Risk or Race:

AN ASSESSMENT OF SUBPRIME LENDING PATTERNS IN NINE METROPOLITAN AREAS

U.S. Department of Housing and Urban Development | Office of Policy Development and Research

Visit PD&R's website

www.hud.gov/policy or www.huduser.org

to find this report and others sponsored by HUD's Office of Policy Development and Research (PD&R). Other services of HUD USER, PD&R's Research and Information Service include listservs, special interest and bimonthly publications (best practices, significant studies from other sources), access to public use databases, and a hotline (1-800-245-2691) for help with accessing the information you need.

Risk or Race:

An Assessment of Subprime Lending Patterns in Nine Metropolitan Areas

Prepared for:
U.S. Department of Housing
and Urban Development
Washington, D.C.

Prepared by:
William C. Apgar Jr.
Christopher E. Herbert
Priti Mathur

August 2009

Foreword

This report, *Risk or Race: An Assessment of Subprime Lending Patterns in Nine Metropolitan Areas,* sheds light on important issues in subprime lending. The report brings together data on lending patterns at the borrower and neighborhood level, the race of borrowers, other economic and demographic characteristics, and the critically important neighborhood credit measures based on newly available borrower credit data.

This past decade has seen major upheaval in the nation's housing and mortgage markets that has had serious consequences for the economy as a whole. Over this period, which saw rapid growth in subprime lending, there was a growing concern that many households that qualified for lower-priced prime loans were inappropriately steered into higher-priced subprime loans.

Indeed, numerous studies over the period raised concerns that subprime lenders were inappropriately targeting economically vulnerable neighborhoods and that some borrowers, particularly minority and/or lower-income borrowers, were paying more than necessary for mortgage financing. Part of this substantial body of literature indicates that neighborhoods that were predominantly African-American had much higher shares of loans originated by subprime lenders than areas where whites predominated.

Despite such evidence such as this, some researchers and industry commentators pointed to the lack of detailed data on credit scores to either corroborate or contradict the evidence on the race of borrowers as a driving factor behind inappropriate subprime lending. This report, using detailed census tract-level data on borrower credit risk, sheds new light on this "risk or race" question. For the first time, lenders were required in 2004 to disclose pricing information for relatively higher-priced loans under the Home Mortgage Disclosure Act (HMDA). The study utilizes census tract-level data on mortgage credit risk (obtained from a credit reporting agency) and loan-level data submitted by financial institutions under HMDA. The analysis employs a variety of models that use alternative definitions of subprime loans and alternative credit risk measures and deploys the best available econometric techniques.

This study presents results from a neighborhood-level analysis of subprime lending in nine metropolitan areas, using two alternative measures of subprime mortgages: the first based on the HUD-developed list of subprime lenders and the second based on the more recently available information on whether or not a loan is "higher-priced"; namely, has a mortgage rate in excess of the established benchmark. It also presents an assessment of the relationship between neighborhood credit scores and mortgage denial rates.

The study finds that overall, the inclusion of neighborhood credit measures did not explain away the troubling finding that race and ethnicity remain an important determinant of the allocation of

mortgage credit. Even after accounting for the impact in the spatial variation in credit scores, the African-American share of the census tract population still had a positive and significant impact on the spatial variation in the tract-level share of higher-priced subprime refinance loans in five of the nine metropolitan areas examined.

Before the inclusion of credit score information, the results showed a very strong association between African-American borrowers and the probability of obtaining a higher-priced loan. In all nine metropolitan areas, for both purchase and refinance loans, when the neighborhood credit measure variables were added, there was very little change in either the significance or the magnitude of the borrower's race or ethnicity.

The findings thus generally support the results found in previous studies. These studies concluded that even after accounting for some neighborhood credit measures along with a variety of borrower and neighborhood characteristics, significant racial and ethnic disparities remained in the use of subprime mortgage lending. This is also important, as the analysis is to reassess the past research. Although it refined the previous findings, it does not refute them; in general, it provides support for the past research and proposes how it can still be effectively used (and expanded on).

Significantly, this study—focusing on a very different set of markets compared to earlier studies—still came to very similar conclusions regarding the existence of racial and ethnic disparities in subprime lending, even after controlling for neighborhood credit scores.

The study also includes an important review of the relevant literature that will serve as a valuable additional aid for researchers and policymakers.

It is important to note that since the study is unable to fully account for borrower-level risk measures, the results, by themselves, cannot be taken as evidence of discriminatory treatment at the individual level. Yet the results are consistent with the possibility that racial and ethnic minorities have been subject to discriminatory treatment. These findings will be important to aid not only efforts to ensure that all households have access to mortgage capital on a fair basis, but also efforts to prevent the excesses in mortgage lending from being repeated in the future.

Raphael W. Bostic
Assistant Secretary for Policy
Development and Research

Table of Contents

Executive Summary

Over the past decade, numerous studies have raised concerns that subprime lenders were inappropriately targeting economically vulnerable neighborhoods and that some borrowers, particularly African-Americans, were paying more than necessary for mortgage financing. Indeed, there is a substantial body of literature consistent with the finding that predominantly African-American neighborhoods had much higher shares of loans originated by subprime lenders than areas where whites predominated, while even controlling for a range of factors that are likely to influence the allocation of mortgage credit. Yet, despite the importance of the topic, efforts to explain the underlying causes of the observed patterns have been limited by the lack of data on mortgage credit risk.

Using newly available data on neighborhood-level measures of credit scores for each of nine metropolitan areas selected to reflect a range of racial and ethnic population characteristics, this paper presents a detailed analysis of what has come to be known as the "risk or race" question. In particular, this report summarizes the results obtained from a series of weighted least squares regressions for models that seek to explain census tract-level variation in the share of higher-priced subprime mortgages for home purchase and refinance. In addition, the paper also reports the findings obtained from a series of logit regression models of the impact that tract-level mortgage credit scores have on the probability that an individual borrower will obtain a higher-priced as opposed to a lower-priced prime mortgage. Following the path-breaking work of Susan Wachter and her colleagues, these models included various measures expected to influence the allocation of mortgage credit, including income and other demographic factors, mortgage credit and other measures of credit risk, and various neighborhood-level housing market variables. For both sets of equations, the key focus is on whether and how the inclusion of tract-level credit score information influences the relationship between racial factors and access to prime mortgages.

In addition, the paper explores several additional methodological issues. First, until the recent release of 2004 Home Mortgage Disclosure Act (HMDA) data on loan-level mortgage pricing, numerous studies examined the spatial variation of the share of subprime loans based on a HUD generated list of institutions that engaged primarily in subprime lending. Moreover, prior to recent enhanced availability of census tract level measures of mortgage credit scores, numerous studies were forced to deploy differing proxy measures for risk factors. Yet even as newer studies now can utilize better measures of subprime loans and risk factors, there is still value in relating the results obtained from these newer studies to previous research. For example, to the extent that the relationship between measures of tract-level patterns of subprime lending based on the HUD-list approach and the newly available measures based on the HMDA concept of "higher-priced" mortgages can be established, there will be much greater opportunity for research on historical patterns of the neighborhood-level increases in the overall growth of subprime lending.

Finally, this study also explores the latest and most advanced econometric techniques to assess the extent to which failure to account for spatial autocorrelation may have biased the estimates of the "race effect" in previous studies. Results presented here suggested for traditional least squares and logit regression models based on aggregate data, spatial autocorrelation does not appear to bias the estimates of coefficients, nor have any important effect on the efficiency of the estimates. Even so, in light of growing evidence that individual subprime lenders may have spatially targeted specific lower-income and/or minority borrowers for the aggressive marketing of subprime mortgages, further research on this topic seems warranted.

Given the substantial variation across metropolitan areas in racial/ethnic settlement patterns and other characteristics likely to influence the spatial distribution of mortgage credit, understandably the estimates of the coefficients included in any single least squares or logit specification are likely to vary from one metropolitan area to another. It is therefore most significant that the estimates of the coefficients relating to the influence of both tract-level and borrower-level race/ethnicity on access to higher-priced (as opposed to lower-priced) mortgages were remarkably consistent across individual market areas. In general, the inclusion of credit score measures did not explain away the troubling finding that even after years of public policy efforts, race and ethnicity remain an important determinant of the allocation of mortgage credit in both home purchase and home refinance markets. While the results generally find a stronger association between African-American borrowers and neighborhoods and higher-priced lending, Hispanic borrowers and neighborhoods were also often found to have a positive association with higher-priced lending, although the associations were smaller and had lower levels of statistical significance.

Overall, these findings confirm that the omission of mortgage credit score information does in fact bias estimates of the effect of race and ethnicity on access to mortgage credit, particularly lower-priced subprime loans. Still, a racial and ethnic association with higher-priced lending remains, even after entering a neighborhood credit metric and controlling for other neighborhood-level risk factors such as household incomes, housing prices, rental shares, share of the population with a college education, and housing turnover rate. Borrower-level models further control for borrower income and gender. In general, larger impacts are found in the refinance market and are associated with the African-American population share.

Although the relatively small size of the coefficients on race and ethnicity does suggest that the magnitude of the racial discrimination may be waning, it is no cause for celebration. This is especially true in light of the recent growth in delinquency and foreclosure rates in lower-income and/or minority neighborhoods. While the mortgage market meltdown is adversely impacting both the prime and subprime markets, the Mortgage Bankers Association National Delinquency Survey indicates that over the last 10 years the probability that a subprime mortgage will enter into foreclosure has been nearly 10 times more likely than a prime

mortgage. As a result, to the extent that mortgage market discrimination may have added—even on the margin—to the concentration of higher-priced subprime loans in minority neighborhoods, the consequences of this discriminatory pattern are coming home to roost in the form of elevated foreclosures undermining the stability and vitality of these areas today.

In an effort to create a bridge of understanding between older and newer studies on the topic, this report presents results from models of the neighborhood variation in subprime lending using two alternative measures of subprime mortgages: the first based on the HUD developed list of subprime lenders, and the second based on the more recently available HMDA information on whether or not a loan is "higher-priced"; namely has a mortgage rate in excess of the established benchmark. The results confirmed the somewhat obvious conclusion that the new data represent a superior measure of subprime lending. This is understandable since the HMDA-based measure reflected loan-level pricing information. In contrast, since most lenders offer a range of both prime and subprime products, the HUD measure of subprime lending combines two types of measurement errors. First, it erroneously classifies as subprime those prime loans made by institutions that on average specialize in subprime lending. In addition, it erroneously classifies subprime loans as prime loans when made by entities that predominantly engage in prime lending, but can and often do account for a large share of subprime activity in any single market area.

Despite these issues, and in tribute to the ingenuity of the creators of the HUD approach, the equations based on the HUD list mirror those based on the more refined measures of the newly available HMDA price information. As a result, the findings suggest that while historical studies based on the HUD approach are not precisely comparable to newer studies, assessing older and newer studies can help place more recent trends into an appropriate historical context.

Similarly, the results clearly demonstrate that tract-level mortgage credit scores offer a superior measure of tract-level mortgage risk factors when compared with measures based on mortgage rejection rates—a measure of credit risk that has been used by previous studies that did not have access to neighborhood- or borrower-level credit scores. Since credit score measures are not widely available and/or are costly to acquire, the results presented here affirm that risk measures based on mortgage denial rates still can play a valuable role in helping neighborhood housing and advocacy organizations monitor recent trends and place these trends into a longer-term historical context.

Finally, the results relating to spatial autocorrelation merit continued scrutiny, especially as new econometric techniques are developed and made widely available. In the equations presented in this report, concerns about the potential for bias resulting from spatial autocorrelation do not appear to be warranted. Detailed examination of this issue found that in about half of the cases, spatial autocorrelation was a potential issue. But further

examination suggested that failure to correct for serial autocorrelation did not bias coefficient estimates to any significant degree.

Given the historically contentious nature of the assessment of any matter involving racial discrimination, the analysis presented here will not quiet the debate on the extent to which racial disparities in the allocation of subprime lending contributed to the concentration of foreclosure-prone subprime loans in lower-income and minority neighborhoods. The variety of models utilizing alternative definitions of subprime loans and alternative credit risk measures, as well as the best available econometric techniques for assessing potential spatial autocorrelation bias, were employed in the hope that the research presented in this study will help cast a new light on existing research and contribute to the ongoing national debate on how best to ensure that all households have access to mortgage capital on a fair basis, but free of the excesses of the recent past.

Section 1: Introduction

1.1 Background

One of the most notable aspects of the residential mortgage market over past decade was the rapid growth of subprime lending and the subsequent mortgage market meltdown as subprime loans began to fail in record numbers. By one measure, subprime lending increased between 1994 and 2006 by nearly 20-fold, from $35 billion to $660 billion, before pulling back sharply in 2007.[1] Initially, the rise of subprime lending expanded home-buying opportunities for minority and lower-income households, especially for those households that otherwise were unable to obtain conventional prime mortgage financing. Yet the aggressive marketing practices of many subprime lenders apparently encouraged some borrowers to take on mortgage loans that they neither fully understood nor were able to repay. As early as 2004, the number of subprime mortgages that were seriously delinquent was on the rise. The resulting surge in foreclosures in many of the nation's low-income and minority neighborhoods is undoubtedly a direct result of the concentration of higher-risk and hence more default-prone subprime loans in these areas.[2]

The ongoing mortgage crisis raises questions about whether mortgage market regulations are sufficient to protect consumers from the aggressive marketing practices of subprime lenders, and whether the concentration of subprime loans in minority communities may have violated "fair lending" standards. Over the past decade, numerous studies raised concerns that subprime lenders were inappropriately targeting economically vulnerable neighborhoods and that some borrowers, particularly minority and/or lower-income borrowers, were paying more than necessary for mortgage financing. Indeed, there is a substantial body of literature that reported that predominantly African-American neighborhoods had much higher shares of loans originated by subprime lenders than areas where whites predominated.

While most analysts acknowledge the illegal actions on the part of some mortgage lenders, many argue that the largest share of the observed differences in mortgage lending outcomes across racial/ethnic lines simply reflected systematic differences in borrower risk and other objective factors that legitimately should influence mortgage pricing and terms. Despite the importance of the topic in fully assessing the underlying causes of the current mortgage crisis, efforts to answer this "risk or race" question have been limited by lack of data on credit risk. While this information is ubiquitous in the lending industry, until recently it has not generally been available to researchers due to the proprietary nature of these data. As a result, even as federal officials debate how to enhance the regulation of mortgage lending to

---

1 Estimates from *Inside Mortgage Finance*.

2 See Joint Center for Housing Studies (2008), *America's Rental Housing: The Key to a Balanced National Policy*.

prevent the current mortgage market debacle from recurring, there remain doubts about the extent (if any) to which that allocation of subprime loans reflected in part the inappropriate use of racial/ethnic characteristics in mortgage lending. Using detailed census tract-level data on borrower credit risk, this study seeks to expand understanding of these important fair lending issues and, in doing so, help guide the ongoing debate as to how best to ensure that mortgage capital is made available on fair and efficient terms in the future.

1.2 Contribution of Proposed Research

The quantitative analyses presented here utilizes newly available census tract-level data on mortgage credit risk obtained for this study from a credit reporting agency and loan-level data submitted by financial institutions under the Home Mortgage Disclosure Act (HMDA) of 1975. As currently amended, HMDA requires mortgage lenders to report information about loan applicant race and income and the geographic location of the property included in the application. For the first time, in 2004 lenders were required to disclose pricing information for loans with APR above a designated threshold. For first-lien mortgages, this threshold was 3 percentage points higher than the rate charged on a Treasury Security of comparable maturity. These loans are characterized as "rate-spread" mortgages or "higher-priced" mortgages. Absent this pricing data, and given the reluctance of mortgage lending institutions to release what they claimed was proprietary information on mortgage credit scores, most previous HMDA based studies of subprime lending were forced to identify subprime loans using a HUD-developed methodology that identified all loans made by lenders known to make primarily subprime loans.

The research presented in this report will build on the existing literature in several ways. Existing work has already confirmed that neighborhood credit measures—most notably information based on credit score distributions but also conventional mortgage denial rates—are significant predictors of subprime lending shares. This study will present results from models of the neighborhood variation in subprime lending using two alternative measures of subprime mortgages: the first based on the HUD-developed list of subprime lenders and the second based on the more recently available information on whether or not a loan is "higher-priced"; namely has a mortgage rate in excess of the established benchmark. While the vast majority of research conducted since 2004 uses HMDA data on "higher-priced" mortgages, efforts to place recent findings into an appropriate historical context require a detailed understanding of the relationship between these two distinctly different approaches to measurement of the spatial pattern of subprime data.

Next, the study will present an assessment of the relationship between neighborhood credit scores and the mortgage denial rate. Prior to the more widespread availability of mortgage credit scores, previous studies were forced to utilize proxies that attempted to substitute for the omitted credit score variable. For example, the mortgage denial rate was included as a proxy for neighborhood credit risk in several studies and was found to be a significant

predictor of subprime lender shares.[3] Since this measure is derived from the HMDA data, it has the significant advantage of being readily available for virtually all markets in the country and for each year beginning in 1993. If a strong link can be identified between HMDA denial rates and neighborhood credit scores, there will be much greater opportunity for research on historical patterns of neighborhood variation in subprime lending (e.g., studies that extend back before the widespread availability of detailed credit scores) to incorporate at least a rough proxy for neighborhood credit risk.

The research presented here will build on an earlier study of the neighborhood variation of subprime lending (Calem, Gillen, and Wachter, 2004) that incorporated both the HMDA denial rate and a measure of tract credit scores. In addition to expanding the analysis to additional market areas, this report will also assess the performance of the denial rate proxy in the context of models that use the HMDA "higher-priced" loan definition of subprime lending, as well as the definition based on the HUD subprime lender list that was used in the initial Calem, Gillen, and Wachter analysis.

Another key contribution of this study will be to investigate the importance of incorporating econometric techniques that test and correct for spatial correlation of variables. One of the arguments about the preponderance of subprime lending in minority communities is that subprime lenders are more effective at marketing to these communities. There may also be "word of mouth" networks that help to market specific lenders among social networks that are geographically based. For these reasons, subprime lender shares may be spatially dependent—that is, their market shares may be higher in contiguous areas because of these marketing patterns, all else equal. Yet no research to date has explored the potential importance of this spatial dependence in modeling subprime market shares. Apgar and Herbert (2006) touch on these issues by testing for spatial correlation in estimated residuals, but they do not fully explore this issue. One important contribution of this study will be to evaluate the importance of employing econometric techniques that allow for such spatial dependence in subprime market shares.

1.3 Outline of the Study

Section 2 reviews the relevant literature. Section 3 describes the data and methodology. Section 4 presents the results of our analysis. Section 5 presents our conclusions. Appendix A presents detailed results of econometric models. Appendix B presents the spatial econometric tests and results.

[3] See Joint Center (2004).

Section 2: Summary of Previous Research

This section presents a brief literature review covering issues related to neighborhood disparities in subprime lending, neighborhood differences in credit and property risk that contribute to the observed subprime lending disparities, and the decision of prime mortgage lenders to enter specific neighborhoods. A comprehensive literature review covering essentially the same topics has recently been conducted (see Apgar and Herbert, 2006). This section will build on that review while focusing on the key questions to be addressed and the analytic approach to be used in this study.

