part of Will county that directly adjoins the southernmost boundary of Cook County. The fact that these tracts are so tightly concentrated indicates that this particular area is very vulnerable and deserves more attention. Perhaps this tight concentration may help policy makers to concentrate their efforts more efficiently.

Since the previous analysis indicates that subprime lending is highly associated with foreclosure rate, in Map 2, I draw the distribution of foreclosures in Cook county and highlight the tracts where the subprime rate is 50 percent or more. The tracts with a high subprime rate almost perfectly predict tracts with high foreclosure rates. This shows that tracts with the highest foreclosure rates in 2006 also had the highest subprime lending rates in 2006. This indicates that policy makers can use subprime rates to predict the high foreclosure tracts. Also, this shows the vulnerability of neighborhoods with a high proportion of subprime lending to foreclosure and raises the question of whether subprime lending is too risky.

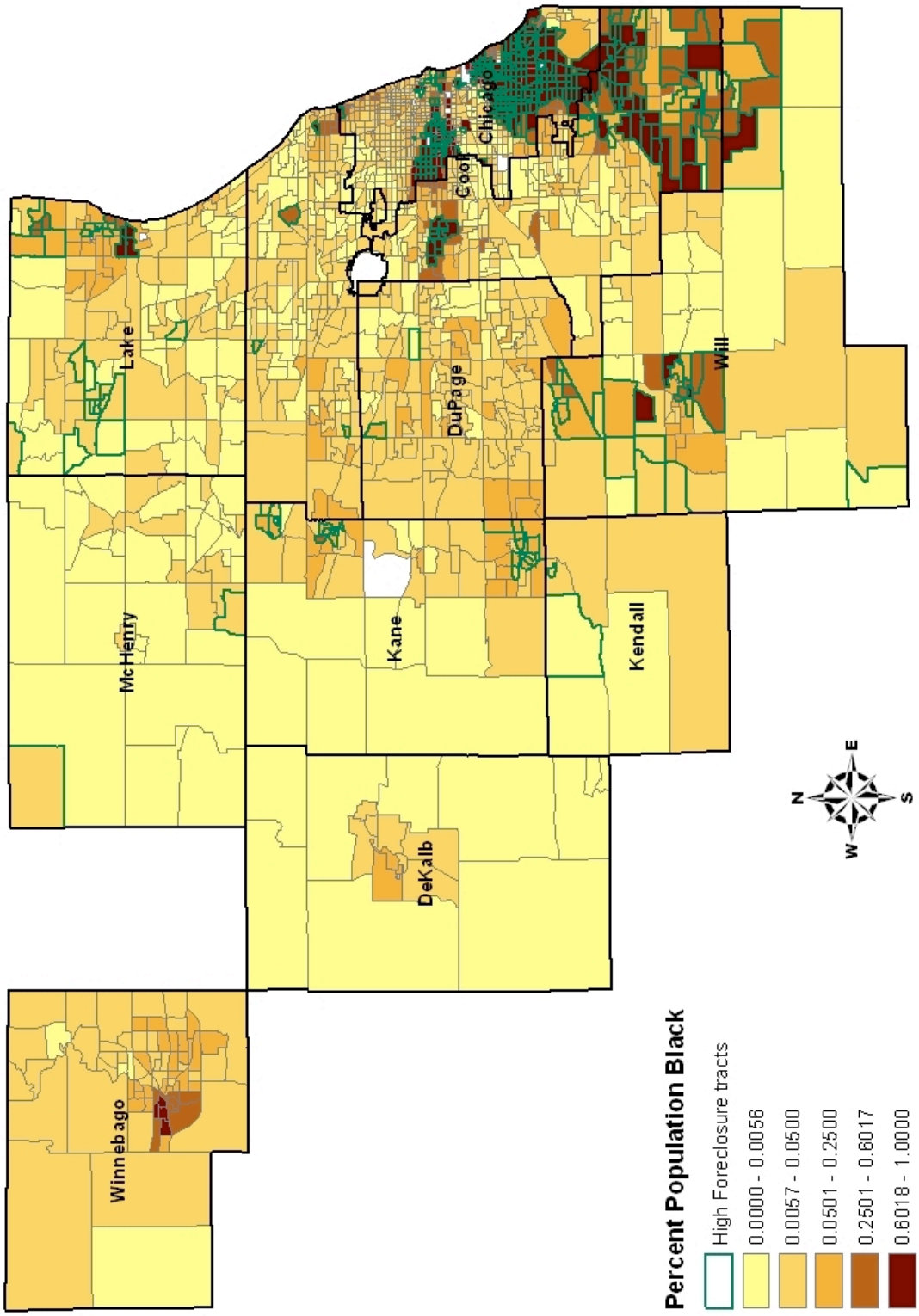
Map 1: The Distribution of Foreclosure Rates



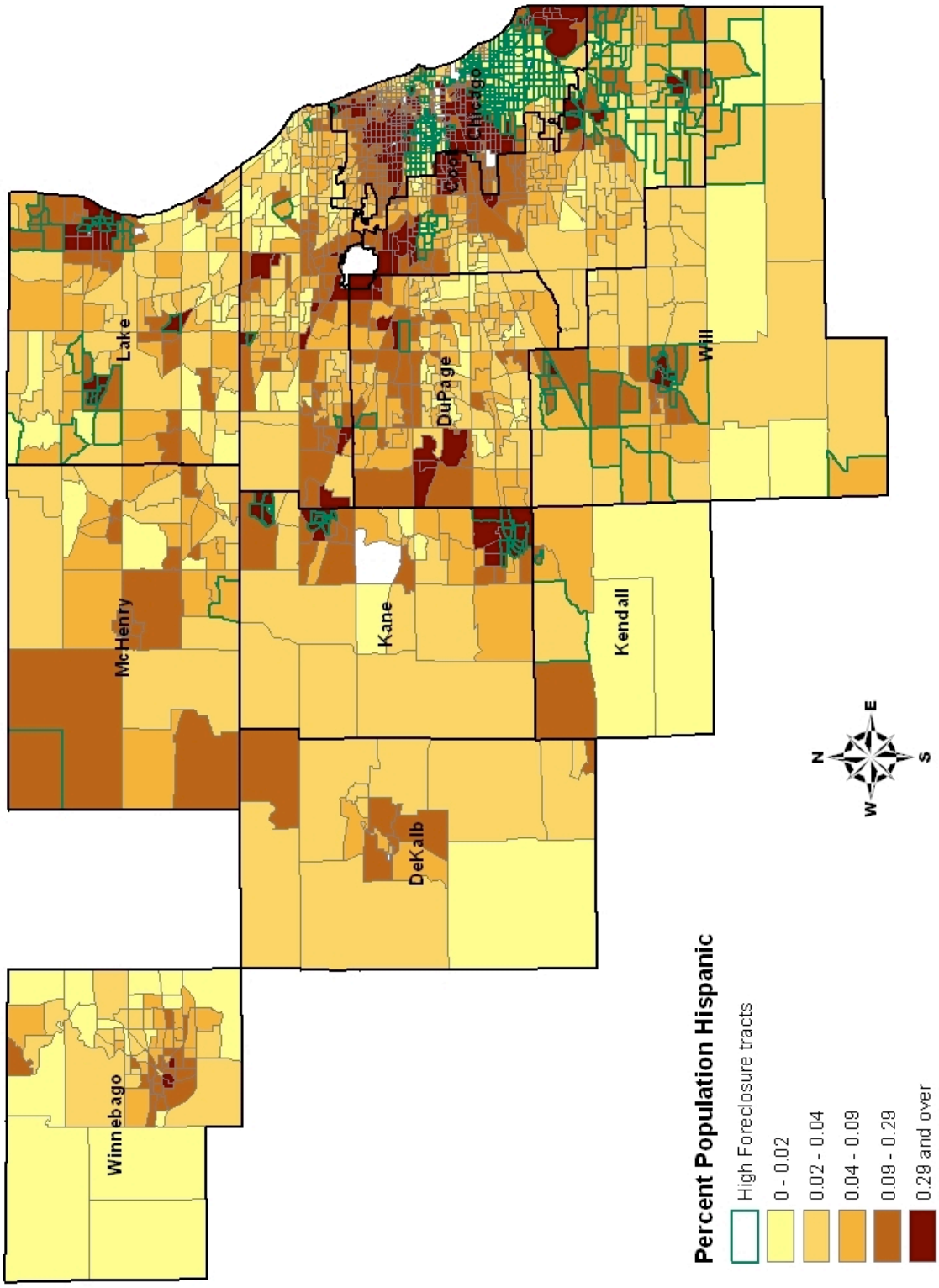
Map 2: The Distribution of Cook County Foreclosure Rates



