1-1-2003

Spatial Analysis of Historical Redlining: A Methodological Explanation

Amy E. Hillier
University of Pennsylvania, ahillier@design.upenn.edu

Postprint version. Copyright Fannie Mae Foundation, 2003. Published in *Journal of Housing Research*, Volume 14, Issue 1, pages 137-168. Publisher URL: http://www.fanniemaefoundation.org/programs/jhr

Spatial Analysis of Historical Redlining: A Methodological Explanation

Abstract
Despite widespread belief that redlining contributed to disinvestment in cities, there has been little empirical analysis of historical lending patterns. The lack of appropriate data and clear definitions of redlining has contributed to this void. This article reviews definitions and methods that have emerged from research on lending in recent years and considers how they can be applied to research on historical redlining. Address-level mortgage data from Philadelphia from the 1940s are analyzed using spatial regression, "hot spot" analysis, and surface interpolation.

Employing multiple definitions of redlining that focus on process and outcome, as well as spatial and statistical relationships in lending, the analyses result in a series of map layers that indicate where redlining may have occurred. In addition to providing some evidence of lending discrimination, this article promotes an explicitly spatial view of redlining that has conceptual and methodological implications for research on contemporary and historical redlining.

Keywords
mortgage discrimination; philadelphia; redlining; spatial analysis

Disciplines
Urban, Community and Regional Planning

Comments
Postprint version. Copyright Fannie Mae Foundation, 2003. Published in *Journal of Housing Research*, Volume 14, Issue 1, pages 137-168. Publisher URL: http://www.fanniemaefoundation.org/programs/jhr

Spatial Analysis of Historical Redlining: A Methodological Exploration

Amy E. Hillier*

Abstract

Despite widespread belief that redlining contributed to disinvestment in cities, there has been little empirical analysis of historical lending patterns. The lack of appropriate data and clear definitions of redlining has contributed to this void. This article reviews definitions and methods that have emerged from research on lending in recent years and considers how they can be applied to research on historical redlining. Address-level mortgage data from Philadelphia from the 1940s are analyzed using spatial regression, "hot spot" analysis, and surface interpolation.

Employing multiple definitions of redlining that focus on process and outcome, as well as spatial and statistical relationships in lending, the analyses result in a series of map layers that indicate where redlining may have occurred. In addition to providing some evidence of lending discrimination, this article promotes an explicitly spatial view of redlining that has conceptual and methodological implications for research on contemporary and historical redlining.

Keywords: Mortgage discrimination; Philadelphia; Redlining; Spatial analysis

Introduction

Redlining is one of the leading explanations for the disinvestment that took place in central cities during the middle decades of the 20th century, along with—and related to—deindustrialization (Bluestone and Harrison 1982), suburbanization (Jackson 1985), and racial segregation (Massey and Denton 1993). Although contemporary redlining research has used various data and methods to determine whether lenders have avoided lending to certain areas or types of areas, the historical redlining charged with contributing to this urban disinvestment has been subjected to very little empirical analysis. Explanations of how and why redlining took place between 1930 and 1970 rely on narratives rather than systematic analysis. How can such empirical tests of historical redlining be conducted, particularly in the absence of the kind of lending data used in contemporary redlining studies?

This article considers how redlining has been defined in research and legislation since the late 1960s, when fair housing legislation made redlining illegal. This article also reviews the types of data and methods that have been used to test for redlining. This framework is used to review existing research on historical redlining. Samples of address-level mortgage data from

* Amy E. Hillier is a Research Associate at the Cartographic Modeling Laboratory and a Lecturer in the Urban Studies Program at the University of Pennsylvania.

The work that provided the basis for this article was supported by funding under dissertation and postdoctoral grants from the U.S. Department of Housing and Urban Development. The author is solely responsible for the accuracy of the statements and interpretations contained in this article. Such interpretations do not necessarily reflect the views of the U.S. government. The author thanks Tony E. Smith, David Eldridge, and two anonymous reviewers for their assistance.