2.1 Early Studies of Neighborhood Lending Patterns

Prior to 2004, readily available Home Mortgage Disclosure Act (HMDA) data did not include the loan pricing and the other loan-level information needed to distinguish a subprime loan from a prime loan. To facilitate examination of mortgage market trends, research staff at the U.S. Department of Housing and Urban Development developed a list of HMDA reporting lenders that primarily specialized in subprime lending.[4] Using this list of subprime lender specialists, numerous studies reported that neighborhoods that were predominantly occupied by African-Americans and/or lower-income households had much higher shares of loans originated by subprime lending specialists than areas where whites and/or higher-income households predominated. For example, in a comprehensive review of neighborhood lending patterns in Chicago in the late 1990s, Immergluck and Wiles (1999) observed that entities that specialized in providing prime loans differentially served higher-income white areas, while subprime lenders concentrated on lower-income and minority communities. Characterizing this as a "dual mortgage market," they concluded that the racial disparities were too great to be explained by differences in the credit quality of the borrowers. Instead, they argued that the observed patterns resulted from the failure of "mainstream lenders" to seek out creditworthy borrowers in lower-income and minority communities.

Numerous other studies supported the "dual mortgage market" hypothesis. Using 2000 HMDA data, Scheessele (2002) reported that while subprime lenders accounted for 16.4 percent of refinance loans in high-income neighborhoods (areas with incomes greater than 120 percent of area median), these lenders accounted for 36.3 percent of these loans in low-income areas (areas with incomes less than 80 percent of area median). The disparity in market share by neighborhood racial composition was even more striking. In 2000, subprime lenders accounted for 14.3 percent of refinance loans in neighborhoods where African-

[4] For description of the HUD subprime list see Scheessele (1999).

Americans comprised less than 30 percent of the population, but 47.8 percent in areas where they accounted for more than half of the population—more than three times as high a share.[5]

Using HMDA data for over 40,000 metropolitan area census tracts, the Joint Center for Housing Studies (2004) provided further documentation of what it called the "prime lending gap" in minority neighborhoods. This study found that the market share of loans made by subprime lending specialists rose over the 1993 to 2001 period in virtually all parts of the country, with particularly sharp increases in low-income and minority communities.

Finally, in a review of HUD-funded studies, Fishbein and Bunce (2000) concluded that a portion of borrowers whose credit would allow them to qualify for lower-cost conventional prime loans were nonetheless receiving subprime loans. They observed that the higher interest rates charged by subprime lenders could not be fully explained by neighborhood and/or borrower risk factors. Instead they concluded that a lack of competition from prime lenders had enabled subprime lenders to gain a growing share of mortgage lending activity in lower-income and minority communities and that racial disparity in lending patterns existed at both the individual borrower and neighborhood level.

Critics of these HMDA-based studies contended that observed lending patterns could simply reflect the uneven spatial distribution of credit-impaired borrowers. Unfortunately, until recently research on this topic was limited by a lack of readily available credit information. One early exception was the study by Avery et al. (1997) that examined the distribution of credit scores by neighborhood characteristics. Working with proprietary data purchased by the Federal Reserve Board of Governors, these researchers reported credit score information for all individuals residing in 994 ZIP codes representing a stratified random sample defined by region, metropolitan status (i.e., central city, suburban, or rural), and median household income. They found that the ZIP code areas with the lowest credit scores included areas with poverty rates over 25 percent, unemployment rates over 9 percent, minority shares of the population over 30 percent, less than 10 percent of the population over age 60, and were located in the East South Central census region. While this study did not examine the relationship between mortgage lending patterns and credit scores, it did show that low-income and minority areas do have generally lower median household credit scores.

Several other studies utilized proprietary information from specific lenders to assess the impact of risk factors on subprime lending. For example, Pennington-Cross, Yezer, and Nichols (2000) analyzed a unique dataset of home buyers that included a variety of risk measures not generally available at the time of the study, including credit history and non-housing debt levels. The study also included geographic area characteristics, although most

[5] Similar findings were reported in *Risk or Race? Racial Disparities in the Subprime Refinance Market* prepared by Calvin Bradford for the Center for Community Change, Washington, DC, May 2002. See also Acorn (2002), which reported on the 60 largest areas.

of these are at the metropolitan area level and not the neighborhood level. The authors concluded that subprime borrowers do, in fact, have higher risk characteristics than borrowers in the conventional or FHA market segments, indicating that the subprime market is appropriately targeted at high-risk borrowers. Nonetheless, they also found that African-American and Asian borrowers have higher probabilities of obtaining subprime financing even after controlling for the risk factors of credit score, debt levels, and income.[6]

2.2 The Growing Availability of Credit Score Information

While credit score information is ubiquitous in the lending industry, only recently has it become widely available for policy research and even today is only available with significant restrictions designed to protect the confidentiality of individual borrower information. Nonetheless, in the last few years there has been a series of studies that have managed to obtain credit data aggregated to the census tract and ZIP code level.

Calem, Gillen, and Wachter (2004) was the first study to incorporate census tract-level measures of credit risk into a regression analysis of subprime lending at both the tract and borrower levels.[7] The researchers obtained tract-level measures of the share of adults with credit scores in the bottom quintile of the credit score distribution in 1999 for the Chicago and Philadelphia metropolitan areas. Their models estimated tract-level subprime shares of purchase and refinance loans as well as the choice of subprime lenders at the borrower level. In addition to the credit score variable, their models incorporated other measures of credit risk, including the denial rate for conventional loan applications in the tract, the number of residential foreclosures as a share of owner-occupied units, and the capitalization rate for residential properties (proxied by the ratio of median rents to median house values).

They found that the measures of tract-level credit risk were associated in a statistically significant manner with subprime lending activity, as expected. In addition, they included as an explanatory variable in their model the share of adults with a college degree, arguing that it is a proxy for the degree of financial literacy, and found the predicted negative association with subprime lending. But, importantly, they also concluded that "even after inclusion of the full set of explanatory variables in the tract-level regressions, the percent of African-American homeowners is strongly, positively correlated with subprime share of neighborhood loans for both cities and both loan products."[8]

---

6 In a closely related study, Pennington-Cross and Nichols (2000) also analyzed average credit scores by a variety of borrower characteristics and concluded that credit scores are lower for lower-income borrowers, but the differences are not large.

7 Though published in 2004, this study first appeared in draft form in 2002. An earlier study by Pennington-Cross, Yezer, and Nichols (2000) incorporated borrower-level credit score measures in an analysis of mortgage choice including subprime as an option. However, it had a relatively small sample of subprime home purchase loans from 1996 and included few neighborhood-level controls.

8 Calem, Gillen, and Wachter (2004), p. 401.

In contrast to their findings for neighborhoods with relatively high shares of African-American households, their results were inconsistent for areas with higher shares of Hispanic and Asian households. While in most cases there was no statistically significant association between subprime lending shares and the share of Hispanic and Asian households in a neighborhood, in some specifications there was actually a statistically significant negative coefficient, indicating that subprime shares were lower in neighborhoods with higher shares of Hispanic and Asian households.

Following Calem, Gillen, and Wachter, several other studies incorporated tract-level measures of credit risk into models examining both tract-level and borrower-level subprime lending patterns. A study by the National Community Reinvestment Coalition (NCRC, 2003) incorporated measures of the shares of households with high or no credit scores for 10 metropolitan areas in modeling the neighborhood subprime share of both purchase and refinance loans in 2001. Other independent variables included tract measures of the racial and ethnic composition, household income levels, residential capitalization rate, and property turnover rate. While the results vary somewhat across the market areas studied, the NCRC study found that even with the inclusion of tract credit risk measures, the share of households that are African-American is generally positively associated with subprime lending shares. As in Calem, Gillen, and Wachter, the Hispanic share is less consistently associated with subprime shares. While the African-American share is statistically significant in 6 of 10 purchase loan models and 9 of 10 refinance loan models, the Hispanic share is only significant in 1 purchase model and 5 refinance models.

Calem, Hershaff, and Wachter (2004) expanded on the first Calem, Gillen, and Wachter study by examining five additional metropolitan areas and two other years of subprime lending data using the same measures of tract credit scores in the previous paper (the shares with very low credit scores and without a credit score). Unlike the previous study, Calem, Hershaff, and Wachter focus exclusively on modeling the probability that individual refinance loans will be made by subprime lenders. They also use a somewhat shorter list of tract-level credit risk measures in that neither the residential foreclosure rate nor the HMDA conventional loan denial rate is included. Mirroring the previous results, the study demonstrated that the inclusion of tract-level credit risk measures reduces but does not eliminate the importance of race in predicting whether a loan is subprime, though in this study the credit score variables are not always statistically significant.[9]

The Joint Center (2004) also included tract-level credit risk measures in a model predicting whether individual borrowers across the country obtained a prime conventional or subprime loan in 2001. They incorporated variables related to borrower characteristics, neighborhood

[9] Wachter, Russo, and Hershaff (2006) employed the same data to examine the factors associated with the percent change in subprime refinance loans between 1997 and 2002.

demographics, neighborhood credit risk measures, and the characteristics of lenders operating in the neighborhood. Unlike the previous studies mentioned, they did not have access to credit score measures, but instead relied on conventional loan denial rates from HMDA, the residential capitalization rate, the turnover rate of owner-occupied properties, and income growth during the 1990s. All of these measures are found to be statistically significant and of the expected sign. Of the credit risk measures, they found that the HMDA denial rate was highly correlated with the probability of a loan being subprime. Consistent with the other studies, the Joint Center study confirmed that even after controlling for tract credit risk, the African-American share in the tract is positively and significantly associated with a loan being nonprime, while the results for Hispanics are mixed.

Most recently, in an analysis of the Dallas metropolitan area, Apgar and Herbert (2006)[10] analyzed the geographic distribution of subprime lending, the location of Alternative Financial Service Providers (AFSPs),[11] and banks using a variety of neighborhood characteristics, including a series of credit risk measures other than credit scores. This study confirmed the positive association between a neighborhood's conventional mortgage denial rate in previous years and its higher share of mortgages originated by subprime lenders. Other measures of neighborhood credit risk examined, including the delinquency and claim rates among FHA loans and the rate of residential foreclosures, were not found to be associated with subprime lender shares of refinance loans, but they were associated with their share of purchase loans. The study also confirmed the significance of the neighborhood turnover rate, the capitalization rate, and the share of households with some college education in predicting subprime lender shares.

2.3 Recent Research Focused on Mortgage Pricing Fairness

The growing availability of credit score data fostered numerous studies concerning the fairness of mortgage pricing for African-Americans and other historically disadvantaged groups. In particular there was a growing concern that many households that qualified for lower-priced prime loans were inappropriately steered into taking higher-priced subprime loans. In one widely cited paper, Lax et al. (2000) used detailed data on borrower and loan characteristics, including measures of risk, and found that the interest rate differential between the prime and "A-loans" could not be justified by differences in credit risk of the respective borrowers—even after making conservative assumptions about the loss rates and servicing costs for the loans.

---

[10] William C. Apgar Jr. and Christopher Herbert, "Subprime Lending and Alternative Financial Service Providers: A Literature Review and Empirical Analysis," U.S. Department of Housing and Urban Development, March 2006.

[11] As defined in the study, AFSPs included check cashers, payday lenders, pawnshops, and rent-to-own stores.

In another paper, Courchane, Surette, and Zorn (2004) examined whether minority borrowers were "inappropriately" channeled into the subprime segment. The study explored mortgage lending patterns using FICO scores and other traditional measures of risk, as well as what the authors described as "borrower self-assessed credit risk factors"[12] gathered from a survey of mortgage borrowers. The paper confirmed that whether borrowers obtained subprime or prime mortgages depended in large measure on risk-related mortgage underwriting variables, including FICO score, LTV, and the ratio of monthly housing costs to income (or "front-end ratio"). Nonetheless, other factors not related to borrower risk (including borrower age, market channel, ethnicity, and shopping behaviors) were also significant and suggested that some borrowers may be improperly receiving subprime mortgages.

Although the work of Lax et al. (2000); Courchane, Surette, and Zorn (2004); and other early studies of mortgage pricing discrimination raised serious public policy concerns, they also highlighted the difficulty of assembling the pricing data needed to complete a study that fully controls for differing borrower risk factors. In an effort to expand the availability of loan price information, the federal regulatory agencies mandated that all HMDA-reported loans include annual percentage rate (APR) pricing information when the loan has an APR above a specific threshold. Since HMDA does not contain information on credit history and other loan risk factors, HMDA data alone do not provide conclusive evidence about the differences in pricing by race and ethnicity. Even so, the inclusion of at least some pricing information in HMDA reporting sparked a new generation of research on the patterns of subprime lending.

For example, in the first major use of these newly available HMDA data, Federal Reserve Board (FRB) researchers examined racial patterns of mortgage pricing at the borrower level.[13] Interestingly, the FRB researchers examined the impact on African-American–white and Hispanic–white gaps in access to higher-priced loans using a series of variables that identified the organization making the loan, the extent to which each organization was subject to regulation under the Community Reinvestment Act, and whether and how they initially sold the loan to a GSE or other secondary market outlets. The FRB researchers concluded that these supply-side variables mattered in that they were highly correlated with the magnitude of the observed racial gap.

The FRB researchers were careful to note that their results point to one of several possibilities. First, these supply-side effects could simply reflect that some lenders specialize in higher-priced lending, and these specialists correctly identified those borrowers who are legitimate candidates for higher-priced loans. Alternately, this segmented marketplace could

---

[12] For example, the survey gathered data on whether the borrowers believe they "have good credit," "pay bills on time," and are "in control of their finances," as well as information on search behavior and adverse life events, such as loss of job.

[13] See Avery, Brevoort, and Canner (2006).

be caused by borrower self-selection, whereby borrowers correctly choose the product-lender combination that best matches their needs and credit histories. Lastly, the FRB raised the possibility that "minority borrowers are incurring prices on their loans that are higher than warranted by their credit characteristics," which could trigger fair lending concerns. In other words, the observed patterns could be the result of illegal targeting of protected classes.

In an effort to explore further the issue of racial differences in access to lower-priced prime loans, FRB researchers collaborated with researchers at the Credit Research Center (CRC) of Georgetown University. The CRC prepared a special analysis using a proprietary dataset from eight subprime lending specialists that included a wide range of borrower characteristics and risk factors but did not identify the specific lender. Using this database and controlling for these risk factors, the analysis explained most of the observed racial gap. Yet even accounting for a detailed set of borrower characteristics and risk factors, the FRB researchers noted that a 1 percentage point difference remained between the white and African-American share of higher-priced home purchase lending.

Building on the Federal Reserve Board's work, the Center for Responsible Lending (CRL) conducted a study that utilized publicly available HMDA data augmented by information from a large, proprietary nonprime database to develop a database of 177,487 loans that contained borrower, loan, property, and pricing characteristics, including credit scores (Gruenstein-Bocian et al., 2006). Unlike the CRC data used in the 2005 FRB study that represented the activity of only eight selected lenders, the CRL study presented data on a broad cross section of the nonprime market. It also addressed prevailing interest rates and state-specific information, focusing solely on disparities within the nonprime marketplace. While controlling for risk, the results showed that African-American and Hispanic borrowers are more likely than white borrowers to receive higher-rate nonprime home loans, even after accounting for differences in risk. Those mirror results of a previous CRL study that demonstrated that for nonprime purchase loans, borrowers with prepayment penalties paid higher interest rates than similarly situated borrowers without these penalties, even though the presence of a prepayment penalty should enable the lender to offer financing at a lower rate.[14]

2.4 Reassessing Historical Patterns of Subprime Lending

In combination, the mounting evidence of pricing disparities in the subprime marketplace, along with a growing awareness of the extent to which abusive subprime lending has contributed to the current mortgage crisis, has fueled calls for expanded regulation of mortgage lending. Yet despite the urgency, even with the addition of new credit risk

[14] See Ernst (2005). This study used multivariate regression models to estimate separate results for each year 2000–2002 for fixed-rate loan products using the Loan Performance Asset-Backed Securities (ABS) Database of securitized subprime loans, which includes FICO scores and LTV and DTI ratios.

variables, fully assessing the impact of risk factors on mortgage lending patterns remains difficult. Staten (2005), for example, dismissed these new HMDA-based studies (even those that included measures of mortgage credit risk) as being fundamentally flawed in that they could never fully control for differences in legitimate risk factors, such as credit histories and loan-to-value (LTV) ratios. Others presented methodological critiques. For example, Black and Schlottmann (2006) critiqued the CRL study and questioned whether the matched records between the two databases are representative of overall lending patterns. Others questioned whether the various proposals designed to eliminate the perceived mortgage pricing disparities would do more harm than good. For example, Wallace, Elliehausen, and Staten (2005) argued that legislative solutions to perceived mortgage lending abuses were akin to "throwing the baby out with the bath."

As noted previously, this study seeks to shed additional light on these important policy issues. This study combines newly available census tract-level data on mortgage credit risk and loan-level HMDA data to assess the spatial patterns of mortgage lending in nine metropolitan areas selected to represent the diversity of mortgage and neighborhood market dynamics in the United States. To create a bridge of understanding between older and newer studies on the topic, this study will present results from models of the neighborhood variation in subprime lending using two alternative measures of subprime mortgages: the first based on the HUD-developed list of subprime lenders, and the second based the more recently available information on whether or not a loan is "higher-priced"; namely has a mortgage rate in excess of the established benchmark.

In addition, the study will present a detailed assessment of the relationship between neighborhood credit scores and the mortgage denial rate. If a strong link can be identified between HMDA denial rates and neighborhood credit scores, there will be much greater opportunity for research on historical patterns of neighborhood variation in subprime lending (e.g., studies that extend back before the widespread availability of detailed credit scores) to incorporate at least a rough proxy for neighborhood credit risk. Finally, this study will incorporate the latest and most advanced econometrics to test, and to the extent needed correct for spatial correlation of variables. In doing so, it will assess the extent to which failure to account for spatial autocorrelation may have biased the estimates of the "race effect" in previous studies of the spatial patterns of subprime lending.

Given the elevated attention placed on the topics, the analysis presented here will not quiet the debate on the extent to which racial disparities in the allocation of subprime lending contributed to the concentration of foreclosure-prone subprime loans in lower-income and minority neighborhoods. By presenting the results obtained from the estimation of a variety of models that utilize alternative definitions of subprime loans and alternative credit risk measures, as well as deploying best available econometric techniques for assessing potential spatial autocorrelation bias, this study aims to cast a new light on existing research and

contribute to the ongoing national debate on how best to ensure that all households have access to mortgage capital on a fair basis.

Section 3: Data and Methodology

In this section we describe the process used to select metropolitan areas to be included in the study, the data used in the analysis, and the methodology used to analyze the prevalence of higher-priced lending.

3.1 Data Sources

Three primary data sources are used in this study:

1) Average FICO credit scores at the census tract level;
2) HMDA data on mortgage originations and application denial rates; and
3) Decennial Census data 2000.

Each of these data sources is discussed in turn.