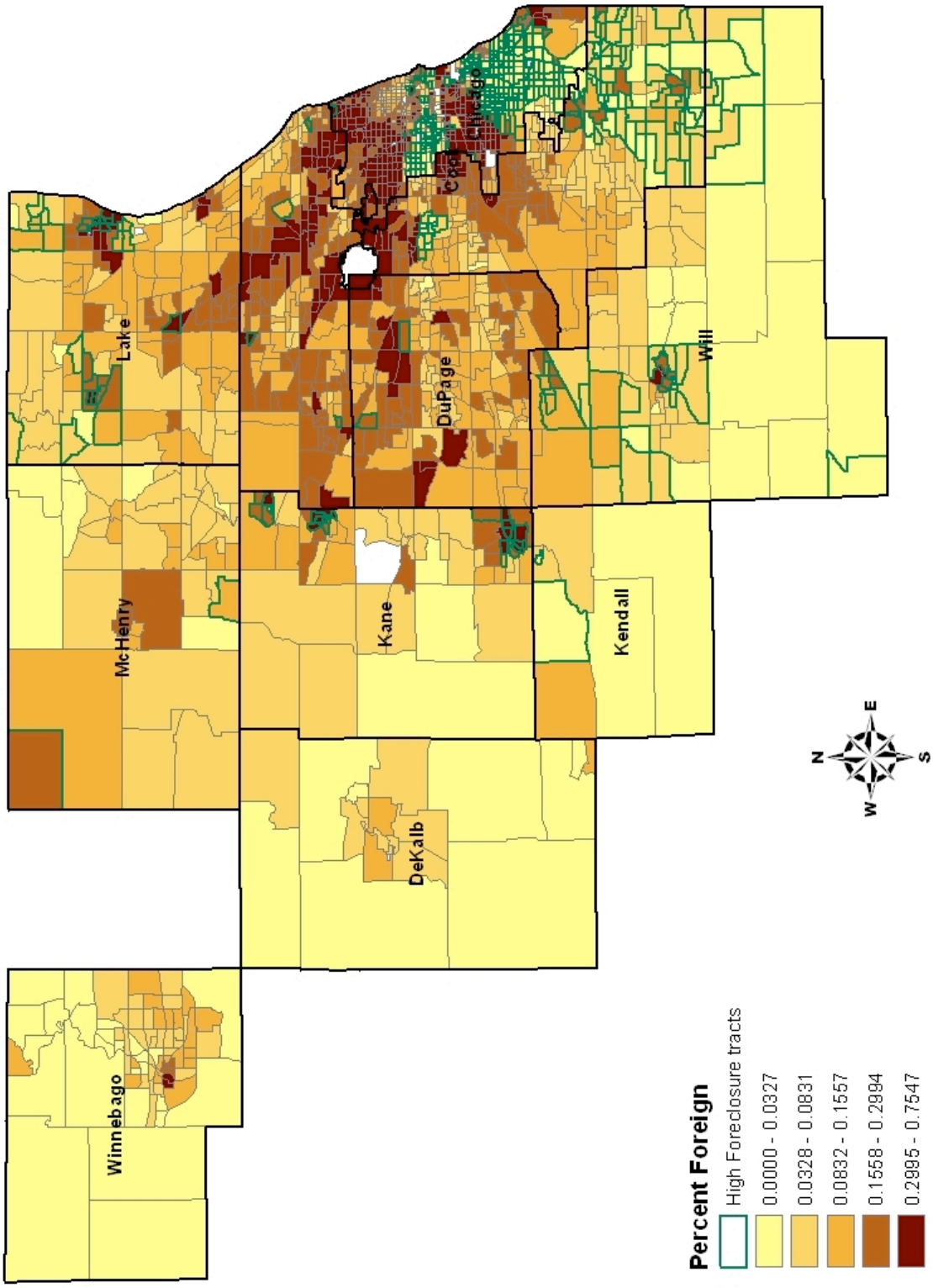
Map 3: The Distribution of the Black Population Share



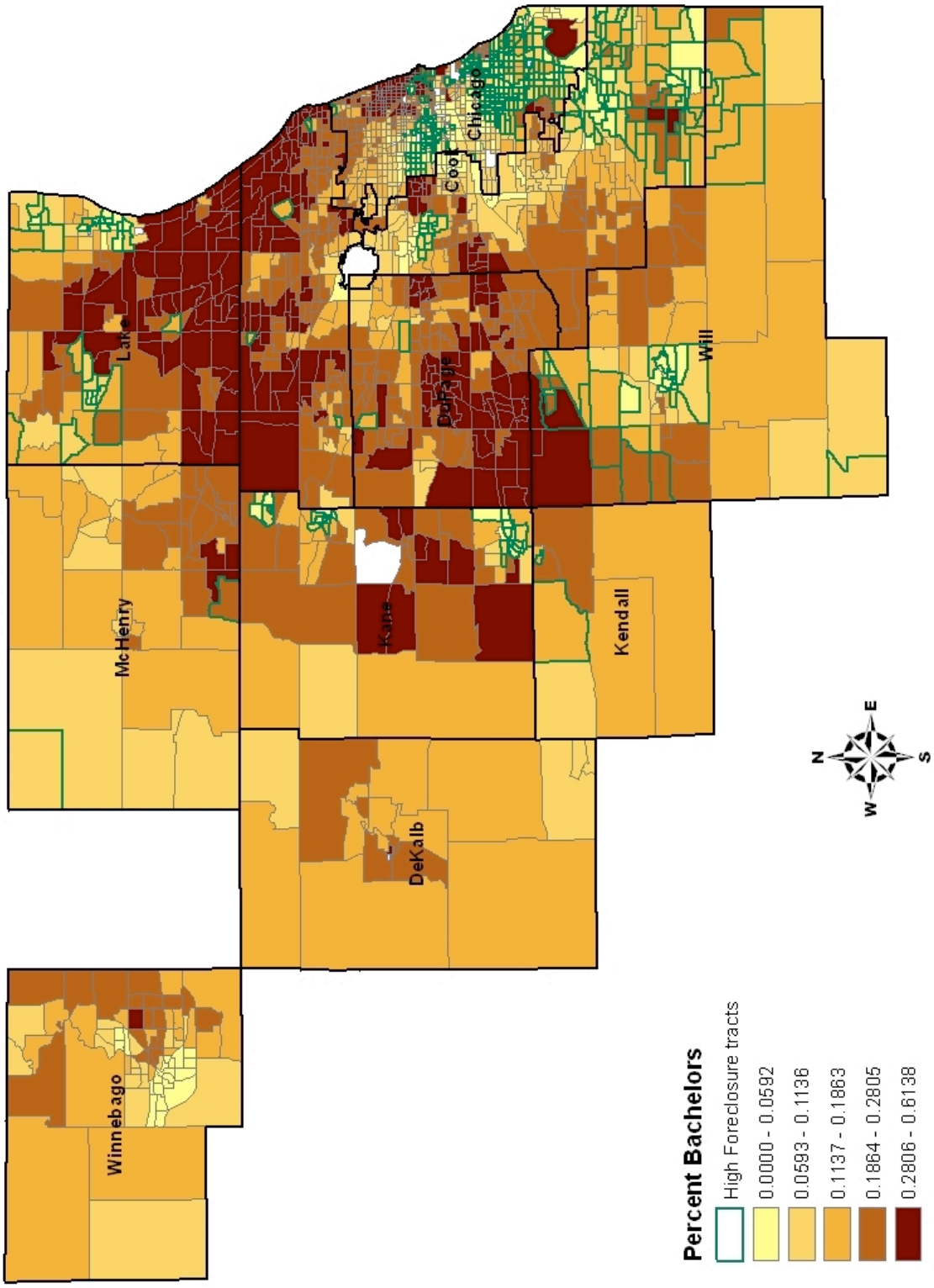
Map 4: The Distribution of the Hispanic Population Share



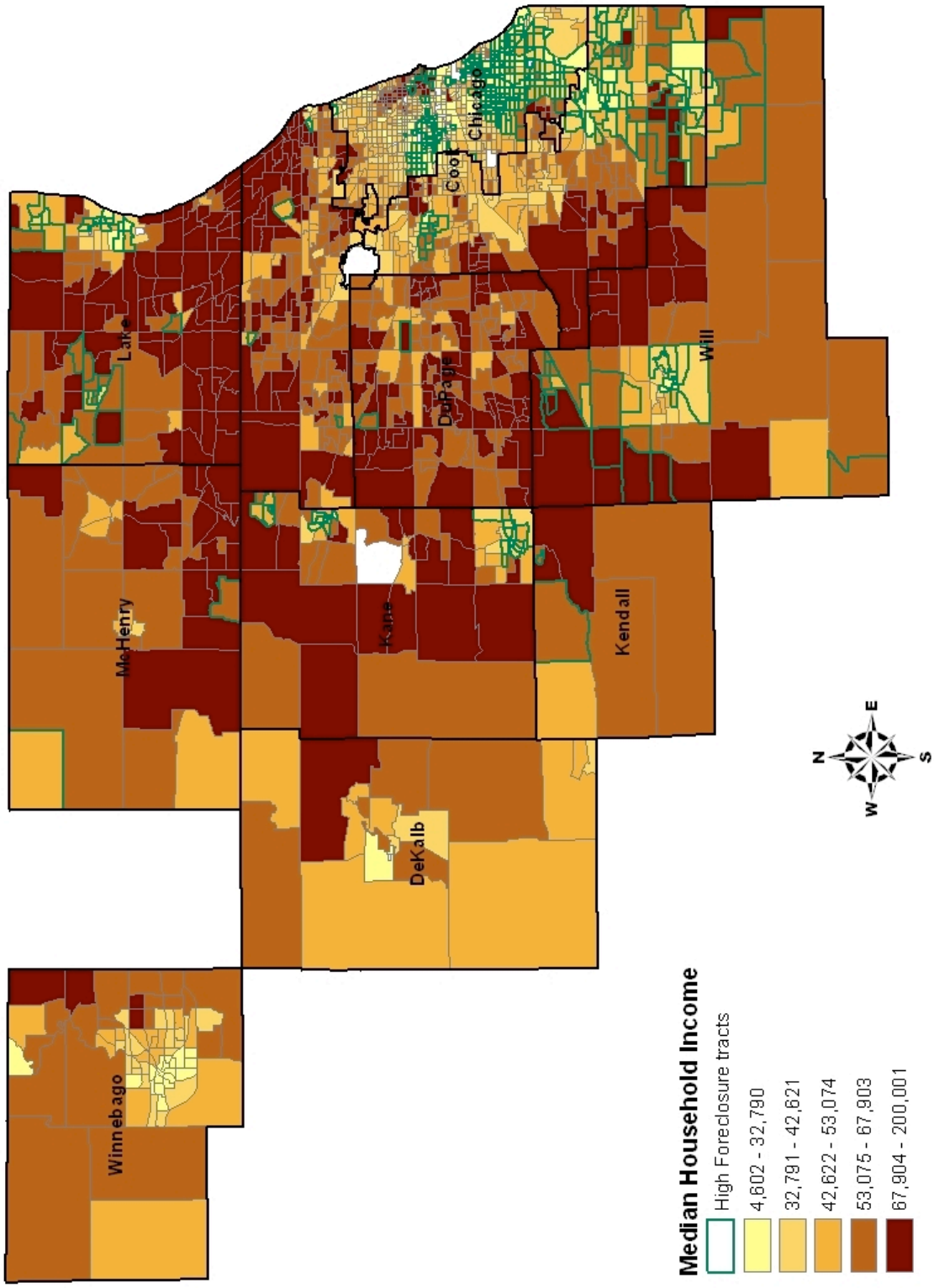
Map 5: The Distribution of the Foreign Population Share