Philadelphia between 1940 and 1950 are described and tested using several exploratory quantitative methods. These methods include a spatial autoregressive (SAR) model that incorporates a spatial weight matrix into traditional regression; Ripley's local K function, which looks for statistically significant "hot spots"; and Kriging, which interpolates a continuous surface from a series of mapped points. Results from the different methods are compared, and implications for research on historical and contemporary redlining are discussed.

Definitions of Redlining in Legislation and Contemporary Research

Scholars, journalists, and fair housing activists generally agree that redlining involves ideas about creditworthiness that have little or nothing to do with the mortgage applicant and everything to do with the location of the property (Marcuse 1979). Discrimination against mortgage applicants on the basis of their race may have a similar negative effect on neighborhoods with high percentages of racial minorities, but this type of racial discrimination should be distinguished from spatial, or place-based, discrimination (Holloway and Wyly 2001; Zenou and Boccard 2000), which involves discrimination on the basis of the racial composition of the neighborhood in which a property is located (Black 1999). Most research on mortgage discrimination, including the Boston Fed study (Munnell et al. 1996), is designed to identify discrimination against individual mortgage applicants on the basis of their race, not the neighborhood (Holloway and Wyly 2001; Reibel 2000). A limited amount of research has considered the relationship between discrimination against individuals and neighborhoods (Holloway 1998; Holloway and Wyly 2001; Reibel 2000; Schill and Wachter 1993), but the distinction remains important for conceptual and methodological reasons, as well as for public policy (Tootell 1996; Turner and Skidmore 1999).

Beyond this consensus that redlining involves discrimination on the basis of neighborhood rather than individual characteristics, several different ideas about redlining have emerged in recent years. The dominant conceptualization holds that redlining can be process based or outcome based (Yinger 1995). Another less explicit distinction within the redlining literature is whether the focus is on spatial relationships or statistical associations between racial composition and mortgage outcomes. These different conceptualizations do not necessarily conflict, but they emphasize different aspects of redlining and have important implications for data analysis.

The distinction between process-based and outcome-based redlining corresponds loosely to the distinction in the mortgage discrimination literature between disparate treatment and disparate impact. Process-based redlining occurs when a specific discriminatory act by the lender can be identified during the mortgage-seeking process (Yinger 1995). Although loan decisions generally have received the greatest attention, process-based redlining theoretically can be identified at any point in the application process, including during the initial inquiry (Bogdon and Bell 1999). The discrimination does not necessarily involve poor service or overt discrimination; it might involve giving different information to different applicants or requiring different types of information from applications (Squires and O'Connor 2001; Yinger 1995). Squires explained that even when institutions have objective, race-neutral underwriting standards, those standards are not necessarily applied consistently, and the final decision often involves “subjective dimensions” (1994, 73). These types of practices are the target of the 1968 Fair Housing Act and the 1974 Equal Credit Opportunity Act (Ross and Yinger 2002; Yinger 1995).

Outcome-based redlining, on the other hand, occurs when neighborhoods with high percentages of racial minorities have less access to mortgages. This definition does not depend on identifying the particular stage in the mortgage decision-making process in which discrimination occurs, and it leaves open the possibility that an actor other than the lender may have introduced discrimination into the lending process (Ross and Yinger 2002; Yinger 1995). The Community Reinvestment Act (CRA) of 1977, which mandates that depository financial institutions make credit available to all areas from which they accept deposits, makes this form of discrimination illegal (Ross and Yinger 2002).

The unit of analysis is also central to distinguishing process-based from outcome-based redlining. Studies of process-based redlining (Harrison 2001; Holloway 1998; Holloway and Wyly 2001; Ling and Wachter 1998; Reibel 2000; Schill and Wachter 1993; Tootell 1996; Wyly 2002) employ individual-level data, whereas studies of outcome-based redlining (Avery, Beeson, and Sniderman 1999; see Schill and Wachter 1993 for a review) use aggregate data (Turner and Skidmore 1999).

Efforts to distinguish redlining research as either process or outcome based takes attention away from an equally important distinction within the research—focusing on spatial relationships or statistical associations. Redlining was originally a spatial concept, referring to specific areas that were not receiving appropriate amounts of mortgage credit. More recent redlining research refers to discrimination against certain types of areas, without the same attention to identifying specific redlined areas on a map.