Credit Score Data

Measures of the distribution of credit scores at the census tract level are critically important for the study. The model used to produce these scores is a proprietary product of the Fair Isaac Corporation, and the credit information needed to produce the scores is maintained by each of the three credit reporting agencies. Although the study design called for credit score information aggregated to the census tract level, contractual agreements prevented the credit reporting agencies from selling tract-level aggregate measures. In lieu of readymade tract aggregates, budget considerations dictated that a random selection of individual depersonalized credit scores for homeowners in each census tract of the metropolitan areas included in the study be purchased to support distributional metrics for each census tract. The following steps were taken in obtaining the credit score data to meet the project's budget for the number of records to be purchased:

- Scores were obtained only for the heads of households in owner-occupied housing units, rather than all households;
- A random sample of one out of every five owner-occupants was selected; and
- Mid-sized metropolitan market areas rather than the largest market areas were selected (the market selection process is described further below).

Credit scores were derived from archival history data posted as of June 2004. This period was chosen to be able to examine differences in results for models that utilize the higher-priced loan definition provided by the 2004 HMDA data for subprime loan identification in comparison with models that employ HUD's list of subprime lenders (that was still being

compiled in 2004). In all, 1,622,834 credit scores for homeowners in the nine metropolitan areas were obtained.

Previous studies have found that both the share of households with low scores and the share with no scores are statistically significant. The credit risk metric used in our analysis is the share of owner-occupants in a census tract with credit scores below 620—a common cutoff for eligibility for prime mortgages. This is similar to methodology used in other studies that have incorporated such measures.[15] We explored using a cutoff of 660 instead, but the results were very similar to those based on a cutoff of 620. The measure based on a cutoff of 620 was preferred as the results were somewhat more statistically significant. In addition to the share of households below this level, we also created a variable to measure the share of owner-occupants for which a credit score could not be generated to examine whether the lack of sufficient credit history was also associated with the share of higher-priced loans.

Exhibit 1 presents summary information on average FICO scores at the census tract level for each of the nine market areas included in this study, as well as the average scores for census tracts grouped by the share African-American, the share Hispanic, and census tract median household income as a percent of the metropolitan area median household income. Across the nine market areas included in the study, average census tract-level FICO scores range from a low of 640 in Memphis to a high of 695 in Providence. Memphis had by far the lowest average FICO scores; as the next-lowest markets, Tulsa and Toledo, had average scores of 668 and 670, respectively. There were three market areas with average scores around 680 (Albuquerque, Tacoma, and Baltimore) and another two that had average scores around 690 (Harrisburg and Omaha). As will be discussed more below, in keeping with its low average FICO score, Memphis also had by far the highest share of higher-priced mortgages among the markets studied. The next-highest shares were in Tulsa and Toledo, also areas with lower average FICO scores. Omaha also had a fairly large share of higher-priced refinance mortgages, despite having an average FICO score that is among the highest of these markets.

[15] Calem, Gillen, and Wachter (2004) and Calem, Hershaff, and Wachter (2004) included the share of individuals with very low credit scores, defined as scores in the bottom quintile of the distribution, while NCRC (2003) included the share of individuals with scores in the upper three quintiles of the credit score distribution.

Exhibit 1

Average Census Tract FICO Scores by Metropolitan Area and Selected Census Tract Characteristics (# of Tracts in Parentheses)

Selected Tracts	Albuquerque	Baltimore	Harrisburg	Memphis	Omaha	Providence	Tacoma	Toledo	Tulsa
	Avg FICO Score	Avg FICO Score	Avg FICO Score	Avg FICO Score	Avg FICO Score	Avg FICO Score	Avg FICO Score	Avg FICO Score	Avg FICO Score
All Census Tracts	**678 (185)**	**680 (607)**	**688 (109)**	**640 (283)**	**691 (237)**	**695 (347)**	**680 (153)**	**670 (173)**	**668 (263)**
Census Tracts Categorized by Share African American									
<5%	680 (176)	717 (205)	710 (68)	700 (36)	702 (168)	705 (276)	697 (78)	697 (104)	681 (166)
>=5% & <25%	659 (9)	697 (177)	679 (23)	678 (75)	685 (46)	657 (66)	663 (72)	654 (32)	657 (78)
>25% & <50%	NA (0)	666 (66)	649 (10)	644 (41)	653 (8)	624 (5)	618 (3)	620 (15)	612 (4)
>=50 %	NA (0)	621 (159)	591 (8)	600 (131)	603 (15)	NA (0)	NA (0)	595 (22)	600 (15)
Census Tracts Categorized by Share Hispanic									
<5%	661 (5)	681 (573)	704 (89)	642 (252)	698 (181)	712 (236)	694 (93)	679 (127)	675 (198)
>=5% & <25%	707 (47)	677 (33)	621 (19)	626 (31)	674 (45)	669 (79)	657 (60)	644 (44)	650 (62)
>25% & <50%	681 (76)	660 (1)	598 (1)	NA (0)	645 (8)	638 (24)	NA (0)	589 (2)	623 (3)
>=50 %	653 (57)	NA (0)	NA (0)	NA (0)	650 (3)	617 (8)	NA (0)	NA (0)	NA (0)
Census Tracts Categorized by Income as Percent of Area Median Income (AMI)									
<80% AMI	644 (58)	644 (291)	638 (29)	603 (124)	646 (70)	654 (107)	654 (59)	613 (52)	636 (83)
>= 80% & <100% AMI	673 (46)	696 (138)	690 (31)	635 (44)	698 (69)	694 (68)	689 (52)	675 (47)	664 (67)
>=100% & <120% AMI	689 (30)	713 (97)	709 (27)	655 (39)	709 (44)	714 (78)	699 (30)	700 (40)	679 (38)
>= 120% AMI	712 (53)	732 (95)	726 (22)	694 (76)	726 (54)	726 (94)	716 (12)	711 (34)	701 (75)

Exhibit 1 also shows for each market area the average FICO score for census tracts grouped by tract share of African-American households, tract share of Hispanic households, and tract median household income as a percent of the metropolitan area median household income. These tabulations show that there is a fairly strong association between these neighborhood characteristics and average FICO scores, with lower FICO scores in tracts with higher minority shares and higher FICO scores in tracts with higher relative incomes. These tabulations also illustrate that the cross-market differences in average FICO scores for metropolitan areas are in part due to differences in the composition of the population in these market areas. For example, across all of the markets that have census tracts that are majority African-American, the average FICO score in these neighborhoods is about 600. Thus, the lower marketwide average FICO score for Memphis, in part, reflects the fact that it has a much higher proportion of these high-share minority tracts than the other markets studied.

Exhibit 2 presents correlation coefficients between tract-level average FICO scores and tract-level characteristics, including the tract shares of African-American and Hispanic households, the tract median household income, and the HMDA denial rate for prime mortgage applications. FICO scores correlate fairly highly with all of these variables, but the highest correlation is with median household income, which averages 0.84 across the nine market areas. There are also fairly large negative correlations with both share African-American (-0.68) and share Hispanic (-0.44), although excluding the two markets areas with very low correlations with the share Hispanic (Baltimore and Memphis), the average correlation in the remaining markets was -0.56.

HMDA Data

Under the Home Mortgage Disclosure Act, mortgage lenders exceeding specified thresholds for asset size and volume of lending activity in metropolitan areas are required to file annual reports on all mortgage applications they receive and the disposition of these applications. The reported information includes the census tract where the property is located; the race, ethnicity, gender, and income level of the applicant; selected characteristics of the loan and property; and the disposition of the application. Beginning with reports filed for the year 2004, lenders are also required to identify whether the interest rate on an originated loan exceeds a higher-priced threshold based on a comparison of the loan's annual percentage rate to the interest rate on Treasury notes of comparable maturity as of the date the loan is originated. The threshold is set at 3 percentage points for first-lien mortgages and 5 percentage points for other liens. Loans identified as higher-priced in HMDA are widely taken to be subprime mortgages, although given variations in interest rates on individual loans as well as the complex relationship between mortgage rates and Treasury rates, this identification is by no means perfectly accurate.

Exhibit 2

Correlation Coefficients for Average Census Tract FICO Scores and Selected Tract Characteristics by Metropolitan Area

Census Tract Characteristic	Albuquerque	Baltimore	Harrisburg	Memphis	Omaha	Providence	Tacoma	Toledo	Tulsa	Average All Metro Areas
Median Family Income	0.84	0.84	0.83	0.80	0.87	0.91	0.80	0.81	0.85	0.84
HMDA Prime Mortgage Denial Rate	-0.64	-0.77	-0.81	-0.74	-0.75	-0.75	-0.48	-0.85	-0.67	-0.72
Share of Population with Bachelor's Degree	0.82	0.72	0.60	0.76	0.69	0.74	0.68	0.64	0.68	0.70
Share African-American	-0.12	-0.84	-0.84	-0.87	-0.70	-0.65	-0.71	-0.76	-0.60	-0.68
Share Hispanic	-0.59	0.00	-0.72	-0.07	-0.38	-0.77	-0.65	-0.46	-0.35	-0.44
Correlation Between FICO Scores and HMDA Higher-Price and HUD Subprime Loan Share										
Share of Purchase Loans that are Higher-Priced	-0.13	-0.69	-0.80	-0.68	-0.76	-0.81	-0.64	-0.66	-0.54	**-0.64**
Share of Refinance Loans that are Higher-Priced	-0.31	-0.78	-0.81	-0.79	-0.82	-0.74	-0.68	-0.84	-0.64	**-0.71**
Share of Purchase Loans made by Subprime Lenders	0.06	-0.54	-0.67	-0.53	-0.63	-0.77	-0.49	-0.38	-0.28	**-0.47**
Share of Refinance Loans made by Subprime Lenders	0.04	-0.66	-0.69	-0.59	-0.69	-0.74	-0.59	-0.59	-0.45	**-0.55**

Source: Tract characteristics from Census Summary File 3; Prime mortgage denial rate based on applications in 2004 HMDA made to lenders not on HUD's subprime lender list in each market area.

HMDA data available through 2003 did not include any variable that directly identifies whether the loan is subprime or not. To address this shortcoming, HUD developed an approach that classifies loans as subprime by identifying lenders who primarily originate subprime loans.[16] This methodology has been used by all previous studies that have relied on HMDA data to examine subprime lending patterns. In the 1990s this methodology was probably fairly accurate, as subprime lending was largely conducted by specialized lenders. However, over time, primarily prime lenders have entered the subprime business in order to diversify their product lines and to tap this growing market. As a result, in recent years the approach of using HUD's list of lenders that are primarily subprime may not be as accurate in identifying subprime loan volumes.

Decennial Census Data

Data from the 2000 Decennial Census at the censustract level provide a variety of neighborhood characteristics to control for mortgage risk. Following previous research, information on tract demographic and housing stock characteristics are independent variables in the analysis of higher-priced loan shares.

3.2 Market Selection

The primary factors used in site selection were to provide geographic representation of each of the nine census divisions and to provide a range of market conditions with respect to the share of higher-priced loans reported in the 2004 HMDA data. As noted above, mid-sized markets were also selected to stay within the available budget for credit score data.

The metropolitan areas selected for the study and selected characteristics of these areas are shown in Exhibit 3.[17] In general, the markets selected had higher-priced loan shares that were above the share for the United States as a whole, although in most cases the shares were only slightly higher than the U.S. shares. Memphis stands out as having very high shares of higher-priced loans—more than twice the overall U.S. share. Among refinance loans, Baltimore, Omaha, Toledo, and Tulsa also had higher-priced loan shares that were in excess of 20 percent (or more than 33 percent above the overall U.S. share). Among purchase mortgages, Tulsa was the only market area other than Memphis that had a higher-priced loan share that was above 20 percent. But for the most part, the markets were not selected to include areas with the highest shares of higher-priced loans in the country.

---

[16] For the list of subprime lenders and a discussion of the methodology used to identify subprime lenders, see http://www.huduser.org/datasets/manu.html.

[17] The metropolitan area definitions used in the study were those defined in June 2003 based on results of the 2000 Decennial Census. These definitions were used in reporting the 2004 HMDA data.

Exhibit 3

Selected Demographic Characteristics of Metropolitan Areas Selected for Study

Metropolitan Area	Total Population	Higher-Priced Loan Share of Refinance Loans (%)	Higher-Priced Loan Share of Purchase Loans (%)	Share Non-Hispanic White (%)	Share African-American (%)	Share Hispanic (%)	Median Family Income (1999 $)
Albuquerque	729,554	14.2	12.8	48.0	2.1	41.5	46,037
Baltimore	2,552,994	21.1	14.8	66.3	27.1	2.0	66,909
Harrisburg	629,401	17.2	14.8	86.5	7.5	3.2	51,892
Memphis	1,205,204	35.3	30.7	51.9	43.3	2.2	47,440
Omaha	767,041	22.5	16.1	83.8	7.5	5.2	54,596
Providence	1,582,997	13.1	16.2	84.4	3.2	2.6	52,297
Tacoma	700,820	15.7	16.3	76.0	6.8	5.5	60,195
Toledo	659,188	23.8	17.0	81.0	11.8	5.5	50,286
Tulsa	859,532	26.7	22.7	73.8	8.6	4.6	46,479
US	281,421,906	15.7	14.4	69.1	12.0	12.5	50,046

Note: Metropolitan area definitions used are those defined as of 2003. However, the median family income shown is for the 1999 metropolitan area definitions used to report results of the 2000 Decennial Census.

Source: Metropolitan area higher-priced loan shares from tabulations of 2004 HMDA data by ArcBridge Consulting; U.S. higher-priced loan shares from Avery et al. (2005). Population and income variables from 2000 Decennial Census, Summary File 3.

The metropolitan areas selected are generally not markets with significant minority populations. Baltimore and Memphis both have significant African-American populations, with a population share that is well in excess of the U.S. share. Albuquerque also has a very large Hispanic population, which is several times the U.S. share. The remaining six metropolitan areas all have African-American and Hispanic population shares that are below the overall U.S. share. While the focus of this study is on racial/ethnic differences in the use of higher-priced loans, many of the market areas selected have lower African-American and Hispanic population shares than the United States as a whole. Thus, this study is somewhat unique in looking at smaller markets, with generally more moderate higher-priced loan shares and below average minority population shares.

3.3 Methodology

Dependent Variables and Modeling Approach

This study focused on two main tract-level dependent variables: the share of mortgage originations in a census tract that are subprime and the choice by individual borrowers of a subprime loan. For the census tract-level analysis we estimate an ordinary least squares regression model but use the number of loan originations in the tract as weights. The use of weights is intended to account for the fact that tracts with a small number of originations will have much greater variation in the subprime loan share, and so we do not want these observations to be overly influential in the model. For the borrower-level analysis, a logit model will be estimated where the dependent variable is 1 when the borrower obtained a subprime loan and 0 otherwise.

As discussed above, we use two methods for identifying subprime loans: the higher-priced loan designation in the HMDA data and HUD's list of subprime lenders. While much of the previous research has focused on the share of refinance loans that are subprime, we will also examine the share of purchase mortgages that is subprime, given that the subprime purchase market was growing in importance by 2004.

Independent Variables

Exhibit 4 presents the means and standard deviations of the variables used in tract-level analysis.[18] Drawing upon previous research, the following are the primary tract-level independent variables that are used in the analysis:

- ***Share of Population that is African-American or Hispanic:*** Previous research has found a strong positive association between the share of African-Americans in a neighborhood and the likelihood of using subprime lenders even after controlling for a variety of characteristics. The association with the share Hispanic is less consistent, although there is also generally a positive association.
- ***Median Family Income:*** Since the risk of mortgage default would be expected to fall with increases in income, neighborhood income levels have generally been found to have a negative association with subprime lending shares. Generally, this variable has been measured in log form given the approximate log-normal distribution of incomes.
- ***Share of adults age 25 and older with a bachelor's degree:*** Several previous studies have found a negative relationship between the share of adults with a

[18] Note that the figures shown in Exhibit 4 do not match the figures shown in Exhibit 3 because Exhibit 3 presents data for the metropolitan area as a whole, while Exhibit 4 presents average census tract values for these market areas.

college degree and the use of subprime lending, with areas with higher shares of college graduates less likely to use subprime lenders.

Exhibit 4

Summary Statistics for Census Tract Variables

(Census Tract mean values with standard deviation in parentheses)

	Albuquerque	Baltimore	Harrisburg	Memphis	Omaha	Providence	Tacoma	Toledo	Tulsa
Share Purchase Loans Originated by Subprime Lenders (%)	10.11 (6.74)	10.34 (8.11)	9.22 (6.54)	23.24 (18.35)	12.59 (10.70)	22.87 (14.08)	22.86 (8.31)	10.56 (7.64)	13.02 (9.43)
Share Higher-Priced Refinance Loans	14.17 (9.61)	21.08 (15.41)	17.15 (10.66)	35.27 (21.90)	22.50 (13.69)	13.13 (7.50)	15.66 (7.15)	23.78 (15.63)	26.71 (15.26)
Share Higher-Priced Purchase Loans	12.79 (10.96)	14.78 (12.60)	14.77 (9.96)	30.71 (22.54)	16.07 (13.13)	16.17 (11.25)	16.25 (6.42)	16.95 (14.60)	22.74 (14.65)
Share Refinance Loans Originated by Subprime Lenders (%)	18.70 (8.49)	22.69 (13.13)	18.69 (10.34)	29.76 (18.99)	21.41 (12.23)	27.25 (12.34)	23.16 (8.88)	22.71 (13.58)	21.71 (12.21)
Black, Non-Hispanic (%)	2.06 (2.04)	29.75 (34.51)	11.16 (19.75)	47.21 (36.19)	8.31 (17.83)	3.56 (5.40)	6.37 (6.61)	16.08 (27.00)	9.22 (18.69)
Hispanic (%)	39.62 (21.34)	1.89 (2.22)	2.96 (4.09)	2.09 (3.00)	5.41 (9.02)	7.38 (12.17)	5.16 (3.53)	4.59 (4.89)	4.44 (5.14)
Other Race, Non-Hispanic (%)	9.99 (16.83)	4.25 (3.70)	3.26 (2.25)	2.43 (2.61)	3.43 (2.50)	5.77 (5.47)	11.45 (6.53)	2.81 (2.06)	12.61 (5.70)
Tract Median Family Income (Ln)	10.70 (0.38)	10.87 (0.46)	10.79 (0.31)	10.59 (0.57)	10.85 (0.36)	10.78 (0.40)	10.83 (0.28)	10.72 (0.38)	10.72 (0.38)
Share Pop. Age 25 or Older With Bachelor's Degree	14.58 (8.01)	15.38 (9.74)	14.55 (7.78)	12.96 (10.39)	18.15 (10.07)	13.95 (7.94)	13.00 (6.79)	12.13 (7.97)	14.94 (10.08)
Rental Rate	29.36 (21.26)	33.07 (23.11)	32.71 (19.95)	35.88 (23.93)	32.02 (21.45)	41.44 (24.69)	33.98 (22.25)	32.61 (20.16)	30.57 (20.51)
Capitalization Rate	0.53 (0.42)	0.53 (0.22)	0.51 (0.19)	0.57 (0.29)	0.55 (0.24)	0.39 (0.10)	0.46 (0.23)	0.52 (0.24)	0.58 (0.22)
Homeowner Turnover Rate	12.88 (18.38)	9.72 (6.21)	7.82 (3.62)	9.93 (11.96)	9.19 (10.15)	10.14 (6.99)	15.00 (12.10)	6.50 (3.74)	7.27 (5.84)
Share Households Credit Score <620	30.65 (12.21)	30.36 (17.73)	27.59 (15.92)	46.92 (19.34)	26.40 (14.32)	24.60 (12.21)	29.14 (10.86)	33.97 (17.81)	34.72 (13.62)
Share Households Unscorable	21.75 (6.60)	23.53 (13.78)	20.79 (11.33)	26.57 (11.17)	17.03 (10.61)	16.72 (9.83)	28.42 (10.14)	24.87 (14.05)	27.27 (10.67)
Prime Mortgage Denial Rate	18.70 (10.90)	16.96 (10.59)	16.24 (7.81)	25.42 (13.12)	25.99 (10.73)	14.86 (5.66)	17.02 (5.35)	32.33 (11.79)	21.01 (9.40)

Sources: See text for sources for each variables.