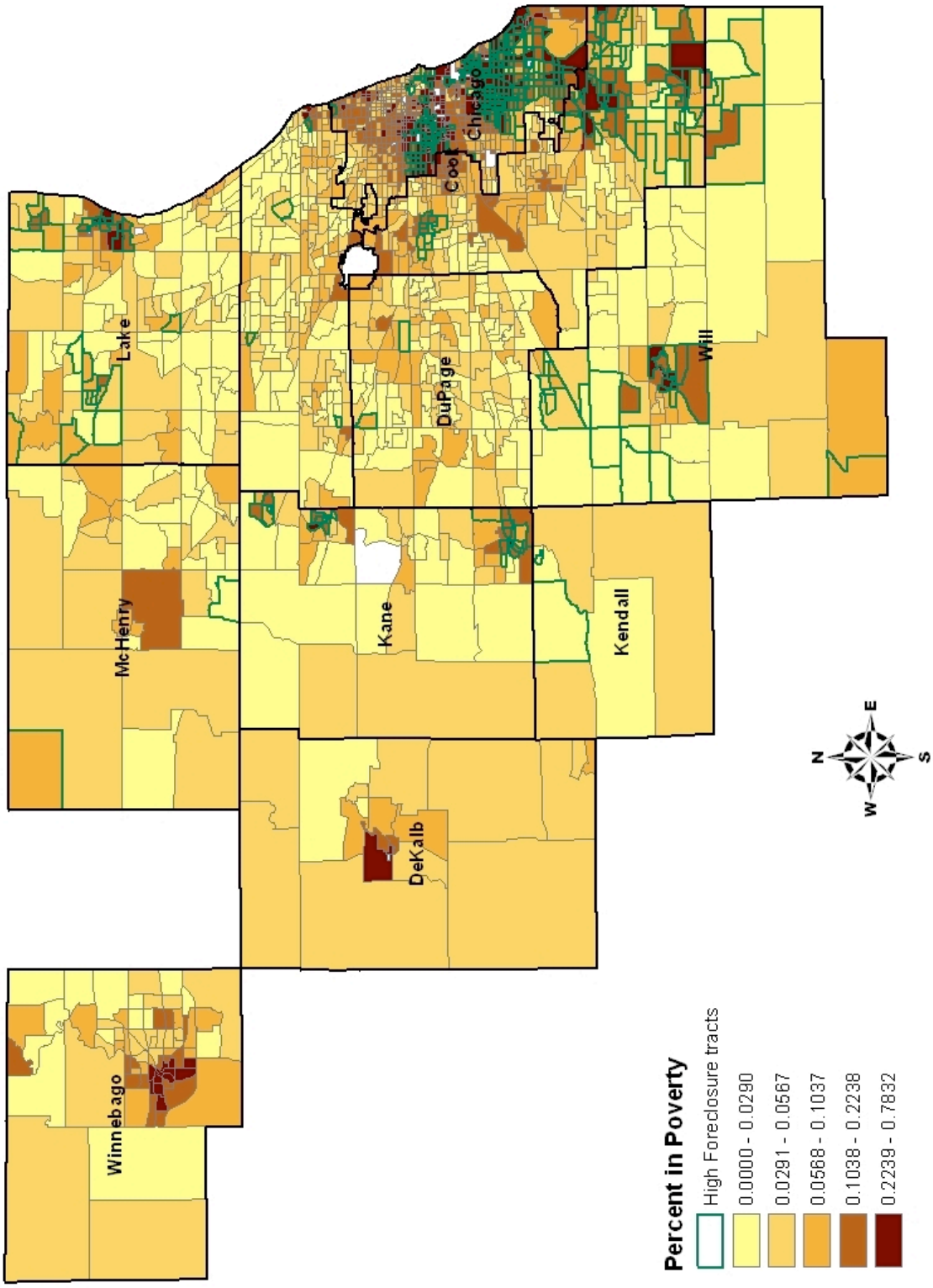
Map 6: The Distribution of the College Graduates Share



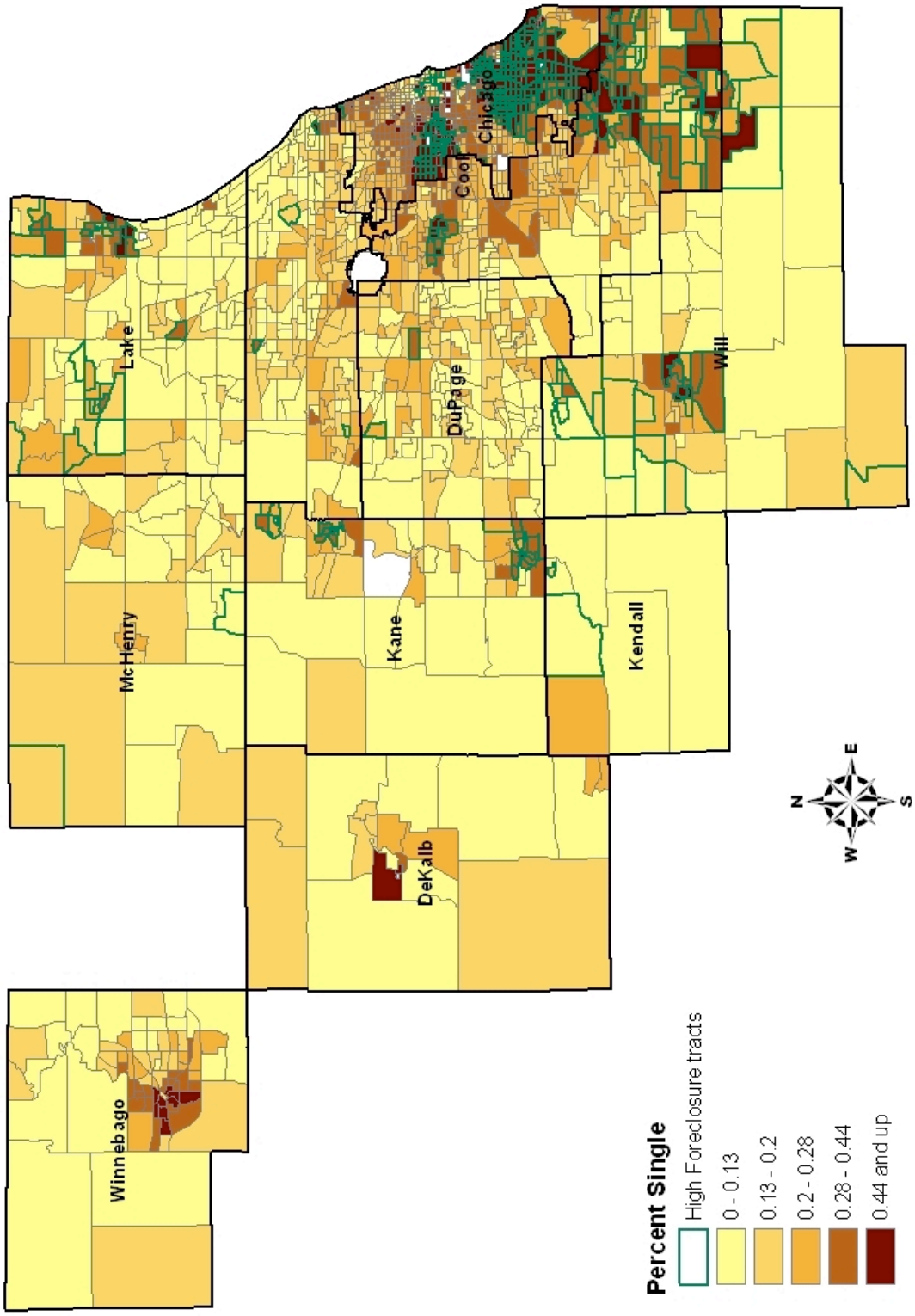
Map 7: The Distribution of Median Income



Map 8: The Distribution of the Poverty Share



Map 9: The Distribution of the Single Householders Share



Map 3 examines the distribution of black population where tracts with more than 3 foreclosures per 100 owner-occupied housing units (high foreclosure) are highlighted in green. In Chicago and Cook County, the high foreclosure tracts are all in areas with the highest percent population black (60-100 percent). The fact that blacks are concentrated in high foreclosure areas indicates that they are an important affected group of people. Map 4 presents the distribution of Hispanics and highlights high foreclosure tracts. In Cook county and Chicago, there is a very small percentage of Hispanics in high foreclosure tracts. However, in other counties tracts with very high percentages of Hispanics generally coincide with high foreclosure tracts. Map 5 shows the distribution of percent foreigners and highlights high foreclosure tracts. The distribution of foreigners is somewhat similar to that of the percent Hispanic, which could be just because many of the foreigners are Hispanic.