Originally, redlining was considered a literal and geographic process. Community activists in Chicago’s Austin neighborhood coined the word “redlining” in the late 1960s while organizing residents around what they perceived as unfair lending practices. They used the word “redlining” to refer to the red lines that savings and loan associations had drawn around areas they refused to service (Pogge 1992). In 1968, the President’s National Advisory Panel on Insurance in Riot-Affected Areas also found evidence that lenders were drawing red lines on maps. The panel quoted from an underwriting guide that warned against providing insurance to areas with certain high-risk characteristics: “A good way to keep this information available and up to date is by the use of a red line around the questionable areas on territorial maps centrally located in the Underwriting Division for ease of reference by all Underwriting personnel” (President’s National Advisory Panel on Insurance in Riot-Affected Areas 1968, 6). The Douglas Commission found the same thing in 1969: “There was evidence of a tacit agreement among all groups—lending institutions, fire insurance companies, and FHA [Federal Housing Administration]—to block off certain areas of cities within ‘red lines,’ and not to loan or insure within them” (National Commission on Urban Problems 1969, 101).

Concern about geographic disparities in lending fueled efforts to supplement the Fair Housing Act with federal legislation aimed specifically at redlining. The 1975 Home Mortgage Disclosure Act (HMDA), which requires that lending institutions file information about loan amounts, disposition, and location, includes the following explanation: "The Congress finds that some depository institutions have sometimes contributed to the decline of certain geographic areas by their failure pursuant to their chartering responsibilities to provide adequate home financing to qualified applicants on reasonable terms and condition" (Canner 1982, 11). The 1977 CRA, which states that mortgage lenders have a legal obligation to provide credit to low-income neighborhoods in areas where they are chartered, is even more specifically targeted at this kind of redlining—just as it is targeted at outcome-based, rather than process-based, redlining. Under CRA, lenders are required to prepare statements that include maps of their service areas with explanations of the types of products they offer (Squires 1994). This spatial conceptualization of redlining is also consistent with the geography (Harvey 1985; Smith, Caris, and Wyly 2001) and historical sociology literature (Bartelt et al. 1987) that points to capital flows and the "spatiality of credit" (Wyly 2002, 5) as an explanation for urban decline.

Investigations of redlining that uncover maps with red lines are the exception, but at the core of this original conceptualization is the idea that redlining is something spatial. When lending institutions redline an area, they avoid lending to whole spatially contiguous areas, not just small blocks with similar characteristics scattered throughout a city. In their search for more subtle forms of discrimination against certain areas, researchers more recently have looked for statistical associations between racial composition and mortgage outcomes, without the same concern for spatial relationships (Avery, Beeson, and Sniderman 1999; Harrison 2001; Holloway 1998; Holloway and Wyly 2001; Ling and Wachter 1998; Phillips-Patrick and Rossi 1996; Reibel 2000; Schill and Wachter 1993; Shlay 1989; Tootell 1996; Wyly 2002). If, after controlling for a host of individual and neighborhood-level factors, areas with higher percentages of racial minorities have higher mortgage denial rates, there is evidence of redlining. This approach is consistent with "neighborhood effects" research (Ellen and Turner 1997) in that it incorporates contextual information about a neighborhood without considering the spatial proximity of neighborhoods with similar outcomes. Rather than identifying a spatially contiguous area that has been underserved, this view of redlining considers whether neighborhoods with racial minorities are underserved overall. Although these spatial and contextual views of redlining are conceptually quite different, they may lead to the same conclusions. Because of the extreme levels of segregation in American cities during the second half of the 20th century (Massey and Denton 1993), areas with high levels of racial minorities often are clustered. Conceptually, however, these views of redlining are quite different and require different types of analyses to uncover them.