- ***Rental Rate:*** The share of housing units occupied by renters (or conversely the share occupied by owners) has generally been included in previous studies on the assumption that higher rental rates will increase the risk of falling home prices, increasing the risk of mortgage default and so making the use of subprime mortgages more likely. However, there have also been findings of a positive association between subprime lending rates and homeownership rates, which may be attributable to lenders aiming marketing efforts at neighborhoods with higher homeownership rates where the number of potential customers is greater.
- ***Residential Capitalization Rate:*** Most previous studies of neighborhood subprime lending shares have included the residential capitalization rate as an explanatory variable. In general, a capitalization rate is the rate of financial return on a financial asset given by the ratio of income to the value of the asset. Since the rate of return on investments should be equivalent after taking into account the risk of the investment and the expected appreciation rate of the asset, lower capitalization rates are taken as an indication of lower risk and/or higher expected appreciation. Thus, lower capitalization rates indicate less risk for investors in residential property. The most common measure in subprime lending studies is the ratio of median gross rent in a tract to median property values. These measures have consistently been significant predictors of subprime lending shares, with higher capitalization rates associated with higher subprime lending shares.
- ***Owner-Occupied Turnover Rate:*** The share of owner-occupied units that turned over has consistently been found to be a strong predictor of subprime lending shares, with greater turnover thought to indicate greater demand for owner-occupied homes, reducing the risk of mortgage default, and so lowering the incidence of subprime lending. The turnover rate is estimated by dividing the number of home purchase mortgages in 2004 from HMDA by the number of owner-occupied housing units for the census tract from the 2000 Decennial Census.
- ***Neighborhood Credit Score Distribution:*** Two measures of credit scores are used: the share of homeowners with scores below 620, the level generally identified as the cutoff for qualification for prime mortgages, and the share of households without a credit score. A higher share of households with scores below 620 is expected to increase the likelihood of using subprime loans. Similarly, since households without a credit score would be less likely to qualify for a prime mortgage, it would also be expected that a higher share of unscored households would increase the use of subprime lending.
- ***Prime Mortgage Denial Rate from HMDA:*** As discussed in Section 2, a number of previous studies has used measures based on the share of mortgage applications

to prime lenders (i.e., lenders not included on HUD's list of subprime lenders) that are denied as a proxy for a neighborhood's credit risk. To examine how well the HMDA denial rate proxies for credit scores, we estimate census tract-level prime mortgage denial rates using the 2004 HMDA data.[19]

Exhibit 5 presents summary statistics for the borrower-level characteristics included in the logit models in addition to the tract-level variables described above. The borrower characteristics included in HMDA that are used to the models predicting whether a higher-priced (or subprime) mortgage is chosen are the following:

- ***Race:*** Given the findings of previous studies, the primary focus of this study is on the association between African-American borrowers and the use of higher-priced loans. A variable indicating whether the borrower is African-American will be included in the borrower-level models. The expectation is that the coefficient on this variable will be positive, indicating a greater likelihood of using higher-priced loans even after controlling for other borrower and neighborhood characteristics. In some cases lenders do not have to (or fail to) report borrowers' race. An indicator will also be included for these borrowers to see if the failure to report race is associated with greater use of higher-priced loans.
- ***Ethnicity:*** Indicator variables for borrowers who are Hispanic or do not have their ethnicity reported will also be included.
- ***Gender:*** Indicator variables will be included for female and male borrowers and for borrowers where gender is not identified.
- ***Income:*** Borrowers with lower incomes would be expected to be more likely to use higher-priced loans. Given that borrower incomes are not reported in all cases and we would like to include these observations in the model, income is included as a categorical rather than a continuous variable. Reported borrower income is measured as a percent of the area median income (AMI), with categories created for those below 80 percent of AMI, between 80 percent and 100 percent of AMI, 100 percent and 120 percent of AMI, and 120 percent of AMI or higher. Those with missing incomes are a separate category.

[19] Previous studies have generally used denial rates from one or more years prior to the year being studied to avoid any simultaneity between the denial rate and the subprime loan share. However, since 2004 was the first year in which HMDA data used newly defined census tract definitions, it would be difficult to link these data to HMDA data reported in previous years.

Exhibit 5

Summary Statistics for Borrower Characteristics
(Percent of Borrowers)

Census Tract Characteristic	Albuquerque	Baltimore	Harrisburg	Memphis	Omaha	Providence	Tacoma	Toledo	Tulsa
Race									
African-American	1.7	16.5	4.6	30.1	3.3	4.4	4.7	5.8	4.6
White	78.3	66.1	83.2	58.1	87.0	79.2	76.2	84.5	80.5
Other Race	4.8	6.0	3.4	2.7	2.7	2.7	8.0	1.7	6.1
Race Not Identified	15.2	11.4	8.7	9.1	7.0	13.7	11.1	8.0	8.9
Ethnicity									
Non-Hispanic	53.9	76.5	79.5	78.5	77.5	71.3	77.9	81.1	78.0
Hispanic	29.3	3.1	1.6	2.5	4.9	7.7	4.5	2.7	4.7
Ethnicity Not Identified	16.8	20.4	18.9	19.1	17.6	21.0	17.5	16.2	17.3
Gender									
Female Borrower	32.1	30.8	27.8	32.9	25.9	28.0	26.3	26.6	27.0
Male Borrower	64.1	64.3	68.4	62.7	70.8	67.3	68.8	69.7	68.7
Gender Not Identified	3.8	4.9	3.8	4.4	3.3	4.7	4.9	3.7	4.3
Income									
Income LT 80% of AMI	29.7	26.4	37.3	29.2	29.9	17.7	37.7	33.6	34.0
Income 80–99% of AMI	13.5	14.3	14.9	14.9	14.0	15.8	18.8	15.1	13.6
Income 100–119% of AMI	11.5	13.1	11.1	11.1	13.1	15.8	14.6	11.9	11.2
Income 120% AMI or more	41.6	41.2	32.6	41.4	39.8	45.3	26.1	36.2	37.1
Income Missing	3.7	5.0	4.0	3.2	3.0	5.1	2.7	3.2	4.0

Source: 2004 HMDA data.

Econometric Tests for Spatial Autocorrelation

Many studies of neighborhood subprime mortgage lending, like the one by Calem et al. (2004), adopted a regression approach with the dependent variables reflecting the mortgage characteristics and the independent variables reflecting a set of mortgage applicants' socioeconomic-demographic or neighborhood characteristics. While the selection of variables was reasonable in most of these studies, the regression models adopted—either linear using ordinary least squares estimation methods or nonlinear (such as logistic regressions) using maximum likelihood estimation methods—may suffer from a methodological problem since observations of adjacent neighborhoods are likely dependent upon each other. Clearly, housing values of neighboring houses are interdependent and marketing efforts of lenders may also be spatially dependent. In the context of mortgage lending analysis, the types of regression models used in previous studies assume that the variables, both dependent and independent, are independent among observations, but we know that this assumption is likely violated in most cases.

The presence of spatial dependence, referred to in spatial econometrics as spatial autocorrelation (SA), will not result in biased parameter estimates of standard regression models, but the estimated variances of the estimates usually fall below the true values, producing artificially stronger statistical results than are warranted. One way to conceptualize this problem is that the presence of SA essentially creates duplicative or redundant sample observations. Observations providing similar information because of their spatial proximity reduce the effective sample size or, alternatively, inflate the degrees of freedom such that the consequent probabilities for testing the significance of parameter estimates are no longer reliable in the presence of SA (Anselin and Griffith, 1988). In other words, SA may not create biased coefficient estimates, but it can create inefficient estimates. Appendix B provides further details on the approaches used both to test for SA and to estimate models accounting for SA, as well as the results obtained from these models.

Section 4: Analysis

4.1 Census Tract-Level Analysis

Exhibit 6 presents a summary of the OLS regression results estimating the census tract-level share of higher-priced loans in both the refinance and purchase loan markets. The exhibit compares results for models with and without the share of households with FICO scores below 620 in the tract and the share of individuals that did not have a credit score to see how the inclusion of credit measures affects the model results. Full details on the estimation results are presented in Appendix A.[20]

Focusing first on the results for the higher-priced share of refinance loans, in general the regression models do a fairly good job of explaining variation in higher-priced lending shares across tracts with an average of R-squared of 0.73 in models without the credit score measures and 0.75 in models including the credit measures.[21] In the models without credit measures, the tract median household income is most commonly statistically significant with the expected sign (in eight of nine metropolitan areas), while the share of adults with a college education, the rental share of housing units, and the capitalization rate are also statistically significant and of the expected sign in six of nine metropolitan areas. (In one market the capitalization rate coefficient is statistically significant, but not of the expected sign.) The homeowner turnover rate is statistically significant and of expected sign in only three metropolitan areas.

Of the racial/ethnic population share measures, the share African-American is statistically significant and positive in six of nine metropolitan areas. The three market areas without a statistically significant association are the areas with a relatively small African-American population, with the average census tract having fewer than 10 percent African-Americans (Albuquerque, Providence, and Tacoma). The share Hispanic is positive and statistically significant in four metropolitan areas and negative and statistically significant in one metropolitan area despite their relatively small Hispanic populations averaging less than 5 percent of the population per census tract. Albuquerque is an exception, with an average of more than 40 percent of the average census tract comprised of Hispanics, but with no

---

[20] The models were also tested to determine if spatial autocorrelation was evident in the pattern of higher-priced lending across census tracts. The test results are presented in Appendix B. In general, spatial autocorrelation was found in a minority of the estimated models. However, since the results from the spatial correction models were not materially different from the OLS results, we present the OLS results here.

[21] R-squared is the share of the variation in the dependent variable that is explained by the regression model. An R-squared measure of 0.73 indicates that the models explain 73 percent of the variation across census tracts in higher-priced refinance shares.

statistically significant association between these shares and the higher-priced refinance loan shares.

Exhibit 6 summarizes results for the models estimating the share of higher-priced purchase mortgages as well. In general, the models are less accurate in predicting the higher-priced purchase mortgage shares, with R-squares that average 0.63 in models without credit score measures. Moreover, fewer of the independent variables are statistically significant. In particular, tract median household income ceases to be statistically significant in all but three of nine metropolitan areas as does the share of rental units. The share of adults with a college education is significant for eight of nine market areas and the homeowner turnover rate is significant for seven of nine market areas. The differences in the importance of the independent variables suggest that there is a different process producing higher-priced loan shares in the purchase and refinance markets.

Exhibit 6

Summary of Regression Results for Tract-Level Higher-Priced Refinance and Purchase Share Models
(Number of Statistically Significant Coefficients that are +/- Across Nine Markets)

Variable (Expected Sign in Parentheses)	Refinance		Purchase	
	Without Credit Scores	With Credit Scores	Without Credit Scores	With Credit Scores
African-American Population Share (+)	6/0	5/0	7/0	3/0
Hispanic Population Share (+)	4/1	3/2	4/0	2/2
"Other" Race Population Share (+)	3/0	1/1	3/0	3/0
Tract Median Household Income (Ln) (-)	0/8	0/5	0/3	0/2
Share Pop. With College (-)	0/6	0/5	0/8	1/6
Rental Rate (-)	0/6	0/7	0/3	0/5
Capitalization Rate (+)	6/1	5/0	5/0	3/0
Homeowner Turnover Rate (-)	0/3	3/1	0/7	0/6
Share Households Credit Score <620 (+)	NA	8/0	NA	8/0
Share Households Unscorable (+)	NA	2/1	NA	0/0
Average R-squared	0.73	0.75	0.63	0.66

Note: Table shows the number of market areas where the sign of the coefficient is statistically significant and positive to the left of the slash, and the number of market areas where the sign of coefficient is statistically significant and negative to the right of the slash. See Appendix A, Exhibits A-1 and A-2 for complete modeling results.

However, there are similarities in the coefficients for the African-American and Hispanic population share variables, with seven of nine markets having a statistically significant positive association with the African-American share and four of nine having a statistically significant positive association with the Hispanic share.

When the credit score measure is added to the models, it is highly statistically significant and of the expected sign in eight of the nine metropolitan areas for both refinance and purchase mortgages. Clearly, credit scores are an important factor in variation of the prevalence of higher-priced mortgages across neighborhoods. The share of individuals that do not have a credit score is only statistically significant and of the expected sign in two market areas, is significant and negative in one market for refinance loans, and is not statistically significant in any market for purchase mortgages. The addition of the credit score variables does not have much impact on the overall explanatory power of the models, with the R-squared measure generally only increasing a few percentage points for both refinance and purchase mortgages. However, the inclusion of these variables does somewhat reduce the explanatory power of the other tract-level explanatory variables, with fewer of them statistically significant and many of them smaller in magnitude when they are significant. The variables that are most consistently significant and of the expected sign are the share of adults with a college education and the rental share of housing units. The capitalization rate tends to be more significant in the refinance models, while the homeowner turnover rate is more significant in the purchase models.

Our concern is mostly with the effect of including credit scores on the estimated association between the share of African-Americans and Hispanics in the tract and the higher-priced lending shares. For the most part, including credit scores does reduce this association for refinance mortgages, but in most markets there remains a statistically significant association. In the models estimating the higher-priced share of refinance mortgages, the share African-American is still statistically significant and positive in five metropolitan areas and the share Hispanic is still statistically significant and positive in three market areas. Nevertheless, the magnitude of the coefficients on these minority population share measures is smaller in all market areas.

For models estimating the higher-priced share of purchase mortgages, the impact of including credit scores is larger, with only three market areas having a statistically significant and positive association with the share African-American and two market areas having a statistically significant and positive association with the share Hispanic after the credit measures are included. Again, even in markets where the coefficients remain statistically significant, they are smaller in magnitude. There are also two markets in both the refinance and purchase markets where the share Hispanic becomes statistically significant and negative when credit scores are included.

Exhibit 7 summarizes the changes in the coefficients on tract share African-American and Hispanic. In the models of the higher-priced share in the refinance market, across all metropolitan areas, the average coefficient on share African-American declines from 0.14 to 0.06 (a decline of 59 percent), while the average coefficient on the share Hispanic declines from 0.09 to 0.00 (a decline of nearly 100 percent). Considering only metropolitan areas with statistically significant coefficients in the models without credit measures, on average the African-American coefficient is reduced by 40 percent and the Hispanic coefficient is reduced by 34 percent. In the models of the higher-priced share in the purchase market, the average coefficient on the African-American population share declines from 0.21 to 0.09 (a decline of 57 percent), while the average coefficient on the Hispanic population share declines from 0.22 to 0.06 (a decline of 75 percent). Considering only markets with statistically significant coefficients when credit measures are excluded from the models, the declines in coefficients are 57 percent and 61 percent, respectively, for African-American and Hispanic population shares.

Exhibit 7

Summary of Change in Racial/Ethnic Coefficients from Inclusion of Credit Score Variables in Models of Higher-Priced Loan Shares

	African-American			Hispanic		
Metropolitan Area	**Without Credit Scores**	**With Credit Scores**	**Percent Change**	**Without Credit Scores**	**With Credit Scores**	**Percent Change**
Refinance						
Albuquerque	-0.40	-0.42	4%	0.03	0.00	-86%
Baltimore	**0.22**	**0.14**	-36%	-0.16	**-0.17**	-7%
Harrisburg	**0.27**	**0.18**	-35%	-0.05	-0.08	-62%
Memphis	**0.29**	**0.26**	-9%	0.18	0.15	-16%
Omaha	**0.27**	**0.12**	-55%	**0.26**	**0.12**	-51%
Providence	0.08	0.04	-49%	**0.17**	**0.08**	-51%
Tacoma	0.11	0.00	-103%	**0.31**	**0.27**	-15%
Toledo	**0.25**	**0.12**	-52%	**0.51**	0.13	-75%
Tulsa	**0.17**	0.08	-54%	**-0.41**	**-0.50**	-20%
Average	0.14	0.06	-59%	0.09	0.00	-98%
Purchase						
Albuquerque	0.19	0.13	-32%	-0.04	**-0.08**	91%
Baltimore	**0.18**	**0.09**	-50%	0.07	0.05	-38%
Harrisburg	**0.19**	0.07	-62%	**0.59**	**0.53**	-10%
Memphis	**0.33**	**0.13**	-62%	**0.42**	**-0.03**	-108%
Omaha	**0.24**	0.04	-82%	**0.30**	0.03	-89%

Exhibit 7

Summary of Change in Racial/Ethnic Coefficients from Inclusion of Credit Score Variables in Models of Higher-Priced Loan Shares

Providence	0.05	-0.02	-145%	**0.32**	**0.20**	-36%
Tacoma	**0.25**	0.13	-46%	0.21	0.14	-34%
Toledo	**0.23**	**0.12**	-50%	0.26	-0.10	-138%
Tulsa	**0.20**	0.10	-47%	-0.13	-0.24	80%
Average	0.21	0.09	-57%	0.22	0.06	-75%

Note: Statistically significant coefficients shown in bold. See Appendix A, Exhibits A-1 and A-2 for complete modeling results.

Just because a number of the racial/ethnic coefficients are *statistically significant* does not mean that the difference in higher-priced loan shares associated with changes in a tract's racial/ethnic composition are *large in magnitude*. It is also difficult to compare the magnitude of the association across markets based solely on the size of the estimated coefficient since there are large differences across markets in the range of minority population shares across tracts. For example, the size of the coefficient in the refinance model on the African-American share of tract population is similar in Baltimore and Omaha, but since there is less variation across Omaha's neighborhoods in the African-American population share the racial coefficient does not explain as much of the inter-neighborhood variation in higher-priced loan shares in the market area.

Exhibit 8 presents estimates of the magnitude of the association of a two standard deviation change in a tract's share African-American or Hispanic in the metropolitan areas where the coefficients for these variables were statistically significant.[22] In general, larger impacts are found in the refinance market and are associated with the African-American population share. For refinance mortgages, the higher-priced share of mortgages is estimated to increase from 4.3 to 18.9 percentage points with a two standard deviation increase in the African-American population share across the four market areas with a statistically significant association with this variable. For the Hispanic population share, the estimated increase is generally fairly small, with three of the markets having an increase of only 1.9 to 2.2 percentage points in the higher-priced share from a two standard-deviation increase in the

[22] The mean and standard deviation of the tract share of the population that is African-American or Hispanic is presented in Exhibit 2 in Section 3 of this study. In markets such as Omaha where the African-American share only averages 8.3 percent across all census tracts, the standard deviation of this measure is 17.8 percentage points. In contrast, in Memphis, the average share African-American is 47.2 percent and the standard deviation of this share is 36.2 percentage points. Thus, in Memphis, there is much greater variation across tracts in the African-American population share, and so greater variation in the higher-priced lending share that is associated with these population measures.