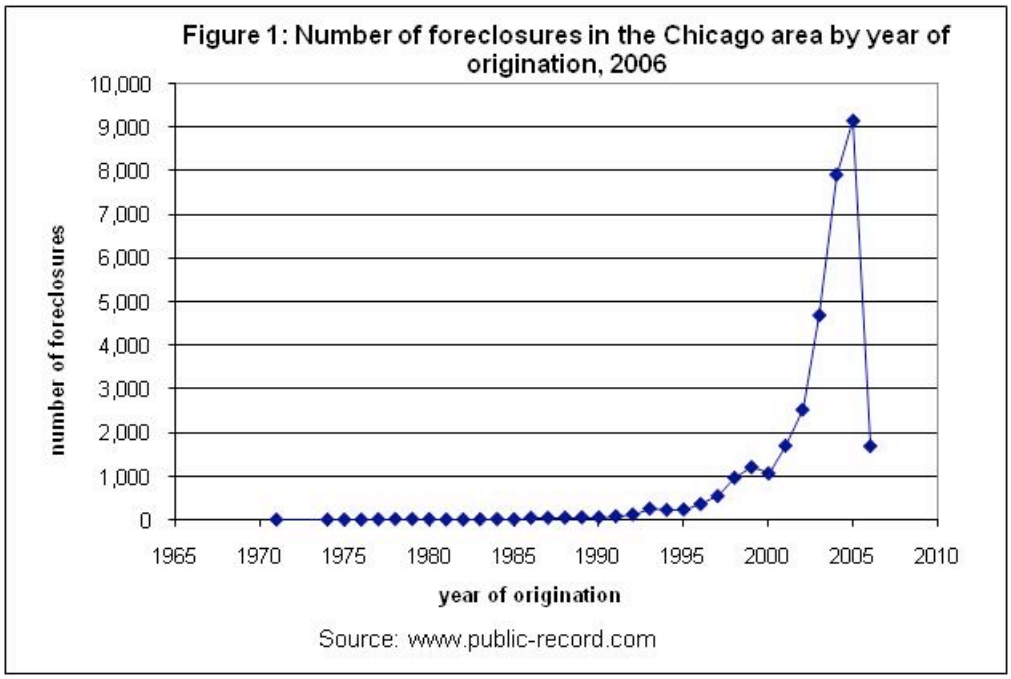
Map 6 presents the distribution of the college graduates share. Areas with high foreclosure rates have a low percent of people with bachelor's degrees. From the picture it seems as if the area surrounding Chicago from the north and from the west has very high concentrations of educated people. Map 7 shows the distribution of income with tracts with the highest foreclosure rate highlighted. In Chicago, the high foreclosure tracts have very low income. However, some tracts on the border of Will county and Cook county have pretty high incomes, but a high foreclosure rate. In other counties, the high foreclosure tracts are sometimes in areas with pretty high income. Similarly to tracts with high percent bachelors, tracts with high income seem to envelope Chicago from the North and from the west.

Map 8 examines the distribution of the percent in poverty. Even though a lot of high foreclosure tracts have high percent in poverty in Chicago, the poverty share is not nearly as good a predictor of high foreclosure tracts as the subprime or the black share. Map 9 presents the distribution of single householders. In Cook county areas with the highest foreclosure rates generally have the highest percent of householders that are single, which is 44 percent and higher.

The percent single unlike the percent in poverty seems to be a pretty good predictor of high foreclosure tracts.

Quick Foreclosures

From the maps it appears that foreclosures are highly concentrated in certain areas. However, it is difficult to tell if that is due to abusive lending or the high risk of the areas. Following Duda et al. (2005), I examine the year of origination of 33,053 foreclosure proceeding filings from www.public-record.com in Chicago in 2006. Concern of abusive lending may be partly justified if a disproportionate amount of recent mortgages are in foreclosure. Figure 1 shows the number of foreclosures in 2006 by the year of mortgage origination. Most mortgages in foreclosure were originated recently: about 86 percent of them were originated between 2000 and 2006. In fact, the largest number of defaults came from mortgages originated in 2005, which is just one year before foreclosure. It is really worrying that about a third of all foreclosures are quick foreclosures (foreclose the same year or the year after origination). This indicates that there might be a serious problem with lending standards, since most foreclosures are coming from extremely recent mortgages and not from older ones.

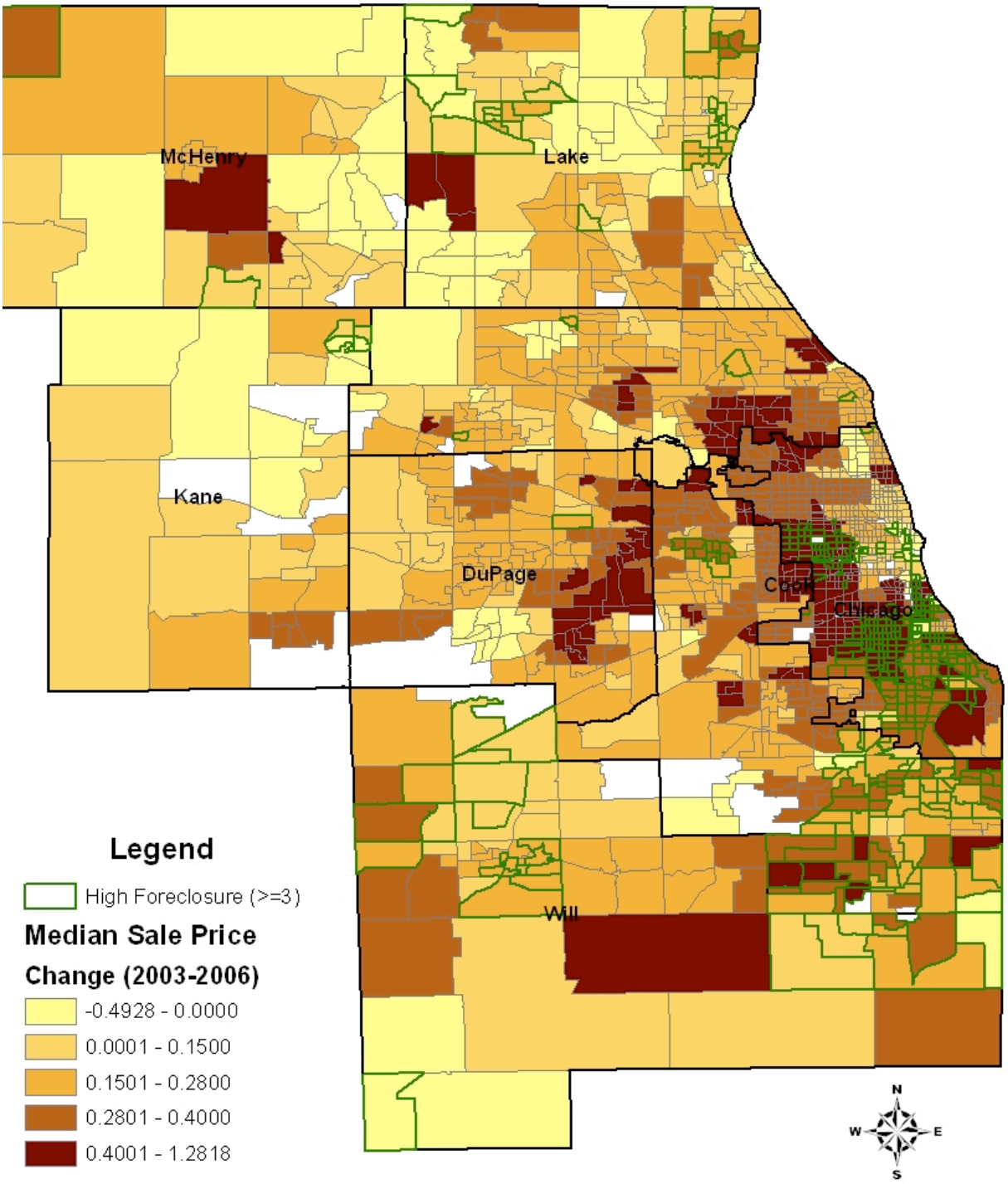


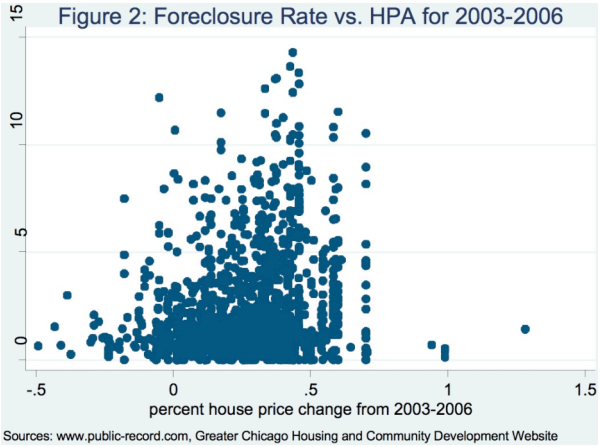
Since there is a fairly large percent of quick foreclosures in 2006, it is useful to examine their geographical distribution. Map 10 presents the distribution of quick foreclosures in 6 Chicago area counties, and highlights the high foreclosure areas in green. Please note that white tracts did not have house price change data. Areas with the highest foreclosure rates are generally not the areas with the highest quick foreclosure rates. In fact, the highest quick foreclosure tracts surround Chicago from the North and from West: North Chicago, north Cook county, a south part of Lake county and Dupage county. From the spatial analysis, this area also has the highest concentration of the wealthiest and the most educated people. This is merely speculation, but the wealthier areas might be more at risk of foreclosure in the future due to irresponsible lending.

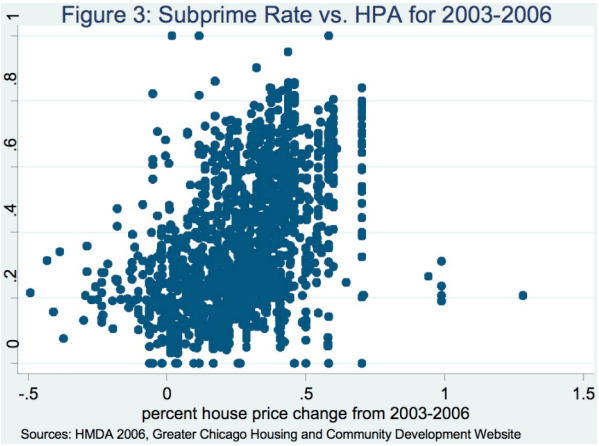
Housing Price Appreciation

It is interesting to explore the relationship between foreclosures and housing price appreciation (HPA), as suggested by some studies. Please note that this study measures HPA as the percent change of median house sale prices from 2003-2006. Results from the literature review suggest that there is a negative relationship between these factors. Using the foreclosure proceedings data in Chicago, Figure 2 shows a plot of the default rate against HPA at the tract level. The figure does not reveal any particular relationship between these two factors. Mayer et al. (2008) indicate that there is a positive relationship between subprime lending and appreciation. Figure 3 shows a plot of percent subprime owner-occupied against HPA at the tract level. The figure reveals a positive linear relationship between subprime lending and appreciation. However, there is a lot of variation in subprime lending for each HPA value. This relationship seems to make sense because when the housing market is strong lenders are more likely to be looser about their lending standards and thus grant more subprime loans.