These different approaches to understanding redlining—as a process or an outcome and as a spatial relationship or a statistical association—overlap considerably. They do not conflict; they complement one another. Redlining studies can be distinguished in both ways simultaneously—as either process based or outcome based and as either spatial or contextual. Because most studies examine lender discrimination using applicant-level data and consider neighborhood characteristics rather than location, they would be defined as process based and contextual. Studies that examine lending outcomes using aggregate data and that identify the actual location of underserved areas would be considered outcome based and spatial.

Data and Methods Used in Contemporary Redlining Research

The growing number of studies on redlining—which is still dwarfed by the number of studies on lending discrimination—is directly attributable to the availability of lending data from the past 20 years. To a large extent, the available data influence the conceptualization of redlining adopted. To test for process-based redlining, researchers need access to application-level data that include information about the characteristics of the applicant, property, and neighborhood in which the property is located. Outcome-based studies, on the other hand, use aggregate data. Spatial investigations can use individual-level or aggregate data, but they require paper maps or specific information about the location of properties. Tests for statistical association also can use either individual-level or aggregate data, but they only require information about the characteristics of neighborhoods, not their location.

Almost all studies of contemporary redlining use HMDA data. Amendments in 1980 to HMDA mandated the collection and distribution of census tract–level data on lending by certain types of lenders. Subsequent amendments in 1989 required that lenders report whether they accept or reject a mortgage application. Lenders also must report the income, sex, and race of applicants. The result is a huge data set, appropriate for process-based (individual) and outcome-based (aggregate) studies. Researchers have supplemented these data with data about individual applicants from lender records (Tootell 1996) and with data about neighborhoods (tracts) from the census. In the absence of the enriched HMDA data set, redlining studies from before 1989 used some additional sources of information on lending. For example, Bradbury, Case, and Dunham (1989) used deed records for their study of lending patterns in Boston. However, studies since 1989 have relied primarily on HMDA data.

HMDA data make loan disposition (process-based studies) or census tracts (outcome-based studies) the most obvious choices as units of analysis. Bradbury, Case, and Dunham (1989) used neighborhood statistical areas created by the census in 1980 instead of census tracts, but that is the exception. HMDA data also make the decision to accept or reject a mortgage application (process-based studies) and the mortgage acceptance or rejection rate (outcome-based studies) the most obvious choices for dependent variables. However, prior to the 1989 HMDA amendments, researchers used as dependent variables the aggregate number of loans, the aggregate value of mortgages, the ratio of the aggregate number of loans to deed transfers or separately owned structures, and the type of loan made (see the review of literature in Schill and Wachter [1993]). The choice of independent variables varies little across redlining studies, and it generally involves some version of census tract–level racial composition data.

The control variables distinguish redlining studies the most, and omitted variable bias is the most common critique of redlining studies (Ross and Yinger 2002). Researchers have used the data set created for the Boston Fed study (Tootell 1996), which includes information about credit history, probability of default, cost of default, and other characteristics of applicants, such as age, marital status, and number of dependents. When aggregated, these rich applicant and loan data are used to supplement census data. A variety of information about neighborhoods also has been used, including information on income, housing value, age and condition of housing, unemployment and education levels, age of population, duration of residency, and tax arrearage (see Schill and Wachter 1993). Other studies (Calem 1996; Lang and

Nakamura 1993) included information about the number of housing sales in an effort to test the effect of information about particular areas on lending decisions. Attention to interaction variables also distinguishes some redlining research, with several studies focusing on the interaction between race and racial composition in loan dispositions (Holloway 1998; Holloway and Wyly 2001; Reibel 2000; Schill and Wachter 1993).

Analyses of data in these studies nearly always involve some form of multiple regression. Studies trying to explain the lender's decision to accept or reject frequently used logistic regression (Holloway and Wyly 2001; Reibel 2000; Wyly 2002). Often, studies used multiple levels, introducing neighborhood characteristics into models after individual characteristics were analyzed separately (Avery, Beeson, and Sniderman 1999; Holloway 1998). Phillips-Patrick and Rossi (1996) critiqued single-equation models, arguing that models with simultaneous equations are needed. Geographic Information Systems (GISs) sometimes were used as tools for displaying the spatial distribution of independent, dependent, and control variables (Holloway 1998; Holloway and Wyly 2001; Wyly 2002), but GISs rarely were used to conduct spatial analyses, and statistical models rarely included spatial data other than tract-level variables.