Hispanic population share. However, in Harrisburg there is an estimated 7 percentage point increase associated with a two standard deviation change in the Hispanic share of the population in the tract.

In the purchase market, across the three market areas with a statistically significant coefficient, the higher-priced share increases from 5 to 9.1 percentage points with a two standard deviation increase in the African-American population share. In the two metropolitan areas with a statistically significant Hispanic population share coefficient, the share of higher-priced purchase loans is estimated to increase from 3.8 to 3.9 percentage points for a two standard deviation increase in the Hispanic population share.

Exhibit 8

Change in Higher-Priced Loan Share Associated with Two Standard Deviation Change in Minority Share of Neighborhood Population

Metropolitan Area	Racial/Ethnic Group	Refinance Mortgages	Purchase Mortgages
Baltimore	African-American	9.6	5.9
Memphis	African-American	18.9	9.1
Omaha	African-American	4.3	NA
Toledo	African-American	6.5	5.0
Harrisburg	Hispanic	7.0	3.8
Omaha	Hispanic	2.2	NA
Providence	Hispanic	2.0	3.9
Tacoma	Hispanic	1.9	NA

Note: NA denotes areas where the estimated effect is not applicable since the coefficient was not statistically significant and/or of the expected sign.

Comparison of Results Using HUD Subprime Lender List

As discussed in Section 2, prior to 2004, the HMDA data did not include any indication of whether a loan was higher-priced. Beginning in the late 1990s, HUD staff developed an annual list of lenders that primarily originated subprime loans as a means of identifying subprime loans in HMDA. Since much of the existing literature examining subprime lending patterns across neighborhoods or borrowers used HMDA data prior to 2004, these studies relied on the HUD list of primarily subprime lenders as a means of flagging subprime loans. In this section, we estimate tract-level models of the share of loans originated by HUD subprime lenders and compare these results to the estimates presented above for higher-priced loan shares to see whether there is any substantial difference in the conclusions reached regarding the association of a neighborhood's racial/ethnic composition and the incidence of higher-priced or subprime lending.

Exhibit 9 compares summaries of regression results estimating the share of higher-priced loans versus models estimating the share of loans made by HUD-identified subprime lenders. In general, the models predicting loans made by HUD-identified subprime lenders are less accurate than the models of higher-priced loans, having lower average R-squares and fewer statistically significant variables. However, in many respects, the results are similar in terms of which variables are statistically significant. In terms of the racial/ethnic population share variables, the HUD subprime lender list tends to produce more statistically significant findings. The models based on the HUD subprime lender list find a statistically significant positive association with the African-American population share in six metropolitan areas for

refinance mortgages and four metropolitan areas for purchase mortgages, compared to five market areas for refinance loans and three market areas for purchase loans using the higher-priced loan share measure. For Hispanics, the results are similar, with the HUD subprime lender list producing slightly fewer statistically significant coefficients for refinance mortgages (2 versus 3 for the higher-priced loan share) but more for purchase mortgages (5 compared to 2 for the higher-priced loan share).

Exhibit 9

Summary of Regression Results for Tract-Level Higher-Priced Share and HUD Subprime Lender Share Models
(Number of Statistically Significant Coefficients that are +/- Across Nine Markets)

Variable (Expected Sign in Parentheses)	Refinance		Purchase	
	HMDA Higher-Priced	HUD Subprime List	HMDA Higher-Priced	HUD Subprime List
	5/0	6/0	3/0	4/0
Hispanic Population Share (+)	3/2	2/0	2/2	5/1
"Other" Race Population Share (+)	1/1	1/1	3/0	2/1
Tract Median Household Income (Ln) (-)	0/5	1/2	0/2	0/2
Share Pop. With College (-)	0/5	0/6	1/6	0/4
Rental Rate (-)	0/7	0/3	0/5	0/5
Capitalization Rate (+)	5/0	5/0	3/0	4/0
Homeowner Turnover Rate (-)	2/1	0/0	0/6	0/5
Share Households Credit Score <620 (+)	8/0	6/0	8/0	7/0
Share Households Unscorable (+)	2/1	2/1	1/2	0/3
Average R-squared	0.75	0.63	0.66	0.50

Note: Table shows the number of market areas where the sign of the coefficient is statistically significant and positive to the left of the slash, and the number of market areas where the sign of coefficient is statistically significant and negative to the right of the slash.

Exhibit 10 shows the effect on the magnitude of the coefficients for the racial/ethnic population shares from including credit scores in models predicting the share of loans originated by HUD-identified subprime lenders. In general, the impact of credit scores on changes in the racial/ethnic coefficients is less than for models estimating the share of HMDA higher-priced loans (Exhibit 7). In the refinance market, the average coefficient on the share African-American declines 27 percent, compared to an average of 72 percent for the higher-priced share model; for Hispanics, the average decline is 14 percent compared to

98 percent for the higher-priced share model. In the purchase market, the average coefficient on the share African-American declined by 48 percent, compared to 57 percent in the higher-priced share model; for Hispanics, the average decline in the coefficient is 28 percent compared to 75 percent in the higher-priced share model.

Exhibit 10

Summary of Change in Racial/Ethnic Coefficients from Inclusion of Credit Score Variables in Models of HUD Subprime Lender Shares

Metropolitan Area	African-American			Hispanic		
	Without Credit Scores	With Credit Scores	Percent Change	Without Credit Scores	With Credit Scores	Percent Change
Refinance						
Albuquerque	0.14	0.13	-10%	0.01	0.00	-75%
Baltimore	**0.21**	**0.13**	-36%	0.06	0.14	119%
Harrisburg	**0.25**	**0.18**	-30%	-0.18	-0.11	-41%
Memphis	**0.39**	**0.42**	9%	**0.69**	**0.78**	13%
Omaha	**0.20**	**0.13**	-36%	**0.19**	**0.21**	14%
Providence	0.06	0.00	-105%	**0.16**	-0.01	-107%
Tacoma	**0.19**	0.06	-68%	**0.32**	0.24	-27%
Toledo	**0.21**	**0.11**	-46%	0.08	-0.16	-294%
Tulsa	**0.18**	**0.19**	2%	-0.14	-0.06	-56%
Average	0.20	0.15	-27%	0.13	0.11	-14%
Purchase						
Albuquerque	0.10	0.04	-61%	-0.03	**-0.06**	92%
Baltimore	**0.13**	**0.08**	-42%	**0.24**	**0.27**	14%
Harrisburg	**0.09**	-0.01	-115%	**0.47**	**0.35**	-24%
Memphis	**0.37**	**0.20**	-46%	0.39	0.05	-87%
Omaha	**0.20**	**0.13**	-34%	**0.33**	**0.27**	-16%
Providence	0.20	0.11	-43%	**0.43**	**0.31**	-27%
Tacoma	0.25	0.02	-92%	0.21	0.19	-6%
Toledo	**0.08**	0.05	-46%	-0.06	-0.17	187%
Tulsa	**0.21**	**0.23**	13%	**0.39**	**0.47**	20%
Average	0.18	0.09	-48%	0.26	0.19	-28%

Note: Statistically significant coefficients shown in bold.

In short, models using the HUD lender list to define subprime loans find a larger association between a census tract's African-American or Hispanic population share and the prevalence of subprime lending. The average coefficients for the African-American population share in the HUD subprime lender models are 0.15 for refinance mortgages and 0.09 for purchase mortgages. In comparison, these coefficients in models using the higher-priced loan share average 0.06 and 0.09, respectively. Similarly, the average coefficients on the Hispanic population share in the HUD lender models are 0.11 for refinance mortgages and 0.19 for purchase mortgages, compared to 0.00 and 0.06 in the higher-priced loan share models.

Use of Prime Mortgage Denial Rate in Place of Credit Score

In the absence of credit scores, several studies of subprime lending patterns have created an estimate of neighborhood credit risk based on the denial rate for prime mortgages. To the extent that these denial rates serve as a good proxy for credit scores, this approach has the advantage of making available a credit risk measure for virtually all market areas in all years. As shown in Section 3, the average correlation between tract-level denial rates and average FICO scores is 0.70—relatively high, but still with a significant amount of variation ranging from a low of 0.48 to a high of 0.82. In this section we examine differences in the estimated racial/ethnic association with higher-priced lending when the prime mortgage denial is used in place of credit scores.

As expected, the denial rate is strongly associated with subprime lending shares, with positive and statistically significant coefficients in all but one case. Exhibit 11 compares the estimated coefficients for the African-American and Hispanic population shares in models of higher-priced loan shares alternatively using denial rates and credit scores. In general, models using denial rates produce much larger coefficients on the racial/ethnic population share variables. The average coefficients for the African-American population share are 0.18 in the refinance market and 0.16 in the purchase market, compared to 0.06 and 0.09, respectively, when credit score measures are used. The average coefficients for the Hispanic share are 0.05 in the refinance market and 0.16 in the purchase market, compared to 0.00 and 0.06, respectively, when credit score measures are used. Thus, while the denial rate does provide an indication of a neighborhood's credit risk, it does not provide the same degree of control as credit scores and results in higher estimates of the association between a neighborhood's racial/ethnic composition and higher-priced lending shares.

Exhibit 11

Summary of Change in Racial/Ethnic Coefficients from Using Denial Rates in Place of Credit Scores

Metropolitan Area	African-American			Hispanic		
	With Denial Rates	With Credit Rates	Percent Change	With Denial Rates	With Credit Rates	Percent Change
Refinance						
Albuquerque	0.29	-0.42	-242%	-0.05	0.00	-108%
Baltimore	**0.15**	**0.14**	-4%	-0.11	**-0.17**	57%
Harrisburg	**0.19**	**0.18**	-9%	-0.20	-0.08	-62%
Memphis	**0.22**	**0.26**	19%	0.12	0.15	23%
Omaha	**0.22**	**0.12**	-44%	**0.21**	**0.12**	-40%
Providence	0.06	0.04	-29%	**0.13**	**0.08**	-37%
Tacoma	0.11	0.00	-103%	**0.29**	**0.27**	-8%
Toledo	**0.23**	**0.12**	-48%	**0.47**	0.13	-73%
Tulsa	**0.11**	0.08	-32%	**-0.43**	**-0.50**	14%
Average	0.18	0.06	-68%	0.05	0.00	-96%
Purchase						
Albuquerque	0.29	0.13	-55%	-0.05	**-0.08**	53%
Baltimore	**0.12**	**0.09**	-24%	0.12	0.05	-60%
Harrisburg	**0.11**	0.07	-37%	**0.42**	**0.53**	26%
Memphis	**0.20**	**0.13**	-36%	0.25	**-0.03**	-113%
Omaha	**0.17**	0.04	-74%	**0.23**	0.03	-86%
Providence	0.00	-0.02	-522%	**0.27**	**0.20**	-24%
Tacoma	**0.25**	0.13	-46%	0.20	0.14	-32%
Toledo	**0.18**	**0.12**	-36%	0.15	-0.10	-165%
Tulsa	**0.15**	0.10	-32%	-0.15	-0.24	57%
Average	0.16	0.09	-46%	0.16	0.06	-65%

Note: Statistically significant coefficients shown in bold.

4.2 Borrower-Level Analysis

As described in Section 2, many of the studies examining higher-priced or subprime lending patterns have focused on the shares of these loans aggregated at the census tract level. To the extent that these analyses find disparities in the prevalence of these loans by a neighborhood's racial/ethnic composition this pattern may be due to some combination of lenders targeting these neighborhoods by their marketing efforts as well as discriminatory treatment of individual borrowers regardless of where they live but because of racial/ethnic segregation a neighborhood lending pattern results. Since the HMDA data reports at the borrower level and includes the applicants' race/ethnicity, these data can also be used to control for both neighborhood and borrower race/ethnicity. Several previous studies have estimated borrower-level models of subprime loan choice. In this section we present the results of logit models of the use of a higher-priced loan as a function of both borrower and neighborhood characteristics.

Exhibit 12 presents a summary of the logit model results with and without credit variables. (Full model results are presented in Appendix A.) Given the large number of observations, many of the borrower-level coefficients are highly statistically significant. The pattern of statistical significance of the census tract variables is very similar to the tract-level models, with the share of adults with a college education and the rental share of housing units most strongly associated with higher-priced loans in both the purchase and refinance models, while the homeowner turnover rate is strongly associated with higher-priced loans in the purchase model. But our primary concern is with the coefficients on borrower and neighborhood race/ethnicity. Focusing first on model results without the credit score measures, the results show a very strong association between African-American borrowers and the probability of obtaining a higher-priced loan, as the coefficient on this variable is highly statistically significant and positive in all nine metropolitan areas for both purchase and refinance loans. The coefficients for Hispanic borrowers are also positive and significant in five markets for refinance loans and six markets for purchase loans, although in two markets the coefficient is negative and statistically significant for refinance loans. The coefficients on the tract-level population measures are also positive and significant in five markets for the share African-American for both refinance and purchase loans, and for Hispanics, the coefficients are positive and significant in three markets for refinance loans and four markets for purchase loans.

When the credit score variables are added to the markets there is very little change in either the significance or magnitude of the coefficients on the borrower's race or ethnicity. On average, the coefficients on African-American and Hispanic borrowers change by only 2 percent. There is a much larger impact, however, on the significance and magnitude of the tract-level measures. Once credit scores are added, the tract share African-American is only positive and significant in a single market for refinance loans, while the share Hispanic is only positive and significant in one market for purchase loans. Thus, while the logit models

find significant differences in the probability of obtaining a higher-priced loan based on the borrower's race/ethnicity, after controlling for the neighborhood's credit score there is no residual racial/ethnic neighborhood association.

Exhibit 12

Summary of Regression Results for Borrower-Level Higher-Priced Share Models (Number of Statistically Significant Coefficients that are +/- Across Nine Markets)

Variable (Expected Sign in Parentheses)	Refinance		Purchase	
	Without Credit Scores	With Credit Scores	Without Credit Scores	With Credit Scores
Borrower Characteristics				
African-American (+)	9/0	9/0	9/0	9/0
Race Not Identified (?)	9/0	9/0	8/1	7/1
Hispanic (+)	5/2	5/2	6/0	6/0
Ethnicity Not Identified (?)	2/4	1/4	0/7	0/7
Single Female Borrower (?)	7/0	7/0	9/0	9/0
Single Male Borrower (?)	8/0	8/0	9/0	9/0
Gender Not Identified (?)	2/2	2/2	4/1	4/1
Income LT 80% of AMI (+)	6/1	6/1	5/4	4/5
Income 80–99% of AMI (+)	7/0	7/0	6/1	6/1
Income 100–119% of AMI (+)	8/0	8/0	8/0	8/0
Income Missing (?)	0/9	0/9	1/5	1/5
Census Tract Characteristics				
African-American Pop. Share (+)	5/0	1/1	5/0	0/2
Hispanic Pop. Share (+)	3/1	0/2	4/1	1/4
Other Race Pop. Share (+)	3/1	1/1	4/0	2/1
Tract Median Household Income (Ln) (-)	0/8	0/5	0/6	0/2
Share Pop. With College (-)	0/9	0/8	0/9	0/8
Rental Rate (-)	0/6	0/6	1/4	1/7
Capitalization Rate (+)	5/0	3/0	4/0	3/0
Homeowner Turnover Rate (-)	3/2	1/3	0/8	0/7
Share Households Credit Score <620 (+)	NA	9/0	NA	8/0
Share Households Unscorable (+)	NA	1/2	NA	2/1
Average Pseudo R-squared	0.0968	0.0981	0.0837	0.0857

Note: Table shows the number of market areas where the sign of the coefficient is statistically significant and positive to the left of the slash, and the number of market areas where the sign of coefficient is statistically significant and negative to the right of the slash.

Source: Appendix A, Exhibits A-3 and A-4.

To assess the magnitude of the differences in the probability of obtaining a higher-priced loan due to a borrower's race/ethnicity, we used the estimated logit coefficients to calculate the estimated probability of using a higher-priced loan for three prototypical borrowers.[23] For each borrower we assumed that they were single male borrowers with income between 80 percent and 99 percent of the area median income and that they lived in a census tract with average characteristics for their metropolitan area. The only factor that varied for these borrowers is whether they were white (non-Hispanic), African-American (non-Hispanic), or white Hispanic.

Exhibit 13 presents the increase in the probability of obtaining a higher-priced loan for African-Americans and Hispanics relative to the probability for white non-Hispanics. In general, the logit model results find a much larger association between African-Americans and the probability of obtaining a higher-priced loan. Across all nine market areas, on average African-Americans were 13.1 percentage points more likely to obtain a higher-priced refinance loan and 12.5 percentage points more likely to obtain a higher-priced purchase loan. For Hispanics, the average increase in the probability of higher-priced loans was much lower—2.1 percentage points for refinance loans and 5.3 percentage points for purchase loans. While these increases may not seem large in absolute size, given the much higher foreclosure rate among higher-priced loans, these differences have important implications for the chance that these borrowers will lose their homes.

The magnitude of the association between *borrower* race and, to a lesser extent, ethnicity and the prevalence of higher-priced loans is generally much larger than the association found based solely on *neighborhood* racial/ethnic composition as summarized in Exhibit 7 above. In these borrower-level models, neighborhood race/ethnicity is essentially a nonfactor once neighborhood credit scores are included. These results indicate that the race/ethnicity of the borrower is an important independent factor in the use of higher-priced loans. Neighborhood racial/ethnic composition may still be a factor, but once borrower race and neighborhood credit scores are accounted for—both of which are highly correlated with borrower race—there is no remaining independent effect of the neighborhood's racial/ethnic composition. It is interesting that the inclusion of credit scores has essentially no effect on the magnitude of the borrower coefficients for race/ethnicity. However, it seems quite likely that if borrower-level credit scores were available, these coefficients would also be reduced.

---

[23] In a logit model the dependent variable is the natural logarithm of the odds ratio—that is, the ratio of the probability of the outcome of interest to one minus this probability. To estimate the probability of a higher-priced loan for a prototypical borrower, we multiply the estimated logit coefficients by sample average values of the continuous independent variables along with the stated assumptions about the dummy variables. We then raise the resulting number to the base e to get the odds ratio, and we then convert this into the underlying probability of obtaining a higher-priced loan.

Exhibit 13

Increase in Probability of HMDA Higher-Priced Loan Due to Borrower Race/Ethnicity
(Percentage points)

Metropolitan Area	Refinance		Purchase	
	African-American	Hispanic	African-American	Hispanic
Albuquerque	11.0	4.1	7.2	5.1
Baltimore	11.1	4.8	13.9	7.7
Harrisburg	10.5	13.5	11.2	6.5
Memphis	18.9	-4.6	20.5	6.8
Omaha	14.5	-3.2	10.6	1.6
Providence	6.3	1.3	10.9	9.9
Tacoma	11.5	5.5	13.0	7.8
Toledo	15.3	2.0	6.2	2.0
Tulsa	18.5	-4.9	18.8	0.0
Average	13.1	2.1	12.5	5.3

Notes: Estimates made for single male borrower with income between 80 percent and 99 percent of area median income and at mean of tract characteristics for each market area. Probability is calculated from the log odds ratio predicted by the logit model.

The following coefficients were not statistically significant: Hispanic refinance borrowers in Toledo and Hispanic purchase borrowers in Omaha, Tulsa, and Toledo.