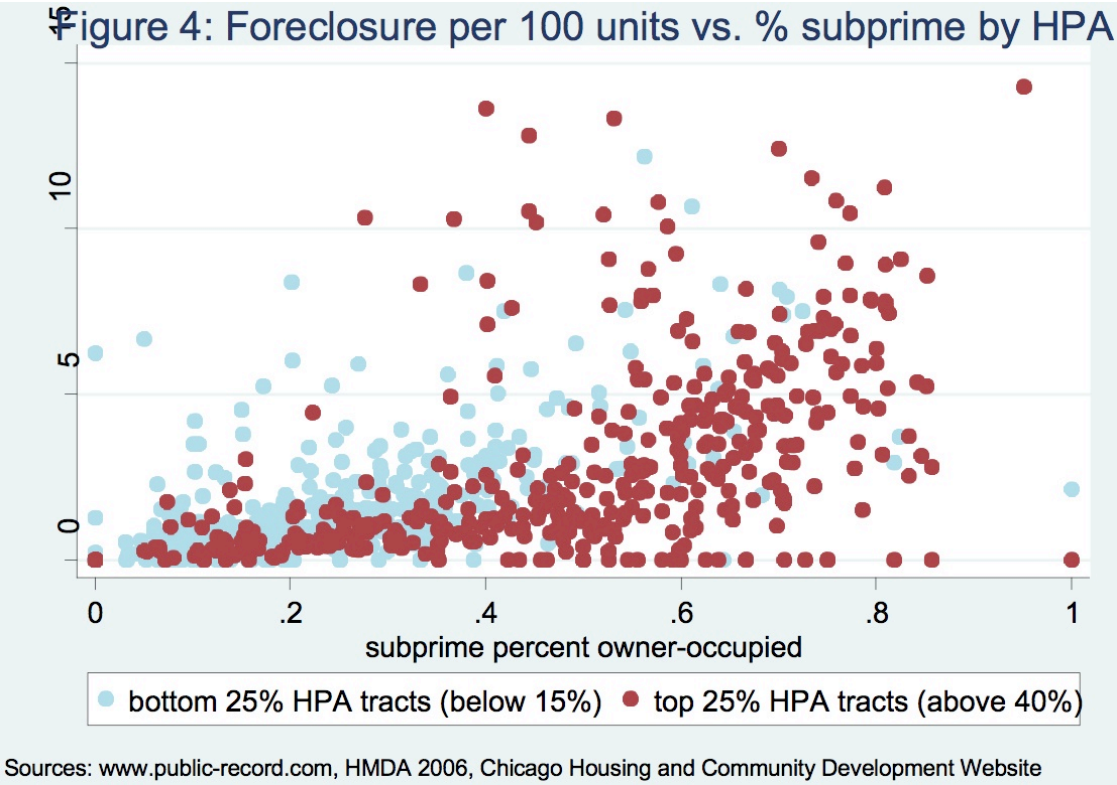
Map 10: The Distribution of the Median Sale Price Change for 2003-2006







Following Immergluck (2008), I explore the relationship between foreclosures, subprime lending and housing price appreciation at the tract level. Figure 4 shows a plot of the foreclosure rate against the subprime rate for tracts with the highest (top 25%) and the lowest (bottom 25%) house price change values. Tracts with low appreciation are mostly concentrated in areas with lower percent subprime lending (less than 45%). However, about half of high appreciation tracts had greater than 45 percent subprime lending. Moreover, the high appreciation tracts with lower subprime lending rates have lower rates of foreclosure compared to lower appreciation tracts with similar levels of subprime lending. The high appreciation tracts have higher foreclosure rates in areas with high subprime lending, but it is impossible to compare that to the foreclosure rates for lower appreciation tracts because there are too few of them with such high subprime lending. These results indicate that subprime lending is a crucial player in the relationship between foreclosure and housing price appreciation.

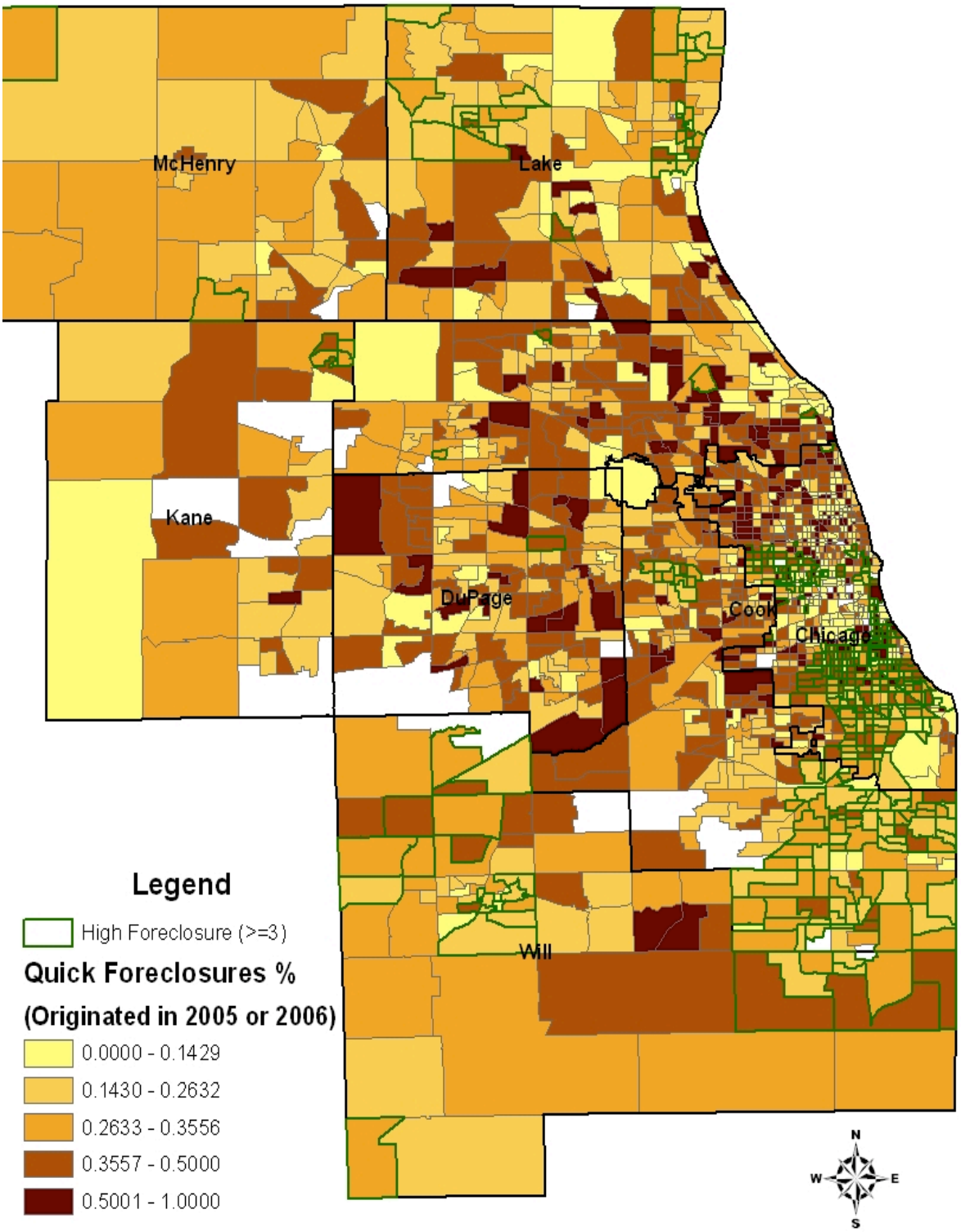


Finally, Map 11 represents the geographic distribution of house price change, where high foreclosure tracts are highlighted in green. Please note that white tracts did not have house price change data. The relationship between the foreclosure rate and the house price appreciation rests unclear in this map. In Chicago, a lot of the high foreclosure tracts had high or the highest house price appreciation, while in other high foreclosure tracts house price appreciation wasn't as high. These results call for more careful statistical analysis in order to uncover some hidden or ambiguous relationships in the data.

STATISTICAL ANALYSIS

It is important to keep in mind that this study deals with the neighborhood and not the individual determinants of foreclosure. However, the microeconomic story of the borrower is at the heart of this problem, since determinants of foreclosure are individual in nature. Foreclosures arise because borrowers with mortgage contracts make a decision not to pay off their loan. This decision is based on the unique circumstances of the individual, and not on neighborhood characteristics. Indeed, the neighborhood where an individual resides is a partial reflection of him, but not a

Map 11: The Distribution of Quick Foreclosures



complete one. Thus, even though it is possible to speculate about the individual determinants of foreclosure from this study, the analysis would not be particularly convincing.

Results from the exploratory data analysis indicate that some neighborhood characteristics are associated with foreclosures; however, it is unclear which ones are the most important in explaining foreclosure rates. It is also clear that there is some relationship between foreclosures, subprime lending and house price appreciation, but the direction of the relationship remains mysterious. Firstly, the regression analysis that follows tries to determine the importance of neighborhood characteristics in explaining the foreclosure rate as well as the subprime rate. Secondly, it tries to tease out the relationship between foreclosures, house price appreciation, and subprime lending.

I present ordinary least squares (OLS) tract level regressions relating the foreclosure number and the foreclosure rate to a number of explanatory variables. It is interesting to see if the results are robust to the two measures of foreclosure. The models are:

(1) Foreclosure number = α + β_1*LOAN LEVEL + β_2*DEMOGRAPHIC + β_3*ECONOMIC + β_4*HOUSING + β_5*SALE PRICE CHANGE

(2) Foreclosures per 100 housing units = α + β_1*LOAN LEVEL + β_2*DEMOGRAPHIC + β_3*ECONOMIC + β_4*HOUSING + β_5*SALE PRICE CHANGE

Since it deals with levels, Model (1) includes the number of owner or renter occupied units. With this exception, variables included in the LOAN LEVEL matrix are the only explanatory variables that are different across both models. Following Immergluck (2004), Model (1) LOAN LEVEL variables include the number of subprime or prime owner-occupied loans and the number of subprime or prime not owner-occupied loans. Model (2) LOAN LEVEL variables include

percent subprime owner-occupied or not owner-occupied loans. In both models SALE PRICE CHANGE is crucial to determine the relationship between house price appreciation and foreclosure.