Comparison of Results Using HUD Subprime Lender List and Denial Rates

As with the tract-level analysis, we also explored how the use of the HUD subprime lender list in place of the HMDA higher-priced loan designation affected the estimated racial/ethnic coefficients. In general, the modeling results are quite similar, although the estimated racial/ethnic coefficients predicting the use of a HUD-identified subprime lender are generally larger than the coefficients predicting the use of a higher-priced loan. Exhibit 14 presents estimates of the increase in the probability of obtaining a loan from a HUD subprime lender due to the borrower being African-American or Hispanic. On average, the increases in these probabilities are larger than those shown in Exhibit 13 for higher-priced loans. Thus, as with the tract-level analysis, the use of the HUD subprime lender list produces somewhat higher estimates of racial/ethnic disparities in subprime lending than is found using the HMDA higher-priced loan definition.

Exhibit 14

Increase in Probability of HUD Subprime Lender Loan Due to Borrower Race/Ethnicity
(Percentage points)

Metropolitan Area	Refinance		Purchase	
	African-American	Hispanic	African-American	Hispanic
Albuquerque	11.0	4.1	7.2	5.1
Baltimore	11.1	4.8	13.9	7.7
Harrisburg	10.5	13.5	11.2	6.5
Memphis	18.9	-4.6	20.4	6.7
Omaha	14.5	-3.2	10.6	1.6
Providence	6.3	1.3	10.9	9.9
Tacoma	11.5	5.5	13.0	7.8
Toledo	15.3	2.0	6.2	2.0
Tulsa	18.5	-4.9	18.7	-0.1
Average	13.1	2.1	12.5	5.2

Notes: Estimates made for single male borrower with income between 80 percent and 99 percent of area median income and at mean of tract characteristics for each market area.

The following coefficients were not statistically significant: Hispanic refinance borrowers in Toledo and Hispanic purchase borrowers in Harrisburg and Toledo.

We also estimated models using the denial rate in place of the credit score in the logit models. Once again the denial rate was found to be strongly associated with the use of higher-priced loans, with the coefficient positive and statistically significant in all but one of the estimated models. However, just as the tract-level credit measure has little effect on the estimated borrower racial/ethnic coefficients, the use of the denial rate results in essentially no change in the borrower racial/ethnic coefficients. The coefficients on the neighborhood racial/ethnic composition variables also remain largely insignificant in these models.

Section 5: Summary and Conclusions

Over the past decade, numerous studies have raised concerns that subprime lenders were inappropriately targeting economically vulnerable neighborhoods and that some borrowers, particularly African-Americans, were paying more than necessary for mortgage financing. Indeed, there is a substantial body of literature consistent with the finding that predominantly African-American neighborhoods have much higher shares of loans originated by subprime lenders than is true for areas where whites predominate, even after controlling for a range of factors that are likely to influence the allocation of mortgage credit. Yet, despite the importance of the topic, efforts to explain the underlying causes of the observed patterns have been limited by lack of data on mortgage credit risk.

Using newly available data on neighborhood-level measures of credit scores for each of nine metropolitan areas selected to reflect a range of racial and ethnic population characteristics, this paper has presented a detailed analysis of what has come to be known as the "risk or race" question. In particular, this report summarizes the results obtained from a series of ordinary least squares regressions for models that seek to explain census tract-level variation in the share of higher-priced subprime mortgages for home purchase and refinance.

In addition, the paper reports findings obtained from a series of logit regression models investigating the impact that tract-level mortgage credit scores have on the probability that an individual borrower will obtain a higher-priced mortgage rather than a lower-priced prime mortgage to purchase or refinance a home. Following the path-breaking work of Calem, Gillen, and Wachter (2004), these purchase and refinance models include various measures expected to influence the allocation of mortgage credit, including income and other demographic factors, mortgage credit and other measures of credit risk, as well as various neighborhood-level housing market variables. For both refinance and purchase models, the key focus is on whether and how the inclusion of tract-level credit score information influences the relationship between racial factors and access to prime mortgages.

The paper also explores several additional methodological issues. First, until the recent release of 2004 HMDA data on loan-level mortgage pricing, numerous studies examined the spatial variation of the share of subprime loans based on a HUD-generated list of institutions that engaged primarily in subprime lending. Moreover, prior to recent enhanced availability of census tract-level measures of mortgage credit scores, numerous studies were forced to deploy differing proxy measures for risk factors. Yet even as newer studies now can utilize better measures of subprime loans and risk factors, there is still value in relating the results obtained from these newer studies to previous research. For example, to the extent that the relationship between measures of tract-level patterns of subprime lending based on the HUD list approach and the newly available measures based on the HMDA concept of "higher-priced" mortgages can be established, there will be much greater opportunity for research on

historical patterns of the neighborhood-level increases in the overall growth of subprime lending.

Finally, this study also explores the latest and most advanced econometric techniques to assess the extent to which failure to account for spatial autocorrelation may have biased the estimates of the "race effect" in previous studies. Results presented here suggest that for traditional least squares and logit regression models based on aggregate data, spatial autocorrelation appears to neither bias the estimates of coefficients nor have any important effect on the efficiency of the estimates. Even so, in light of growing evidence that individual subprime lenders may have spatially targeted lower-income and/or minority households for the aggressive marketing of subprime mortgages, further research on this topic seems warranted.

5.1 Findings on the Risk or Race Question

Given the substantial variation across metropolitan areas in racial/ethnic settlement patterns and other characteristics likely to influence the spatial distribution of mortgage credit, the estimates of the coefficients included in any single least squares or logit specification are understandably likely to vary from one metropolitan area to another. It is therefore most significant that the estimates of the coefficients relating to the influence of both tract-level and borrower-level race/ethnicity on access to higher-priced (as opposed to lower-priced) mortgages were remarkably consistent across individual market areas. In general, the inclusion of credit score measures did not explain away the troubling finding that even after years of public policy efforts, race and ethnicity remain important determinants of the allocation of mortgage credit in both home purchase and home refinance markets.

Admittedly, since the tract-level variation in mortgage credit scores and the African-American share of the population living in a particular neighborhood are themselves correlated, the inclusion of the credit score generally reduced the magnitude of the coefficient for both variables. Yet, even after accounting for the impact in the spatial variation in credit scores, the African-American share of the census tract population still had a positive and significant impact on the spatial variation in the tract-level share of higher-priced subprime refinance loans in five of the nine metropolitan areas examined.

For models estimating the higher-priced share of home purchase mortgages for African-Americans, the impact of including credit scores was similar. Once again, even after accounting for the impact in the spatial variation in credit scores, in three of the nine metropolitan areas the coefficient on African-American tract-level population share was positive and significant. Notably, these positive and significant tract-level coefficients were recorded in Baltimore, Memphis, and Toledo, market areas with relatively high overall shares of African-American population.

Consistent with previous studies, the coefficients on Hispanic population share were generally smaller and recorded lower levels of statistical significance. For six of nine market areas, the coefficient for the Hispanic population share variable in the home refinance equation had the expected positive sign, even after the inclusion of the tract-level credit score measure. Yet notably, the coefficient on Hispanic population share was positive and statistically significant in only three markets. For the home purchase equations that included credit score measures, the coefficient on the Hispanic share of population was positive and significant in two instances.

Logit models on the probability that a particular borrower will obtain a higher-cost subprime mortgage provided additional insight on racial and ethnic differences in the access to mortgage credit. Before the inclusion of credit score information, the logit regression results showed a very strong association between African-American borrowers and the probability of obtaining a higher-priced loan, as the coefficient on this variable was highly statistically significant and positive in all nine metropolitan areas for both purchase and refinance loans. When the neighborhood credit score variables were added to the models, there was very little change in either the significance or magnitude of the coefficients on the borrower's race or ethnicity.

There was a much larger impact, however, on the significance and magnitude of the tract-level measures. Once credit scores were added, the tract share African-American was positive and significant in only one market for refinance loans, while the share Hispanic remained positive and significant in one market for purchase loans. Thus, while the logit models found significant differences in the probability of obtaining a higher-priced loan based on the *borrower's* race/ethnicity, after controlling for the neighborhood's credit score there remained no residual racial/ethnic *neighborhood* association.

The finding that neighborhood racial and ethnic composition is often not associated with variations in subprime lending once individual borrower race/ethnicity and neighborhood credit scores are accounted for likely reflects differences in both market circumstances and econometric issues. For example, in each of the nine metropolitan areas, there was substantial correlation between tract-level measures of credit score and African-American population shares. Similarly, given the intense spatial segregation of the African-American population, there was also a high correlation between tract-level African-American population share and the probability that an individual African-American borrower lives in a largely African-American neighborhood. As a result, even though racial disparities in lending may involve both a borrower-level dimension and a tract-level dimension, especially since lenders may use tract-level measures to target the marketing of subprime products, it is difficult to separate out individual from neighborhood-level effects.

Overall, these findings confirm that the omission of mortgage credit score information does, in fact, bias estimates of the effect of race and ethnicity on access to mortgage credit,

particularly for lower-priced subprime loans. In general, larger correlations of race and subprime lending are found in the refinance market and are associated with the African-American population share. For refinance markets, the higher-priced share of mortgages is estimated to increase from 4.3 to 18.9 percentage points with a two standard deviation increase in the African-American population share across the four market areas with a statistically significant association with this variable. For the Hispanic population share, the estimated increase is generally fairly small, with three of the markets having an increase of only 1.9 to 2.2 percentage points in the higher-priced share from a two standard deviation increase in the Hispanic population share. However, in Harrisburg there is an estimated 7 percentage point increase associated with a two standard deviation change in the Hispanic share of the population in the tract.

In general, our findings support the results found in previous studies by Calem, Gillen, and Wachter (2004) and Calem, Hershaff, and Wachter (2004). These studies concluded that even after accounting for neighborhood credit scores along with a variety of borrower and neighborhood characteristics, significant racial and ethnic disparities remained in the use of subprime mortgage lending. In contrast to our findings, these earlier studies did find that the neighborhood share of African-Americans was significantly and positively associated with subprime lending shares even after accounting for the individual borrower's race and ethnicity and neighborhood credit risks. In part, this difference from our results may reflect the fact that these earlier studies relied on HUD's list of subprime lenders, which we have found to have a somewhat stronger association with racial/ethnic disparities in lending. The markets included in our study are also smaller and generally have smaller shares of African-Americans than those included in these earlier studies, and so there may be a different market dynamic at work. In that regard, it is notable that our study focusing on a very different set of markets compared to earlier studies still came to very similar conclusions regarding the existence of racial and ethnic disparities in subprime lending even after controlling for neighborhood credit scores. This finding suggests that the lending patterns observed may be fairly common across a wide range of U.S. markets.

Since we are unable to fully account for borrower-level risk measures that are known to affect a borrower's ability to access prime mortgages, our finding of racial and ethnic disparities in the use of higher-priced mortgage financing cannot be taken as evidence of discriminatory treatment. Yet the results are consistent with the possibility that racial and ethnic minorities have been subject to discriminatory treatment. In any event, the growth of subprime lending has disproportionately affected minority and low-income communities. While the mortgage market meltdown is adversely impacting both the prime and subprime markets, best estimates suggest that the probability of entering into foreclosure is as much as 10 times greater for a subprime mortgage than for a prime mortgage. As a result, to the extent that mortgage market discrimination may have added—even on the margin—to the concentration of higher-priced subprime loans in minority neighborhoods, the consequences

of this discriminatory pattern may be coming home to roost in the form of elevated foreclosures that are undermining the stability and vitality of these areas today.

5.2 Methodological Issues and Future Research

The ongoing mortgage crisis raises questions about whether mortgage market regulations are sufficient to protect consumers from the aggressive marketing practices of subprime lenders, and whether the allocation of subprime loans was in accord with what appropriate adherence to "fair lending" standards would support. Unfortunately, a lack of detailed data on current patterns and the historical development of subprime lending makes it difficult to place recent events into a longer perspective.

In an effort to create a bridge of understanding between older and newer studies on the topic, this report presented results from models of the neighborhood variation in subprime lending using two alternative measures of subprime mortgages: the first based on the HUD-developed list of subprime lenders, and the second based on the more recently available HMDA information on whether or not a loan is "higher-priced"; namely, has a mortgage rate in excess of the established benchmark. The results confirmed the readily admitted and somewhat obvious conclusion that the new data represent a superior measure of subprime lending. This is understandable since the HMDA-based measure reflected loan-level pricing information. In contrast, since most lenders offer a range of both prime and subprime products, the HUD measure of subprime lending combines two types of measurement errors. First, it erroneously classifies as subprime those prime loans made by institutions that on average specialize in subprime lending. In addition, it erroneously classifies as prime loans those subprime loans made by entities that predominately engage in prime lending, but can and often do account for a large share of subprime activity in any single market area.

Yet despite these issues, and in tribute to the ingenuity of the creators of the HUD approach, the equations based on the HUD list mirror those using the more refined measures based on the newly available HMDA pricing information. As a result, the findings suggest that while historical studies based on the HUD approach are not precisely comparable to newer studies, assessing older and newer studies can help place more recent trends into an appropriate historical context.

Moreover, the results suggest the importance of conducting further research on exactly what explains the observed differences between the two measures of subprime lending. For example, although the HUD approach may present a less than perfect measure of loan-level subprime activity, examining patterns of lending at the institutional level remains important. By assembling additional data on lending patterns of specific institutions (e.g., the information available on aggregate lender activity published by *Inside Mortgage Finance*), it would be possible to subdivide the HUD list into large, medium, and smaller sized subprime

lending specialists, as well as those specialists that are independent mortgage companies as opposed to those that are subsidiaries or affiliates of more traditional banking operations. Such an exercise would not only represent a more appropriate use of the HUD-list approach, it would also support a potentially most revealing lender-level analysis of the change of subprime lending patterns across neighborhood of varying incomes and racial compositions.

Similarly, the results clearly demonstrate that tract-level mortgage credit scores offer a superior measure of tract-level mortgage risk factors when compared with measures based on mortgage rejection rates. Given this, HUD should work diligently to expand the availability of credit score measures at the tract—and even borrower—level. Yet recognizing that these credit score measures as of now are not widely available and/or are costly to acquire, the results presented here affirm that risk measures based on mortgage denial rates still can play a valuable role in helping neighborhood housing and advocacy organizations monitor recent trends and place these trends into a longer-term historical context.

Finally, the results relating to spatial autocorrelation merit continued scrutiny, especially as new econometric techniques are developed and made widely available. In the equations presented in this report, concerns about the potential for bias resulting from spatial autocorrelation do not appear to be warranted. Detailed examination of this issue found that in about half of the cases, spatial autocorrelation was a potential issue. But further examination suggested that failure to correct for serial autocorrelation did not bias coefficient estimates to any significant degree.

While this report suggests that spatial autocorrelation did not in any significant or important way bias the estimates of the "race effect" in the equations presented here, examination of issues of spatial autocorrelation is a newly emerging field of econometric research that may warrant additional study. For example, as noted earlier, the issue of spatial autocorrelation may very well prove to be important in a more detailed assessment of the lending patterns of particular subprime lenders, their contributions to the recent surge in subprime foreclosures, and the adverse implications these foreclosures are having for areas with historically elevated levels of subprime lending.

Given the historically contentious nature of the assessment of any matter involving potential racial or ethnic discrimination, the analysis presented here will not quiet the debate on the extent to which racial and ethnic disparities in the allocation of subprime lending were due to discriminatory treatment by lenders. Having employed a variety of models that utilize alternative definitions of subprime loans and alternative credit risk measures, as well as having deployed best available econometric techniques for assessing potential spatial autocorrelation bias, the research results of this study are presented in the hope of casting a new light on existing research and contributing to the ongoing national debate on how best to ensure that all households have access to mortgage capital on a fair basis and that the recent excesses in mortgage lending do not reappear in the future.

References

ACORN. 2002. Separate and Unequal 2002: Predatory Lending in America. Philadelphia, PA: ACORN.

Anselin, Luc 1995. Local Indicators of Spatial Association – LISA. Geographical Analysis, 27(2): 93-116.

Anselin, Luc and Daniel A. Griffith. 1988. Do Spatial Effects Really Matter in Regression Analysis? Papers of the Regional Science Association, 65: 11-34.

Apgar, William C. Jr. and Christopher E. Herbert, "Subprime Lending and Alternative Financial Service Providers: A Literature Review and Empirical Analysis," U.S. Department of Housing and Urban Development, March 2006.

Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner. 1997. The Distribution of Credit Scores: Findings and Implications for the Provision of Financial Services, paper presented at 33rd Annual Conference on Bank Structure and Competition, Federal Reserve Bank of Chicago, May 1997.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner. 2006. Higher-Priced Home Lending and the 2005 HMDA Data. Federal Reserve Bulletin, September 2006.

Bailey, Trevor and Anthony Gatrell. 1995. Interactive Spatial Data Analysis. Essex, England: Longman Group Limited.

Black, Harold A. and Alan Schlottmann. 2006. An Analysis of Unfair Lending: The Effect of Race and Ethnicity on the Price of Subprime Mortgages. Paper prepared for the American Financial Services Association.

Bradford, Calvin. 2002. Risk or Race: Racial Disparities and the Subprime Refinance Market. Washington, DC: Center for Community Change.

Calem, Paul S., Kevin Gillen, and Susan Wachter. 2004. "The Neighborhood Distribution of Subprime Mortgage Lending," Journal of Real Estate Finance and Economics 29(4): 393-410.

Calem, Paul S., Jonathan E. Hershaff, and Susan M. Wachter. 2004. "Neighborhood Patterns of Subprime Lending: Evidence from Disparate Cities," Housing Policy Debate 15(3): 603-622.

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

Ernst, Keith. 2005. Borrowers Gain No Interest Rate Benefits from Prepayment Penalties on Subprime Mortgages. Durham, NC: Center for Responsible Lending.

Fishbein, Allen and Harold L. Bunce. 2000. "Subprime Market Growth and Predatory Lending" paper presented at a conference titled Housing Policy in the New Millennium, October 2000.

Griffith, Daniel. 1987. Spatial Autocorrelation: A Primer. Washington, DC: Association of American Geographers.

Gruenstein-Bocian, D., Ernst, Keith and Li, Wei 2006. Unfair Lending: The Effect of Race and Ethnicity on the Price of Subprime Mortgages. Washington, DC: Center for Responsible Lending, 31.

Lax, Howard, Michael Manti, Paul Raca, and Peter Zorn. 2000. "Subprime Lending: An Investigation of Economic Efficiency," Freddie Mac Working Paper, February 25, 2000.

Joint Center for Housing Studies, Harvard University. 2004. "Credit Capital and Communities," a report prepared for the Ford Foundation.

Joint Center for Housing Studies, Harvard University. 2008. America's Rental Housing: The Key to a Balanced National Policy.

National Community Reinvestment Coalition. 2003. The Broken Credit System: Discrimination and Unequal Access to Affordable Loans by Race and Age.

Pennington-Cross, Anthony, Anthony Yezer, and Joseph Nichols. 2000. Credit Risk and Mortgage Lending: Who Uses Subprime and Why? Research Institute for Housing America, Working Paper No. 00-03, September 2000.

Pennington-Cross, Anthony and Joseph Nichols. 2000. "Credit History and the FHA-Conventional Choice." Real Estate Economics 28(2): 307-336.

Scheessele, Randall M. 1999. "1998 HMDA Highlights," Housing Finance Working Paper Series, HF-009, U.S. Department of Housing and Urban Development, October 1999.