OLS might not be the best to model foreclosure number, because there are around 118 tracts (6.85 percent) with no foreclosures. In this situation it is actually better to estimate a negative binomial regression. The purpose of this model is to model counts, where the variance is greater than the mean (over dispersion around the mean). However, I choose to estimate an OLS instead due to the difficulty in interpretation of the negative binomial coefficients. Negative binomial results are presented in order to compare signs and significance of variables from the OLS regression. This analysis focuses on the standardized coefficients, called the beta values. The beta values signify the effect a one standard deviation change in the explanatory variable on one standard deviation of foreclosure number. These coefficients are useful because they allow for comparison of the relative strength of the explanatory variables on the dependent variable. Table 4 summarizes the results from the regression of foreclosure levels on the explanatory variables.

Foreclosures and Neighborhood Characteristics

The LOAN LEVEL variables have the greatest impact on foreclosures. This is reasonable since foreclosures are decisions by people with loans. Similarly to Immergluck (2004), the number of subprime owner-occupied loans has the greatest effect on the number of foreclosures. A one standard deviation increase in the number of subprime owner-occupied loans increases the number of foreclosures by 0.8 of a standard deviation of foreclosure number. That is, an increase of 67 in the number of owner-occupied subprime loans results in an increase of 19 in the number of foreclosures in a tract all else constant. This is an enormous effect on foreclosures, considering the average number of foreclosures per tract is 17. However, an increase of 167 in prime owner-occupied loans decreases the number of foreclosures by 7.7.

Table 4: OLS with the number of foreclosures as a dependent, 2006

	Price change for 2003-2006			No price change		
Variables	**Coef**	**SE**	**Beta**	**Coef**	**SE**	**Beta**
Loan Level data						
Subprime owner	0.264 *	0.007	0.800	0.265 *	0.007	0.803
Subprime not owner	0.519 *	0.036	0.242	0.521 *	0.036	0.243
Prime owner	-0.045 *	0.003	-0.320	-0.047 *	0.003	-0.331
Prime not owner	-0.024	0.019	-0.019	-0.013	0.020	-0.010
Denial rate	-8.085 *	2.857	-0.044	-9.702 *	2.870	-0.053
Demographic tract level data						
Population total	-0.227 *	0.034	-0.255	-0.237 *	0.035	-0.266
Recent immigrants	-0.004 *	0.002	-0.052	-0.003 *	0.002	-0.040
Bachelors %	-2.765	5.043	-0.014	-1.795	5.086	-0.009
Public assistance %	-8.190	6.129	-0.026	-9.360	6.177	-0.030
65 and older %	-18.164 *	5.431	-0.047	-22.735 *	5.419	-0.059
Single %	-15.590 *	3.719	-0.134	-17.412 *	3.721	-0.150
Foreign %	-15.727 *	4.960	-0.106	-19.735 *	4.948	-0.133
English bad %	33.569 *	8.677	0.123	34.828 *	8.760	0.127
Black population %	13.265 *	1.885	0.210	12.492 *	1.893	0.197
Hispanic population %	0.435	2.847	0.004	0.458	2.876	0.005
Asian population %	3.725	5.241	0.011	3.144	5.292	0.009
Economic tract level data						
Median household income	-0.851 *	0.305	-0.089	-0.974 *	0.298	-0.102
Time to work aggregate	0.133 *	0.022	0.243	0.141 *	0.023	0.257
Unemployment rate	-6.022	4.612	-0.023	-7.576	4.650	-0.029
Renter burden	0.021	0.040	0.006	0.020	0.040	0.005
Owner burden	-0.026	0.039	-0.007	-0.042	0.039	-0.011
Vacancy rate	17.689 *	7.136	0.031	18.807 *	7.203	0.033
Management %	6.811	3.650	0.049	8.088 *	3.680	0.058
Service %	2.466	5.257	0.008	4.350	5.301	0.014
Housing tract level data						
Median value owner-occupied	0.563	0.515	0.024	0.492	0.487	0.021
Owner occupied	0.006 *	0.001	0.191	0.006 *	0.001	0.186
Renter occupied	-0.002	0.001	-0.041	-0.002	0.001	-0.046
Median year built	0.088 *	0.021	0.057	0.123 *	0.021	0.080
Homeownership %	4.590	2.816	0.051	3.537	2.835	0.039
Median Sale Price Change						
Median Price Change	-9.008 *	1.496	-0.073			
Median Sale Price in 2006	-0.249	0.275	-0.014			
Constant	-164.982	42.360		-234.032	41.232	
Observations:	1711			1718		
Adj. R squared:	0.84			0.84		

Note: divided population total by 100, divided median household income by 10,000, divided time to work by 1,000, divided median value owner-occupied and median sale price in 2006 by 100,000.

* indicates significance at the 0.05 level.

Table 5: Negative Binomial with foreclosure number as dependent

	Price change 2003-2006			No price change		
Variables	**exp(coef)**		**SE**	**exp(coef)**		**SE**
Loan Level data						
Subprime owner	1.00538	*	0.00043	1.00537	*	0.00043
Subprime not owner	1.01282	*	0.00206	1.01227	*	0.00203
Prime owner	0.99912	*	0.00020	0.99923	*	0.00020
Prime not owner	1.00210		0.00116	1.00148		0.00113
Denial rate	2.77809	*	0.57540	2.99592	*	0.61656
Demographic tract level						
Population total	1.00004		0.00002	1.00005		0.00002
Recent immigrants	1.00029	*	0.00010	1.00023	*	0.00010
Bachelors %	0.14457	*	0.05113	0.15508	*	0.05445
Public assistance %	0.11691	*	0.05148	0.13869	*	0.06045
65 and older %	0.08780	*	0.03044	0.11848	*	0.04103
Single %	0.98953		0.24592	1.14180		0.28264
Foreign %	0.59373		0.19583	0.80195		0.26541
Speak English Poorly %	0.46378		0.29286	0.43062		0.26828
Black population %	1.69352	*	0.19489	1.71314	*	0.19669
Hispanic population %	0.65039	*	0.12270	0.62817	*	0.11722
Asian population %	0.29634	*	0.11539	0.31037	*	0.11885
Economic tract level data						
Median household income	1.00000		0.00000	1.00000		0.00000
Time to work aggregate	0.99999	*	0.00000	0.99999	*	0.00000
Unemployment rate	0.26026	*	0.08667	0.28002	*	0.09257
Renter burden	1.00579	*	0.00243	1.00584	*	0.00241
Owner burden	0.99944		0.00265	1.00058		0.00263
Vacancy rate	9.84862	*	4.70308	8.62997	*	4.06403
Management %	1.34806		0.33825	1.25779		0.31222
Service %	0.59333		0.20917	0.53891		0.18782
Housing tract level data						
Median value owner occupied	1.00000	*	0.00000	1.00000	*	0.00000
Owner occupied	1.00031	*	0.00008	1.00030	*	0.00008
Renter occupied	1.00055	*	0.00006	1.00056	*	0.00006
Median year built	1.01027	*	0.00132	1.00822	*	0.00134
Homeownership %	6.53164	*	1.18875	6.75948	*	1.22143
Median Sale Price Change						
Median Price Change				0.65121	*	0.05963
Median Sale Price in 2006				0.99999	*	0.00000
Constant	-14.761		2.631	-18.730		2.578
Observations	1711			1718		
Log Likelihood	-5336.340			-5377.887		

Note: * indicates significance at the 0.05 level

Table 6: Comparison of OLS with Negative Binomial (NB) Results where foreclosure number is a dependent, 2006

Same significance/sign	Significant only in OLS	Significant only in NB	Switching signs
Negative	***Negative***	***Negative***	***Positive in NB***
Median Price Change	Single %	Median Sale Price in 2006	Denial rate
Prime owner	Foreign %	Public assistance %	Population total
65 and older %	Median income	Unemployment rate	Recent immigrants
Positive	***Positive***	Bachelors %	Time to work aggreg
Subprime owner	English bad %	Hispanic population %	
Subprime not owner		Asian population %	
Black population %		Median value owner occupied	
Owner occupied		***Positive***	
Median year built		Homeownership %	
Vacancy rate		Renter burden	
None		Renter occupied	
Prime not owner			
Owner burden			
Management %			
Service %			

Of all the neighborhood characteristics the black share has the largest impact on foreclosures. A one standard deviation increase in this variable results in an increase of 5 in the number of foreclosures, all else equal.

Even though many other coefficients are significant, I first compare them to the negative binomial results. Table 5 shows the results from the negative binomial with exponentiated coefficients. A coefficient less than one indicates a negative impact on foreclosures, while a coefficient above one indicates a positive impact on foreclosures. Table 6 compares OLS with negative binomial results by listing variables for which coefficients were consistent across models, a coefficient was significant in only one model or significant coefficients switched signs. The number of prime owner-occupied loans and the share of population 65 and older have a negative and significant effect on the number of foreclosures in both models. The number of subprime owner-occupied as well as not owner-occupied loans, the black share, the number of owner-occupied units, the median year built and the vacancy rate all have a positive and significant effect on the number of foreclosures in both models. In fact, a one standard deviation increase in the

share of population 65 and over results in approximately one less foreclosure, while a one standard deviation increase in the median year built results in approximately one more foreclosure.

However, the share of single householders, the share of foreigners, the median income and the share of those who speak English poorly are only significant in the OLS. Even though exploratory analysis shows association between these variables and foreclosure, they are not crucial in explaining foreclosure variation. On the other hand, many variables are only significant in the negative binomial, which indicates that those variables are important in explaining foreclosure counts. Variables such as the median sale price in 2006, the share on public assistance, the share of college graduates, the Hispanic and Asian share, the unemployment rate have a negative effect on the number of foreclosures, while homeownership and renter burden all have a positive effect on foreclosure in the negative binomial. The effects are as expected, except for public assistance which actually decreases foreclosures. Similar to exploratory analysis, higher Hispanic concentrations actually result in lower foreclosures. This is surprising since many other studies have found the opposite effect. Unfortunately, since these variables are insignificant in the OLS, it is hard to tell their relative impact on the number of foreclosures. Variables such as the denial rate and the number of recent immigrants even switched signs across models and thus seem to have an ambiguous effect on foreclosure.