Scheessele, Randall M. 2002. "Black and White Disparities in Subprime Mortgage Refinance Lending," Housing Finance Working Paper Series, HF-014, U.S. Department of Housing and Urban Development, April 2002.

Staten, Michael. 2005. The New HMDA Pricing Data: What Can They Tell Us About Pricing Fairness? Credit Research Center, McDonough School of Business, Georgetown University.

Wachter, Susan M., Karl Russo, and Jonathan Hershaff. 2005. "Subprime Lending: Neighborhood Patterns Over Time and in US Cities," University of Pennsylvania Law School, Institute for Law and Economics, Research Paper No. 06-19, July 2006.

Wallace, George, Gregory Elliehausen, and Michael Staten. 2005. "Are Relative Solutions to Abusive Mortgage Lending Practices Throwing Out the Baby with the Bathwater?" Guidance from Empirical Research. Draft paper presented at the 41st Annual Conference on Bank Structure and Competition, Federal Reserve Bank of Chicago, May 2005.

Appendix A: Details of Econometric Results

Exhibit A-1
Summary of Tract-Level Regression Results for HMDA Higher-Priced Refinance Mortgage Shares

	Albuquerque				Baltimore				Harrisburg				Memphis				Omaha			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Black, Non-Hispanic (%)	-0.42		-0.40		**0.14**	***	**0.22**	***	**0.18**	**	**0.27**	***	**0.26**	***	**0.29**	***	**0.12**	***	**0.27**	***
Hispanic (%)	0.00		0.03		**-0.17**	*	-0.16		-0.08		-0.05		0.15		0.18		**0.12**	*	**0.26**	***
Other Race, Non-Hispanic (%)	**-0.14**	*	-0.10		0.07		**0.11**	*	0.11		0.25		-0.40		-0.37		-0.17		-0.19	
Tract Median Household Income (Ln)	**-5.75**	*	**-8.78**	***	**-6.73**	***	**-8.51**	***	**-8.06**	*	**-12.06**	***	**-17.49**	***	**-16.64**	***	-5.19		**-11.55**	***
Share Pop. With College	-0.10		-0.18		**-0.16**	***	**-0.31**	***	-0.07		-0.17		-0.08		**-0.16**	*	**-0.28**	***	**-0.36**	***
Rental Rate	-0.05		**-0.05**	*	**-0.07**	***	**-0.06**	***	**-0.10**	***	**-0.09**	**	**-0.14**	***	**1.74**	***	**-0.06**	**	-0.03	
Capitalization Rate	**4.25**	*	**4.72**	**	**7.36**	***	**9.75**	***	**9.55**	**	**10.21**	**	0.26		-0.13		3.63		**6.35**	***
Homeowner Turnover Rate	0.02		0.02		**-0.16**	***	**-0.10**	***	-0.05		-0.08		**-0.14**	***	**-0.13**	***	-0.03		-0.02	
Share Households Credit Score <620	**0.21**	***			**0.31**	***			**0.39**	***			0.10				**0.33**	***		
Share Households Unscorable	-0.01				**-0.06**	*			-0.21				-0.08				**0.18**	*		
Intercept	**72.15**	**	**110.31**	***	**84.67**	***	**109.77**	***	**95.25**	**	**144.12**	***	**213.66**	***	**205.96**	***	**71.86**	**	**149.49**	***
R-SQ	0.56		0.54		0.83		0.83		0.76		0.73		0.85		0.85		0.80		0.77	
No. Observations	185				610				111				279				237			

	Providence				Tacoma				Toledo				Tulsa			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Black, Non-Hispanic (%)	0.04		0.08		0.00		0.11		**0.12**	***	**0.25**	***	0.08		**0.17**	***
Hispanic (%)	**0.08**	**	**0.17**	***	**0.27**	**	**0.31**	**	0.13		**0.51**	***	**-0.50**	***	**-0.41**	***
Other Race, Non-Hispanic (%)	-0.05		0.02		0.09		**0.17**	**	**0.46**	*	**0.72**	***	-0.02		0.08	
Tract Median Household Income (Ln)	-0.43		**-3.37**	*	-4.53		**-6.95**	**	0.46		-3.28		**-20.95**	***	**-23.12**	***
Share Pop. With College	**-0.19**	***	**-0.25**	***	**-0.13**	*	**-0.32**	***	**-0.33**	***	**-0.43**	***	-0.09		-0.15	
Rental Rate	**-0.08**	***	**-0.03**	*	**-0.10**	***	**-0.09**	***	-0.05		-0.02		**-0.17**	***	**-0.12**	***
Capitalization Rate	**15.34**	***	**16.16**	***	-0.90		0.34		**7.31**	*	**17.63**	***	-2.24		0.17	
Homeowner Turnover Rate	**0.14**	*	0.28		0.01		0.05		-0.22		**-0.31**	*	-0.10		-0.16	
Share Households Credit Score <620	**0.24**	***			**0.33**	***			**0.40**	***			**0.26**	***		
Share Households Unscorable	**0.18**	**			0.00				0.10				0.07			
Intercept	7.07		**43.82**	**	**58.34**	*	**93.21**	***	2.80		50.13		**250.50**	***	**280.43**	***
R-SQ	0.76		0.74		0.73		0.66		0.83		0.80		0.63		0.62	
No. Observations	347				154				172				264			

Notes: Coefficients in bold are statistically significant; *** denotes coefficients statistically significant at 0.01 level; ** denotes coefficients statistically significant at 0.05 level; and * denotes coefficients statistically significant at 0.10 level.

Exhibit A-2
Summary of Tract-Level Regression Results for HMDA Higher-Priced Purchase Mortgage Shares

	Albuquerque				Baltimore				Harrisburg				Memphis				Omaha			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Black, Non-Hispanic (%)	0.13		0.19		**0.09**	***	**0.18**	***	0.07		**0.19**	***	**0.13**	*	**0.33**	*	0.04		**0.24**	***
Hispanic (%)	**-0.08**	**	-0.04		0.05		0.07		**0.53**	**	**0.59**	***	**-0.03**	**	**0.42**	**	0.03		**0.30**	***
Other Race, Non-Hispanic (%)	-0.09		-0.06		**0.13**	*	**0.18**	***	-0.15		0.12		**0.49**	***	**0.61**	***	-0.17		-0.31	
Tract Median Household Income (Ln)	-0.82		-4.28		0.79		-0.75		-0.98		-6.26		**-6.13**	***	-4.18		0.50		**-8.44**	**
Share Pop. With College	**-0.26**	**	**-0.36**	***	**-0.13**	***	**-0.32**	***	-0.01		-0.14		**0.08**	*	**-0.30**	***	**-0.26**	***	**-0.28**	***
Rental Rate	-0.04		-0.04		**-0.04**	***	**-0.03**	*	-0.07	*	-0.05		-0.14		**11.97**	***	**-0.08**	***	-0.04	
Capitalization Rate	0.57		-0.84		**4.81**	***	**7.92**	***	2.55		2.71		2.81		**-0.09**	***	**5.64**	***	**7.76**	***
Homeowner Turnover Rate	**-0.06**	***	**-0.06**	***	**-0.18**	***	**-0.14**	***	-0.25		**-0.34**	*	**-0.06**	**	**-0.10**	***	**-0.05**	**	-0.03	
Share Households Credit Score <620	**0.25**	***			**0.36**	***			**0.51**	***			0.58				**0.21**	**		
Share Households Unscorable	0.04				**-0.08**	**			**-0.28**	*			-0.04				**0.54**	***		
Intercept	21.53		**67.62**	**	-3.64		19.98		17.92		82.85		**65.40**	*	58.87		0.31		**106.82**	***
R-SQ	0.41		0.38		0.69		0.65		0.70		0.64		0.81		0.77		0.72		0.66	
No. Observations	185				610				111				279				237			

	Providence				Tacoma				Toledo				Tulsa			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Black, Non-Hispanic (%)	-0.02		0.05		0.13		**0.25**	***	**0.12**	**	**0.23**	***	0.10		**0.20**	***
Hispanic (%)	**0.20**	***	**0.32**	***	0.14		0.21		-0.10		0.26		-0.24		-0.13	
Other Race, Non-Hispanic (%)	**0.15**	*	**0.27**	***	0.08		0.12		-0.16		0.02		0.11		0.22	
Tract Median Household Income (Ln)	1.44		-3.65		-3.79		**-6.51**	*	2.79		0.34		**-13.16**	***	**-14.74**	***
Share Pop. With College	**-0.30**	***	**-0.45**	***	**-0.23**	***	**-0.38**	***	**-0.24**	**	**-0.32**	***	-0.17		**-0.21**	*
Rental Rate	**-0.08**	***	-0.01		**-0.08**	***	**-0.07**	**	0.00		0.02		**-0.13**	**	-0.08	
Capitalization Rate	**17.40**	***	**18.83**	***	-1.03		-1.17		3.57		**10.42**	***	-1.32		0.81	
Homeowner Turnover Rate	-0.03		0.14		**-0.07**	**	**-0.07**	**	**-0.29**	**	**-0.39**	***	-0.12		**-0.16**	*
Share Households Credit Score <620	**0.52**	***			**0.21**	***			**0.29**	***			**0.19**	*		
Share Households Unscorable	0.11				0.08				0.12				0.14			
Intercept	-15.13		**49.71**	*	54.42		**92.41**	**	-23.64		8.94		**160.19**	***	**182.84**	***
R-SQ	0.80		0.77		0.68		0.65		0.66		0.63		0.50		0.49	
No. Observations	347				154				172				264			

Notes: Coefficients in bold are statistically significant; *** denotes coefficients statistically significant at 0.01 level; ** denotes coefficients statistically significant at 0.05 level; and * denotes coefficients statistically significant at 0.10 level.

Exhibit A-3
Summary of Borrower-Level Logistic Regression Results for HMDA Higher-Priced Refinance Mortgage Shares

	Albuquerque				Baltimore				Harrisburg				Memphis				Omaha			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Borrower Characteristics																				
African-American	0.633	***	0.630	***	0.581	***	0.589	***	0.577	***	0.585	***	0.797	***	0.799	***	0.718	***	0.726	***
	0.382	***	0.376	***	-0.175	**	-0.183	***	-0.640	**	-0.654	**	-0.242		-0.236		0.309	*	0.318	**
Race Not Identified	0.263	***	0.260	***	0.747	***	0.752	***	0.368	***	0.366	***	1.049	***	1.048	***	0.939	***	0.938	***
Hispanic	0.266	***	0.269	***	0.271	***	0.268	***	0.717	***	0.741	***	-0.228		-0.220		-0.203	*	-0.199	*
Ethnicity Not Identified	0.290	***	0.290	***	-0.142	***	-0.146	***	-0.147		-0.148		-0.401	***	-0.400	***	-0.224	***	-0.225	***
Single Female Borrower	0.101	*	0.102	*	0.278	***	0.275	***	0.087		0.089		0.119	***	0.119	***	0.227	***	0.229	***
Single Male Borrower	0.180	***	0.178	***	0.301	***	0.295	***	0.223	***	0.229	***	0.072	*	0.073	*	0.246	***	0.250	***
Gender Not Identified	-0.013		0.002		-0.216	***	-0.229	***	-0.029		-0.021		-0.127		-0.124		-0.040		-0.034	
Income LT 80% of AMI	0.243	***	0.258	***	0.397	***	0.420	***	0.586	***	0.591	***	0.625	***	0.628	***	-0.010		-0.003	
Income 80–99% of AMI	0.392	***	0.409	***	0.347	***	0.365	***	0.397	***	0.397	***	0.346	***	0.351	***	-0.099		-0.104	
Income 100–119% of AMI	0.221	***	0.230	***	0.278	***	0.290	***	0.281	***	0.282	***	0.269	***	0.274	***	0.226	**	0.233	**
Income Missing	-2.890	***	-2.848	***	-1.808	***	-1.780	***	-1.983	***	-1.981	***	-2.347	***	-2.337	***	-1.979	***	-1.972	***
Census Tract Characteristics																				
African-American Pop. Share	-0.026		-0.022		0.003	***	0.008	***	0.006		0.010	***	0.001		0.005	***	-0.003		0.003	*
Hispanic Pop. Share	-0.005	**	0.003		-0.002		-0.005		-0.002		-0.003		0.000		0.006		0.005		0.007	**
Other Race Pop. Share	-0.013	**	-0.008	*	0.006		0.010	***	0.002		0.016		-0.015		-0.010		-0.006		-0.002	
Tract Median Household Income (Ln)	-0.385	*	-0.643	***	-0.415	***	-0.455	***	-0.471		-0.647	**	-0.770	***	-0.701	***	-0.585	***	-0.843	***
Share Pop. With College	-0.017	*	-0.028	***	-0.019	***	-0.030	***	-0.008		-0.016	**	-0.013	***	-0.022	***	-0.021	***	-0.026	***
Rental Rate	-0.003		-0.005	**	-0.004	***	-0.003	***	-0.005	*	-0.004		-0.007	***	-0.006	***	-0.003		-0.001	
Capitalization Rate	0.195		0.178		0.423	***	0.578	***	0.689	***	0.716	***	-0.082		0.139		0.111		0.269	*
Homeowner Turnover Rate	0.004	**	0.005	***	-0.004	**	0.000		0.000		-0.005		-0.006	***	-0.006	***	-0.003		-0.002	
Share Households Credit Score <620	0.024	***			0.022	***			0.025	***			0.013	***			0.018	***		
Share Households Unscorable	-0.009				-0.009	***			-0.018	*			-0.003				0.000			
Intercept	2.000		5.390	***	2.210	**	2.979	***	2.578		4.794		6.938	***	6.474	***	4.871	**	8.043	***
Psuedo R-squared	0.081		0.079		0.110		0.109		0.076		0.075		0.193		0.192		0.098		0.097	
No. Observations	18,180				91,975				12,796				24,178				18,547			

Exhibit A-3 *(Continued)*
Summary of Borrower-Level Logistic Regression Results for HMDA Higher-Priced Refinance Mortgage Shares

	Providence				Tacoma				Toledo				Tulsa			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Borrower Characteristics																
African-American	0.412	***	0.415	***	0.631	***	0.633	***	0.713	***	0.708	***	0.804	***	0.813	***
	-0.137		-0.138		0.215	***	0.215	***	0.156		0.139		0.188	**	0.191	**
Race Not Identified	0.613	***	0.616	***	0.609	***	0.600	***	0.296	***	0.292	***	0.559	***	0.557	***
Hispanic	0.095	*	0.102	*	0.330	***	0.311	***	0.107		0.105		-0.259	*	-0.256	*
Ethnicity Not Identified	0.012		0.013		-0.075		-0.065		0.149	**	0.156	**	-0.411	***	-0.412	***
Single Female Borrower	0.498	***	0.498	***	0.157	***	0.152	***	0.369	***	0.372	***	0.042		0.041	
Single Male Borrower	0.523	***	0.525	***	0.232	***	0.230	***	0.258	***	0.266	***	0.034		0.034	
Gender Not Identified	0.246	***	0.247	***	-0.351	***	-0.372	***	0.411	***	0.409	***	-0.032		-0.033	
Income LT 80% of AMI	0.041		0.046		-0.239	***	-0.247	***	0.502	***	0.512	***	0.666	***	0.670	***
Income 80-99% of AMI	0.245	***	0.246	***	0.063		0.056		0.369	***	0.371	***	0.420	***	0.420	***
Income 100-119% of AMI	0.214	***	0.215	***	-0.001		0.001		0.230	***	0.229	***	0.255	***	0.256	***
Income Missing	-0.937	***	-0.936	***	-2.478	***	-2.456	***	-1.606	***	-1.592	***	-2.147	***	-2.146	***
Census Tract Characteristics																
African-American Pop. Share	-0.001		0.002		-0.007		0.000		-0.004	*	0.003	*	-0.002		0.001	
Hispanic Pop. Share	-0.001		0.004	**	0.009		0.009		-0.006		0.014	*	-0.023	***	-0.017	***
Other Race Pop. Share	-0.005		0.000		0.004		0.011	***	0.039	***	0.055	***	-0.003		0.000	
Tract Median Household Income (Ln)	-0.134		-0.335	***	-0.207		-0.404	**	0.166		-0.106		-1.223	***	-1.322	***
Share Pop. With College	-0.021	***	-0.027	***	-0.015	***	-0.030	***	-0.032	****	-0.038	***	-0.009	*	-0.012	**
Rental Rate	-0.006	***	-0.002	*	-0.006	***	-0.005	***	0.001		0.002		-0.009	***	-0.007	***
Capitalization Rate	1.321	***	1.392	***	-0.024		0.108		0.197		0.712	***	-0.185		-0.087	
Homeowner Turnover Rate	0.005		0.014	***	0.003		0.007	***	-2.399	**	-2.478	**	0.000		-0.002	
Share Households Credit Score <620	0.019	***			0.029	***			0.023	***			0.011	***		
Share Households Unscorable	0.010	**			-0.005				0.003				0.000			
Intercept	-1.624		0.926		0.089		2.878		-4.018	*	-0.783		11.971	***	13.309	***
Psuedo R-squared	0.062		0.061		0.057		0.055		0.106		0.105		0.099		0.099	
No. Observations	59,909				26,904				14,767				15,830			

Notes: *** denotes coefficients statistically significant at 0.01 level; ** denotes coefficients statistically significant at 0.05 level; and * denotes coefficients statistically significant at 0.10 level.