To test the robustness of the results, I estimate a similar regression using the foreclosure rate as a dependent variable in Table 7. The effect of percent subprime owner-occupied is still positive and significant, but it is no longer the largest in magnitude. In fact, the percent black has the largest effect, since a one standard deviation increase in percent black increases the number of foreclosures per 100 units by 1.7. The effect of college graduates is fairly large as well; a one standard deviation increase in percent with bachelors decreases the number of foreclosures per 100 units by one. Similar to the foreclosure levels OLS and negative binomial regressions, the median year built and the vacancy rate have a positive effect the foreclosure rate.

Table 7: OLS with foreclosures per 100 owner-occupied units as a dependent, 2006

	Price change for 2003-2006			No price change		
	Coef	SE	Beta	Coef	SE	Beta
Loan Level data						
Subprime owner %	3.344 *	1.006	0.159	3.133 *	1.003	0.149
Subprime not owner %	0.386	0.516	0.021	0.205	0.514	0.011
Denial rate	1.149	1.225	0.034	0.868	1.228	0.026
Demographic tract level						
Population total	-0.330 *	0.114	-0.203	-0.348 *	0.114	-0.214
Recent immigrants	0.607	0.551	0.043	0.893	0.549	0.064
Bachelors %	-7.754 *	1.971	-0.218	-7.482 *	1.976	-0.211
Public assistance %	4.380	2.359	0.076	4.149	2.362	0.072
65 and older %	-5.270 *	1.950	-0.075	-6.233 *	1.924	-0.089
Single %	-2.662	1.452	-0.125	-2.928	1.443	-0.138
Foreign %	6.535 *	1.941	0.240	5.725 *	1.924	0.211
Speak English poorly %	-25.119 *	3.357	-0.501	-25.057 *	3.366	-0.500
Black population %	3.665 *	0.744	0.316	3.522 *	0.745	0.304
Hispanic population %	4.162 *	1.111	0.227	4.248 *	1.115	0.232
Asian population %	0.274	2.037	0.004	0.085	2.043	0.001
Economic tract level data						
Median household income	0.129	0.118	0.074	0.148	0.114	0.085
Time to work aggregate	0.114	0.070	0.114	0.115	0.070	0.114
Unemployment rate	-1.356	1.798	-0.028	-1.858	1.800	-0.038
Renter burden	0.008	0.016	0.012	0.011	0.016	0.016
Owner burden	-0.006	0.015	-0.009	-0.010	0.015	-0.014
Vacancy rate	14.842 *	2.781	0.143	15.068 *	2.789	0.145
Management %	1.789	1.432	0.070	2.037	1.434	0.079
Service %	-0.939	2.051	-0.016	-0.423	2.055	-0.007
Housing tract level data						
Median value owner occupied	-0.346	0.202	-0.080	-0.278	0.191	-0.064
Median year built	0.022 *	0.008	0.078	0.030 *	0.008	0.107
Homeownership %	-0.280	0.874	-0.017	-0.644	0.872	-0.039
Median Sale Price Change						
Median Price Change	-2.715 *	0.584	-0.119			
Median Sale Price in 2006	0.103	0.108	0.031			
Constant	-40.8212	15.954		-56.7774	15.458	
Observations	1711			1718		
Adj. R squared:	0.28			0.27		

Note: divided population and recent immigrants by 1,000, divided median household income and time to work aggregate by 10,000. divided median value owner-occupied, median sale price in 2006 by 100,000
* indicates significance at the 0.05 level.

The effect of the Hispanic share on foreclosure rates is positive, which contradicts the finding in the negative binomial. Since the coefficients on percent foreign, speak English poorly and single switch across models it is impossible to conclude anything about their effect on foreclosures.

Foreclosures and House Price Appreciation

Even though coefficients on many neighborhood characteristics switch signs across models, the effect of the median 2003-2006 house price change on foreclosure remains the same across all models. In the levels regression, an increase of 20 percent in the price change results in a decrease of 1.75 foreclosures, all else equal. In the rates regression, an increase of 20 percent in the price change reduces foreclosures per 100 units by 0.5. It is curious to see if the price change during different years affected foreclosures differently. Table 8 shows the beta values of the price change variable during different sets of years from OLS regressions with foreclosure levels or foreclosure rates as dependent. The effect of the price change on either the foreclosure number or rate has the highest magnitude during 2003-2006 and is always either negative or not significant. Thus, I conclude that tracts with higher housing price appreciation have lower foreclosure rates, which is in line with what other literature has suggested or found.

Table 8: Foreclosure level and rate by price change, 2006

	Foreclosure measure	
	Levels	Rates
Price Change Set of Years	Beta	Beta
2003-2006	-0.073 *	-0.119 *
2005-2006	-0.009	-0.091 *
2004-2006	-0.072 *	-0.147 *
2003-2004	-0.008	0.008
2003-2005	-0.058 *	-0.041
2004-2005	-0.061 *	-0.063 *

Note: * means significant at the 0.05 level

Moreover, coefficients on other explanatory variables remain virtually unchanged once the price change is added in. In fact, coefficients on explanatory variables remained unchanged for any

combination of years from 2003 through 2006. This indicates that the neighborhood characteristics affect the foreclosure rate independently from house price appreciation.

Determinants of Subprime Lending

It is useful to study the relationship between subprime lending and neighborhood characteristics, since it is significant across all model specifications and has a large positive effect on foreclosures. I fit the following tract-level model to the percent subprime owner-occupied variable.

Percent subprime owner-occupied = α + β_1*Foreclosure rate + β_2*LOAN LEVEL + β_3*DEMOGRAPHIC + β_4*ECONOMIC + β_5*HOUSING + β6*SALE PRICE CHANGE

Table 9 shows the results from the subprime rate regression. Most of the explanatory variables are significant in explaining the variation in the subprime rate. The percent black has the biggest effect on subprime lending, where a one standard deviation increase in the percent black results in a half a standard deviation increase in the subprime rate (9 percent). The percent Hispanic also has a positive effect on the subprime rate, but it is about four times smaller in magnitude than that of the percent black.

The denial rate has the second largest effect on the subprime rate after the percent black, where a one standard deviation increase in the denial rate increases the subprime rate by a third of a standard deviation. This relationship makes sense because the denial rate is a proxy for the credit risk in a neighborhood.

Even though small, the public assistance and the unemployment rate have a negative effect on subprime lending. It is also surprising that the median household income has a positive effect

on subprime lending. As might be expected, the median house value in 2006 and 2000, the percent with bachelor's degrees and the percent in management positions have a negative impact on subprime lending. More expensive, higher educated and more occupationally professional

Table 9: OLS with subprime percent owner as a dependent, 2006

	Price change for 2005-2006			No price change		
	Coef	SE	Beta	Coef	SE	Beta
Loan Level data						
Foreclosure rate	0.215 *	0.059	0.045	0.191 *	0.001	0.040
Denial rate	0.511 *	0.027	0.321	0.530 *	0.027	0.333
Demographic tract level						
Population total	0.114 *	0.028	0.147	0.115 *	0.028	0.149
Recent immigrants	0.235	0.133	0.035	0.217	0.134	0.032
Bachelors %	-0.132 *	0.048	-0.078	-0.149 *	0.048	-0.088
Public assistance %	-0.148 *	0.057	-0.054	-0.175 *	0.058	-0.064
65 and older %	0.021	0.047	0.006	-0.003	0.047	-0.001
Single %	0.001	0.035	0.001	-0.015	0.035	-0.015
Foreign %	-0.061	0.047	-0.048	-0.083	0.047	-0.064
English bad %	-0.116	0.083	-0.049	-0.104	0.084	-0.044
Black population %	0.251 *	0.017	0.457	0.262 *	0.017	0.477
Hispanic population %	0.129 *	0.027	0.148	0.128 *	0.027	0.147
Asian population %	-0.034	0.049	-0.011	-0.028	0.050	-0.009
Economic tract level data						
Median hh income	0.117 *	0.029	0.142	0.089 *	0.028	0.108
Time to work aggregate	-0.055 *	0.017	-0.116	-0.053 *	0.017	-0.112
Unemployment rate	-0.117 *	0.043	-0.051	-0.134 *	0.044	-0.058
Renter burden	0.157 *	0.037	0.049	0.150 *	0.038	0.046
Owner burden	0.142 *	0.037	0.042	0.135 *	0.037	0.040
Vacancy rate	-0.097	0.068	-0.020	-0.096	0.069	-0.020
Management %	-0.243 *	0.034	-0.199	-0.249 *	0.035	-0.204
Service %	-0.036	0.050	-0.013	-0.042	0.050	-0.015
Housing tract level data						
Median value 2000	-0.025 *	0.005	-0.122	-0.036 *	0.005	-0.175
Median year built	-0.632 *	0.193	-0.047	-0.633 *	0.192	-0.047
Homeownership %	0.044 *	0.021	0.056	0.057 *	0.021	0.073
Median Sale Price Change						
Median Price Change	0.111 *	0.021	0.064			
Median Sale Price in 2006	0.000 *	0.000	-0.104			
Constant	1.456	0.379		1.462	0.378	
Observations:	1712			1718		
Adj. R squared:	0.81			0.81		

Note: Divided population total by 10,000, divided foreclosure rate by 100, divided median household income, time to work aggregate, median value owner occupied by 100,000, divided renter burden and owner burden by 100, divided median year built by 1,000

* indicates significance at the 0.05 level.

neighborhoods can be expected to have lower subprime lending rates. Contrasting with its positive effect on foreclosures, the median year built has a negative impact on subprime lending.

The median price change from 2005-2006 has a positive effect on subprime lending, all else equal. This supports the hypothesis that in appreciating areas lenders have looser lending standards and thus grant more subprime loans. I observe a negative effect for three year range specifications and the effect is the most prominent for the price change from 2005 until 2006. The effect is the smallest for the period 2004-2006, while it is higher for the period 2004-2006. This result suggests that subprime lending is most affected by the most recent house price changes. Conversely, the foreclosure rate is more affected by house price changes over longer periods of time. Similarly to modeling foreclosure rates, the addition of any of the house price change variables does not affect coefficients on the rest of the explanatory variables. This indicates that these neighborhood characteristics affect the subprime rate independently of the house price changes.