Exhibit A-4
Summary of Borrower-Level Logistic Regression Results for HMDA Higher-Priced Purchase Mortgage Shares

	Albuquerque				Baltimore				Harrisburg				Memphis				Omaha			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Borrower Characteristics																				
African-American	0.426	***	0.433	***	0.839	***	0.875	***	0.677	***	0.714	***	0.920	***	0.960	***	0.674	***	0.666	***
	0.114		0.121		0.070		0.064		-0.062		-0.088		0.094		0.083		-0.374	**	-0.379	**
Race Not Identified	-0.303	***	-0.302	***	0.602	***	0.612	***	0.684	***	0.699	***	1.045	***	1.062	***	0.972	***	0.970	***
Hispanic	0.312	***	0.313	***	0.513	***	0.514	***	0.425	**	0.446	**	0.336	***	0.364	***	0.121		0.094	
Ethnicity Not Identified	-0.210	***	-0.213	***	-0.332	***	-0.337	***	-0.092		-0.094		-0.390	***	-0.393	***	-0.561	***	-0.565	***
Single Female Borrower	0.686	***	0.676	***	0.709	***	0.705	***	0.462	***	0.468	***	0.572	***	0.574	***	0.704	***	0.704	***
Single Male Borrower	0.788	***	0.784	***	0.846	***	0.842	***	0.370	***	0.380	***	0.573	***	0.572	***	0.743	***	0.745	***
Gender Not Identified	0.803	***	0.801	**	-0.058		-0.063		-0.618	**	-0.613	**	0.207	**	0.214	**	0.961	***	0.963	***
Income LT 80% of AMI	-0.249	***	-0.221	***	-0.308	***	-0.258	***	0.327	***	0.342	***	-0.139	***	-0.100	**	0.575	**	0.614	*
Income 80-99% of AMI	0.174	***	0.196	***	-0.013		0.011		0.409	***	0.399	***	0.093	**	0.125	**	0.411		0.442	
Income 100-119% of AMI	0.367	***	0.381	***	0.087	**	0.102	**	0.429	***	0.416	***	0.233	***	0.257	***	0.388	***	0.388	***
Income Missing	-0.739	***	-0.724	***	0.120	**	0.138	**	0.044		0.067		-0.434	***	-0.410	***	0.365		0.400	
Census Tract Characteristics																				
African-American Pop. Share	0.015		0.019		-0.001		0.005	***	-0.001		0.007	*	-0.004	***	0.006	***	-0.005	*	0.002	
Hispanic Pop. Share	-0.011	***	0.008	***	0.014	**	0.012	**	0.020		0.025	**	-0.015	*	0.005		-0.005		0.005	
Other Race Pop. Share	-0.012	**	-0.008		0.012	**	0.018	***	-0.024		-0.006		0.031	***	0.038	***	-0.003		-0.008	
Tract Median Household Income (Ln)	-0.126		-0.447	**	0.034		0.017		-0.180		-0.428		-0.584	***	-0.413	***	-0.388		-0.968	***
Share Pop. With College	-0.033	***	-0.043	***	-0.022	***	-0.038	***	-0.006		-0.017	*	-0.012	***	-0.033	***	-0.029	***	-0.030	***
Rental Rate	-0.004	**	-0.005	**	-0.002	**	-0.001		-0.003		-0.001		-0.009	***	-0.005	***	-0.005	**	-0.003	
Capitalization Rate	0.002		-0.108		0.233	***	0.441	***	0.049		0.110		-0.093		0.399	***	0.336	**	0.501	***
Homeowner Turnover Rate	-0.006	***	-0.006	***	-0.013	***	-0.009	***	-0.023		-0.027	*	-0.007	***	-0.009	***	-0.008	***	-0.006	***
Share Households Credit Score <620	0.024	***			0.028	***			0.033	***			0.031	***			0.008			
Share Households Unscorable	0.003				-0.012	***			-0.016				-0.003				0.027	***		
Intercept	-0.697		3.558	**	-3.163	**	-2.526		-0.714		2.485		3.786	***	2.660	***	1.281		7.939	***
Psuedo R-squared	0.051		0.050		0.094		0.092		0.063		0.061		0.140		0.137		0.114		0.112	
No. Observations	22,293				63,441				10,125				27,951				16,615			

Exhibit A-4 *(Continued)*
Summary of Borrower-Level Logistic Regression Results for HMDA Higher-Priced Purchase Mortgage Shares

	Providence				Tacoma				Toledo				Tulsa			
	With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores		With Scores		Without Scores	
Borrower Characteristics																
African-American	0.624	***	0.641	***	0.648	***	0.656	***	0.354	***	0.356	***	0.834	***	0.836	***
	0.062		0.066		0.036		0.040		0.073		0.055		-0.020		-0.017	
Race Not Identified	0.263	***	0.267	***	0.120		0.130	*	0.886	***	0.885	***	0.344	***	0.342	***
Hispanic	0.573	***	0.584	***	0.412	***	0.439	***	0.119		0.123		-0.005		-0.004	
Ethnicity Not Identified	-0.028		-0.026		-0.111	*	-0.118	*	-0.195	**	-0.195	**	-0.494	***	-0.490	***
Single Female Borrower	0.900	***	0.904	***	0.428	***	0.428	***	0.537	***	0.546	***	0.427	***	0.425	***
Single Male Borrower	0.901	***	0.906	***	0.538	***	0.524	***	0.484	***	0.489	***	0.411	***	0.413	***
Gender Not Identified	0.505	***	0.510	***	-0.065		-0.075		0.258		0.266		0.041		0.051	
Income LT 80% of AMI	-0.660	***	-0.657	***	-0.009	***	0.139	***	0.225	***	0.239	***	0.144	***	0.141	***
Income 80-99% of AMI	-0.145	***	-0.143	***	0.237	***	0.237	***	0.379	***	0.382	***	0.198	***	0.193	***
Income 100-119% of AMI	0.054		0.057		0.169	***	0.177	***	0.275	***	0.273	***	0.308	***	0.298	***
Income Missing	0.016		0.022		-0.628	***	-0.634	***	-0.823	***	-0.808	***	-0.852	***	-0.844	***
Census Tract Characteristics																
African-American Pop. Share	-0.004		0.001		0.005		0.012	***	0.000		0.008	***	-0.003		0.001	
Hispanic Pop. Share	0.000		0.005	*	0.000		0.003		-0.004	*	0.002		-0.015	**	-0.012	*
Other Race Pop. Share	0.006		0.012	***	0.001		0.005		-0.025		0.011		0.006		0.011	*
Tract Median Household Income (Ln)	-0.214		-0.454	***	-0.172		-0.396	**	0.380		0.100		-1.005	***	-1.129	***
Share Pop. With College	-0.037	***	-0.048	***	-0.023	***	-0.034	***	-3.508	***	-4.088	***	-0.015	***	-0.017	***
Rental Rate	-0.006	***	-0.001		-0.005	**	-0.004	**	0.005	*	0.007	**	-0.008	***	-0.006	***
Capitalization Rate	1.584	***	1.672	***	-0.118		-0.083		0.062		0.506	**	-0.144		-0.007	
Homeowner Turnover Rate	-0.007		-0.001		-0.007	***	-0.006	***	-3.819	***	-4.159	***	-0.006	*	-0.009	**
Share Households Credit Score <620	0.032	***			0.018	***			0.020	***			0.011	***		
Share Households Unscorable	-0.001				0.004				0.010				0.006	*		
Intercept	-0.560		2.619	**	-0.216		2.800		-6.712	*	-3.245		9.220	***	10.861	***
Psuedo R-squared	0.117		0.115		0.039		0.0372		0.081		0.079		0.072		0.071	
No. Observations	29,810				22,073				11,524				16,070			

Notes: *** denotes coefficients statistically significant at 0.01 level; ** denotes coefficients statistically significant at 0.05 level; and * denotes coefficients statistically significant at 0.10 level.

Appendix B: Tests for Spatial Autocorrelation

In this study, a two-stage approach was used to analyze the subprime mortgage lending at the census tract level. The first stage is the spatial exploratory data analysis (ESDA) (Anselin, 1995) and the second stage is the formal modeling stage using spatial regression models. The ESDA stage was used to develop an understanding of the spatial characteristics of the data, especially the magnitudes of spatial autocorrelation (SA) in the variables and whether the presence of SA—no matter how weak it may be—to assess if they may affect the efficiency of the parameter estimates in the models. Specifically, Moran's I and Geary's Ratio were used to test if the variables, both dependent and independent, have significant SA or not (Griffith, 1987). The two indices are widely used for SA diagnosis. Ordinary least squares (OLS) regressions were also estimated, treating the data as if no significant SA exists. The primary purpose of this exercise is not to develop parameter estimates, but to examine the residuals. With no significant SA, the residuals should have homoskedastic properties. But, with SA, a spatial pattern (correlation) will emerge. Therefore, examining the residuals spatially is one of the standard diagnostics for SA. The initial parameter estimates will be used for benchmarks to compare with results from spatial regressions performed in the next stage.

A common approach to handle the SA in the regression framework is to adopt spatial regression models. In these models, additional variables capturing the spatial relationships among observations are included in the form of a spatial weights matrix. Though many forms of spatial weights matrix exist, the simplest form of the spatial weights matrix is a binary matrix C, in which the element c_{ij} will be one if areal units i and j are neighbors, and zero otherwise. This matrix can be incorporated into the regression in various ways, depending on the structure of the model. Regardless of the way the matrix is incorporated into the model, the general purpose is to exploit the spatial relationship information captured by the matrix and the variables in the model to estimate the SA coefficient(s). When the SA is captured by the parameter(s), it is hoped that the residuals of the spatial regression will be independent.

Results and knowledge about the data acquired from the ESDA stage guide the modeling approach in the second stage, specifically in identifying the forms of spatial autocorrelation in regressions (4 forms). If SA is limited to certain variables, simpler spatial regression models may be used (Bailey and Gatrell, 1995). Otherwise, the more complicated models may be needed. The general structure of a spatial regression model is specified as:

$$y = \rho W_1 y + X\beta + \varepsilon \text{, where } \varepsilon = \lambda W_2 \varepsilon + \mu$$

The error term, μ, is assumed to be independent and identically distributed (iid). From this general model, taxonomy of spatial regression models can be formulated. If no SA exists in

all variables ($\rho = \lambda = 0$, the SA coefficients), the model is reduced to a simple linear regression model. If significant SA is only in y (i.e., $\lambda = 0$), then an autoregressive model (AR) can be used. However, it is more likely that significant SA is in both the dependent and independent variables and will be manifested in the residuals. Therefore, the simultaneous autoregressive model (SAR) (i.e., $\rho = 0$) is the most appropriate. Though several methods to estimate the spatial regression models exist, the maximum likelihood method has been the most popular.

The GeoDa tools for Spatial Autocorrelation diagnostics and spatial regression were used to estimate the models. SA requires spatial weights, which are weights based on the spatial location of tracts. The SA tests were performed for tract-level data using both purchase and refinance data for all tracts as well as tracts where the number of loans was greater than 10. The "Queen" model was used to create the spatial weights matrix. The OLS regression was then estimated within the spatial autocorrelation diagnostics for the residuals.

In SA diagnostics, only spatial weights could be accommodated and not the number of loans as used in the ordinary least squares analysis of tract-level loan shares. Instead, the SA analysis was conducted on both all the tracts and the tracts where the number of loans was greater than 10 to remove the influence of tracts with few loans originated. Exhibit B-1 summarizes the diagnostics and critical values used to determine SA.

Exhibit B-1

Spatial Autocorrelation Diagnostics

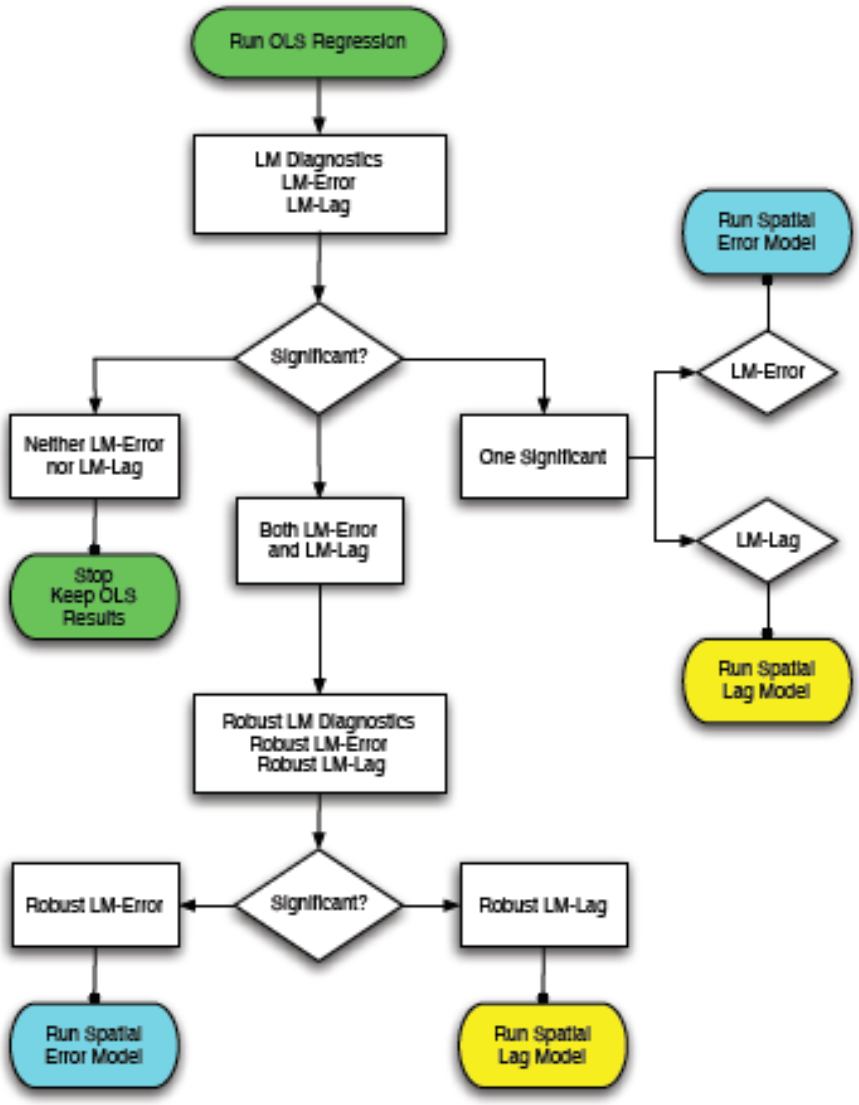
Test	MI/DF	Value(z)	Prob(p)
Moran's I (error)	0.028107	1.1821641	0.2371407
Lagrange Multiplier (lag)	1	1.6137908	0.2039597
Robust LM (lag)	1	1.9306226	0.1646904
Lagrange Multiplier (error)	1	0.3816595	0.5367164
Robust LM (error)	1	0.6984913	0.4032911
Lagrange Multiplier (SARMA)	2	2.3122821	0.3146982

The results were examined in a sequential order. If Moran's I was not significant, it was not necessary to run spatial regression models. If Moran's I was significant, the Lagrange Multiplier tests were examined to determine if the Spatial Lag or Error Model was better. If both Lagrange Multiplier tests were significant, the Robust LM test was examined for both lag and error and the more significant one of these was then chosen.

The decision process in diagnosing spatial dependence is displayed in Exhibit B-2. Exhibit B-3 indicates which model specifications showed evidence of significant spatial autocorrelation. These results are summarized in Exhibit B-4, where it is shown that there is greater evidence of spatial autocorrelation in the models of HUD subprime lender shares than there was for HMDA-defined higher-priced loan shares and there is slightly more spatial autocorrelation in the refinance data than in the purchase data.

Exhibit B-5 shows the change in the racial and ethnic coefficients resulting from the use of modeling approaches to account for spatial autocorrelation in the metropolitan areas where spatial autocorrelation was found in models of the higher-priced share of loans. As noted in Section 3, the presence of spatial autocorrelation would be expected to bias the estimates of the standard errors of the coefficients but not the coefficients themselves. However, the different modeling approach did change the estimated coefficients somewhat. As shown in Exhibit B-5, in general the models correcting for spatial autocorrelation have smaller coefficients on the share African-American and the share Hispanic, although, on average, the differences are small.

Exhibit B-2
Process for Testing and Correcting for Spatial Autocorrelation



Source: GeoDa Manual http://geodacenter.asu.edu/system/files/geodaworkbook.pdf, p.199.

Exhibit B-3

Summary of Results of Tests for Spatial Autocorrelation

		Autocorrelation in Model with Dependent Variable			
Area	**Dataset**	**P HMDA (#Obs)**	**P HUD (#Obs)**	**R HMDA (#Obs)**	**R HUD (#Obs)**
Albuquerque	All Tracts	No (183)	***Yes*** (183)	No (182)	*Marginal* (182)
	Tracts with 10+ Loans	No (173)	*Yes* (173)	No (173)	No (173)
Baltimore	All Tracts	***Yes*** (604)	***Yes*** (604)	***Yes*** (602)	***Yes*** (602)
	Tracts with 10+ Loans	***Yes*** (579)	***Yes*** (579)	***Yes*** (580)	***Yes*** (580)
Harrisburg	All Tracts	No (109)	No (109)	No (109)	No (109)
	Tracts with 10+ Loans	No (108)	No (108)	No (108)	No (108)
Memphis	All Tracts	***Yes*** (273)	***Yes*** (273)	***Yes*** (270)	***Yes*** (270)
	Tracts with 10+ Loans	***Yes*** (250)	***Yes*** (250)	***Yes*** (254)	***Yes*** (254)
Omaha	All Tracts	No (234)	No (234)	No (234)	***Yes*** (234)
	Tracts with 10+ Loans	***Yes*** (229)	No (229)	***Yes*** (230)	***Yes*** (230)
Providence	All Tracts	***Yes*** (345)	***Yes*** (345)	***Yes*** (345)	***Yes*** (345)
	Tracts with 10+ Loans	***Yes*** (341)	***Yes*** (341)	***Yes*** (343)	***Yes*** (343)
Tacoma	All Tracts	No (148)	***Yes*** (148)	No (148)	***Yes*** (148)
	Tracts with 10+ Loans	No (147)	*Yes* (147)	No (143)	*Yes* (143)
Toledo	All Tracts	No (170)	***Yes*** (170)	***Yes*** (170)	***Yes*** (170)
	Tracts with 10+ Loans	No (156)	***Yes*** (156)	***Yes*** (162)	***Yes*** (162)
Tulsa	All Tracts	***Yes*** (263)	***Yes*** (263)	***Yes*** (263)	***Yes*** (263)
	Tracts with 10+ Loans	***Yes*** (255)	***Yes*** (255)	***Yes*** (258)	***Yes*** (258)

Exhibit B-4

Summary of OLS Models with Evidence of Spatial Autocorrelation

	Purchase HMDA Higher-Priced	Purchase HUD Subprime Lenders	Refinance HMDA Higher-Priced	Refinance HUD Subprime Lender
Areas with Significant Spatial Autocorrelation	Baltimore, Memphis, Omaha,[a] Providence, Tulsa	Albuquerque, Baltimore, Memphis, Providence, Tacoma, Toledo, Tulsa	Baltimore, Memphis, Omaha,[a] Providence, Toledo, Tulsa	Albuquerque,[b] Baltimore, Memphis, Omaha, Providence, Tacoma, Toledo, Tulsa
Count	4.5	7	5.5	7.5

[a] These areas showed significant spatial autocorrelation for the truncated models (counts $\geq$ 10) but there was no evidence for the entire dataset.

[b] This area had marginal evidence of spatial autocorrelation in the entire dataset and showed no evidence in the truncated data (counts $\geq$ 10).

Exhibit B-5

Summary of Change in Racial/Ethnic Coefficients in OLS and Models Correcting for Spatial Autocorrelation

Metropolitan Area	African-American			Hispanic		
	OLS	Spatial Correction Model	Difference	OLS	Spatial Correction Model	Difference
Refinance Models						
Baltimore	0.17	0.17	0.00	0.28	0.27	0.00
Memphis	0.30	0.28	-0.01	0.12	0.07	-0.05
Omaha	0.10	0.12	0.03	0.18	0.17	-0.02
Providence	0.09	0.05	-0.04	0.06	0.04	-0.02
Toledo	0.09	0.07	-0.02	0.16	0.16	-0.01
Tulsa	0.03	-0.02	-0.05	-0.25	-0.07	0.18
Average	0.13	0.11	-0.02	0.09	0.10	0.01
Purchase Models						
Baltimore	0.14	0.12	-0.02	0.37	0.40	0.02
Memphis	0.19	0.13	-0.06	-0.32	-0.37	-0.06
Omaha	0.09	0.05	-0.04	0.03	-0.09	-0.11
Providence	-0.01	-0.04	-0.03	0.20	0.11	-0.09
Toledo	NA	NA	NA	NA	NA	NA
Tulsa	0.12	0.08	-0.04	-0.25	-0.17	0.07
Average	0.11	0.07	-0.04	0.01	-0.03	-0.03

U.S. Department of Housing and Urban Development
Office of Policy Development and Research
Washington, DC 20410-6000

July 2011