CONCLUSION AND DISCUSSION

This study enriches the current understanding of the foreclosure phenomenon, by bringing to light evidence to support and extend the results from previous literature. New relationships uncovered in this study substantiate the theories developed in previous studies and provide a much-needed contribution to the field.

Analysis of the spatial distribution of foreclosures shows that defaults are concentrated in "hot spots" in south Chicago and in south Cook County. This indicates that policy makers should pay particular attention to these areas to mitigate the effect of foreclosures. Alarmingly, a third of foreclosures in 2006 came from loans originated in 2005 or in 2006. In fact, 86 percent of foreclosures in 2006 resulted from loans originated in 2000 onwards. This enormous amount of quick foreclosures should be a topic of huge concern. Unlike the findings from Duda et al. (2005), quick foreclosures in Chicago were not concentrated in the high foreclosure areas. However, to

draw any definitive conclusion about this issue, further research on the determinants of quick foreclosures is needed.

Subprime lending consistently has a positive and significant effect on foreclosure across all models. The effect is the largest in the levels model, where a one standard deviation increase in the number of owner-occupied subprime loans results in an increase of 19 in the number of foreclosures in a tract. Even though the effect of subprime lending is smaller in the rates model, these results present a serious cause for concern over the benefits of subprime lending. The fact that the high subprime tracts almost perfectly predict the high foreclosure tracts in the spatial analysis is also quite worrying. These results support evidence from other studies that the risk of subprime lending is much higher than expected. The benefits of granting subprime loans appear to be wiped out by the devastating impact of the unacceptable amount of foreclosures.

As for demographic characteristics, the black share is the most important determinant of foreclosures after subprime lending. In fact, a one standard deviation increase in the black population share increases the number of foreclosures by five and the rate of foreclosure by 1.7. The vacancy rate and the median year built have consistently small but positive effects on foreclosure. It is surprising that these housing characteristics and not incomes have a significant impact on foreclosures. The share of college graduates is also an important determinant of foreclosures, where a one standard deviation increase in it decreases the number of foreclosures per 100 units by one. The percent of population who is 65 and over has a consistently small but negative effect on foreclosures. So, areas with higher concentrations of older people experience fewer foreclosures in Chicago. Contrasting with results from other studies, the Hispanics were not concentrated in the high foreclosure areas. Moreover, since coefficients on the Hispanic share change across models, the effect of this variable on foreclosures is ambiguous. The rest of the neighborhood characteristics either have no effect or have an ambiguous effect on foreclosures.

Similarly to its impact on foreclosures, the black share has the highest impact on subprime lending. However, the share of college graduates is the only other variable that affects subprime lending in a similar way to foreclosures. This indicates that factors that affect foreclosures and subprime lending are different in nature. This could be because the decision to take out a subprime loan is different from the decision to default.

As expected, the median price change has a negative impact on foreclosure and a positive impact on subprime lending, all else equal. It appears that lenders grant more subprime loans in areas where house prices grow at a higher rate. Even if a borrower has a weak financial history, the lender expects him to pay off the loan, since his house appreciated and he can either sell the house or refinance it. Thus, as house prices grow the amount of irresponsible subprime lending increases. However, as house prices grow at a faster rate, the foreclosure rate decreases. At first glance, it seems that a large subprime volume is beneficial and actually results in fewer foreclosures. But the true consequences of irresponsible lending hit communities when house prices start falling. The borrowers the most affected by this phenomenon are the ones who received subprime loans that they were unable to afford. This finding supports the story of California, where an enormous amount of subprime lending was reassuringly accompanied by low foreclosure rates while house prices were rising rapidly until 2006. However, as the prices dropped, California became a state with one of the highest growths in foreclosures, as well as with one of the highest foreclosure rates.

REFERENCES

Apgar, William C., Duda, Mark, Gorey, Rochelle Nawrocki, "The Municipal Cost of Foreclosures: A Chicago Case Study." *Homeownership Preservation Foundation Housing Finance Policy Research Paper Number 2005-1* (February 2005)

Avery, Robert B., Canner, Glenn B. and Cook, Robert E. 2005. "New Information Reported under HMDA and Its Application in Fair Lending Enforcement." *Federal Reserve Bulletin* 91(3): 344-394

Bernanke, Ben S., "Mortgage Delinquencies and Foreclosures." Speech at the Columbia Business School's 32nd Annual Dinner (May 5, 2008) (http://www.federalreserve.gov/newsevents/speech/Bernanke20080505a.htm)

Bunce, Harold L., Gruenstein, Debbie, Herbert, Christopher E., Scheessele, Randall M., "Subprime Foreclosures: The Smoking Gun of Predatory Lending?" U.S. Department of Housing and Urban Developmentn (February, 2001)

Courchane, Marsha J., Surette, Brian J., Zorn, Peter M., "Subprime Borrowers: Mortgage Transitions and Outcomes." *Journal of Real Estate Finance and Economics*, 29:4, 365-392 (December, 2004)

Duda, Mark, Apgar, William C., "Mortgage Foreclosures in Atlanta: Patterns and Policy Issues." NeighborWorks America (December, 2005)

Gerardi, Kristopher, Shapiro, Adam Hale, Willen, Paul S. "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures." *Federal Reserve Bank of Boston Working Papers* (May 2008)

Grover, Michael, Smith, Laura, Todd, Richard M., "Targeting Foreclosure Interventions: An Analysis of Neighborhood Characteristics Associated with High Foreclosure Rates in Two Minnesota Counties." *Federal Reserve Bank of Minneapolis Community Affairs Report* (January, 2007)

Huffman, Mark, "May Foreclosure Filing Rate Highest Ever" ConsumerAffairs.com (June 15, 2008) (http://www.consumeraffairs.com/news04/2008/06/foreclosures_may.html)

Immergluck, Dan, Smith, Geoff, "Risky Business – An Econometric Analysis of the Relationship Between Subprime Lending and Neighborhood Foreclosures." *Woodstock Institute* (March, 2004)

Immergluck, Dan, Smith, Geoff. 2006. "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values," *Housing Policy Debate*, Volume 17:1, 57-80

Immergluck, Dan, Smith, Geoff. 2006. "The Impact of Single-family Mortgage Foreclosures on Neighborhood Crime." *Housing Studies*, 21:6, 851-866

Immergluck, Dan. 2008. "From the Subprime to the Exotic: Excessive Mortgage Market Risk and Implications for Metropolitan Communities and Neighborhoods." *Journal of the American Planning Association*, 74, 59-76

Lanzerotti, Laura, "Homeownership at High Cost: Foreclosure Risk and High Cost Loans in California." *Federal Reserve Bank of San Francisco Working Paper* (July, 2006)

Mayer, Chris, Pence, Karen, "Subprime Mortgages: What, Where, and to Whom?" *Lincoln Land Institute Conference Honoring Chip Case* (December, 2007)

Mueller, Elizabeth, "A Study of Residential Foreclosures in Texas." *Texas Department of Housing and Community Affairs* (September, 2006)

Perkins, Kristin Laurel. 2008. "The Geography of Foreclosure in Contra Costa County, California." Thesis for the Master of City Planning in University of California Berkeley

Quercia, Roberto, Cowan, Spencer M., Moreno, Ana B, "The Cost-Effectiveness of Counseling in Community-Based Foreclosure Prevention." (December, 2005)

Schloemer, Ellen, Li, Wei, Ernst, Keith, Keest, Kathleen, "Losing Ground: Foreclosures in the Subprime Market and Their Cost to Homeowners." *Center for Responsible Lending* (December, 2006)

Wardrip, Keith E., Pelletiere, Danilo, "Neighborhood Poverty and Tenure Characteristics and the Incidence of Foreclosure in New England." *National Low Income Housing Coalition* (June, 2008)

APPENDIX

Foreclosure data cleaning

For each foreclosure filing there was information such as the address of the property, original mortgage date, property type, original mortgage amount, type of mortgage, interest rate, etc. I had to geocode this data from x and y coordinates into census tracts. To accomplish this I used the spatial joining tool in ArcMap in GIS. About 100 foreclosure observations did not have data on x and y coordinates, so I dropped those from the analysis. Thirteen of the observations got coded to census tracts from the wrong counties, so I had to drop them. Some properties appeared more than once in the data, because they had multiple loans from different lenders, or because they appeared first as newly filed foreclosures and later as scheduled for auction. Since I am interested in neighborhood distress and not in individual loans, I deleted 7735 duplicate observations based on the address (street and city). For each property the data had information on property type, whether the building was a single family home, apartment building, condominium, town home or a commercial building. Since, I am interested in outcomes of individual homeowners, that is

foreclosures in owner-occupied residencies, I only kept single-family homes, town homes and condominiums. So, I deleted 5509 properties that were either apartment or commercial buildings. In the end I was left with 33,053 observations. Finally, I aggregated my default data to census tracts and was able to compute a default rate by dividing the number of defaults in a census tract by the number of owner-occupied units. About 40 tracts had no owner-occupied units, so I deleted them from the analysis. These tracts are white in the spatial analysis.

House Price data cleaning

My main job was to assign each tract to a particular municipality. The website where I obtained the data provided a list of census tracts for each municipality, so I had to manually extract this information for each municipality. A tract can belong to multiple municipalities, as it can intersect a couple of them. Since I needed to assign each tract a median housing sale price, if a tract belonged to more than one municipality I calculated the average price of all municipalities it intersected. Table 2 provides a summary of how many tracts belonged to multiple municipalities. Most of the tracts, about 77%, belonged to only 1 municipality, which indicates that averaging out municipality prices shouldn't be too bad.

Number of Municipalities	Number of Tracts	Percent
1	1349	76.6%
2	259	14.7%
3	119	6.8%
4	22	1.2%
5	7	0.4%
6	4	0.2%
9	2	0.1%
Total	**1762**	**100%**