Research on Historical Redlining

Technically, the Civil Rights Act of 1866 made racial discrimination in all contracts illegal, but it was not until 1968 that the U.S. Supreme Court ruled that this protection extended to real estate transactions (Yinger 1995). Passage of the 1968 Fair Housing Act certainly did not end lending discrimination, but along with HMDA (1975) and CRA (1977), this period marked a new era in redlining research. Housing discrimination no longer was considered just unfair; it was illegal. By collecting and distributing data about lending practices, the federal government encouraged community groups, fair housing activists, journalists, and researchers to look for signs of discrimination. Investigations of redlining before 1968, on the other hand, have no meaningful legal definitions of redlining on which to focus. Instead of seeking to document and correct illegal practices, researchers have examined redlining to help explain the disinvestment in central city neighborhoods between, roughly, 1930 and 1970. During this earlier era, the federal government was considered the foe—not the friend—of fair housing, and it was the target of many investigations of redlining. For the purposes of this article, historical redlining refers to discriminatory mortgage practices prior to the passage of the Fair Housing Act.

Unlike researchers of contemporary redlining, researchers have not been explicit about their conceptualizations of historical redlining. Implicit in the existing research on historical redlining are ideas about what redlining is. Research that focuses on maps and underwriting criteria as evidence of discrimination is concerned with process-based redlining, whereas research that focuses on disinvestment in particular areas is concerned with outcome-based redlining. Similarly, research that examines maps can be thought of as focusing on the spatial aspects of redlining, whereas a focus on underwriting criteria that takes neighborhood characteristics—rather than location—into consideration can be thought of as concentrating on association or contextual redlining.

Much of the research on historical lending practices focuses on fairly literal, and certainly spatial, conceptualizations of redlining, and the research focuses nearly exclusively on the lending process rather than on lending outcomes. Researchers have given the greatest attention to the Home Owners' Loan Corporation's (HOLC's) color-coded maps that used red to indicate areas considered hazardous for real estate investments. HOLC created the maps in the late 1930s after making a million loans to homeowners at risk of losing their homes to foreclosure during the Depression. HOLC maps categorized neighborhoods in more than 200 cities across the country according to their stage in decline. Kenneth Jackson (1985) discovered the maps in the late 1970s while researching *Crabgrass Frontier*. He argued that FHA and private lenders used the maps when considering where to make loans, and he connected the maps to the practice of redlining (Jackson 1980; 1985). In recent research, Hillier (2003) suggested that HOLC's role in influencing lender decisions and institutionalizing redlining is overstated, and she pointed to the role other maps and information about neighborhoods played in the decision-making process.

Researchers have uncovered a number of other examples of maps with red lines. Metzger (1999) described an FHA mortgage risk map for Chicago from 1938 that he found in the papers of Ernest Burgess. Mohl (1999) reported finding dozens of "redlining maps" within municipal agency records in Miami, particularly within the Dade county building and zoning department records. Other maps that have been found clearly were created for the same purpose—to indicate neighborhood risk levels—but they did not have red lines. A 1934 map created by the former chief appraiser in Philadelphia for Metropolitan Life Insurance indicated concentrations of African Americans, Jews, and Italians; the age and value of housing; and the economic class of residents (Brewer 1934). Security-First National Bank of Los Angeles created a map applying the neighborhood life cycle theory, categorizing each neighborhood as either subdivision, growth, maturity, decline, or decadence (Smith 1938).

During the 1930s and 1940s, FHA had a large collection of maps because it was directly involved in the Works Progress Administration (WPA) real property surveys, which generated census tract–level and block-level data for cities across the country. One particularly sophisticated FHA map, published in *The Structure and Growth of Residential Neighborhoods in American Cities*, had a series of transparent overlays that showed the relationships between housing conditions, age of housing, and racial composition in Richmond, Virginia (Hoyt 1939). The Mortgage Conference of New York, an association of 37 New York banks created in 1932, also was involved in making and sharing maps. In a criminal and civil antitrust suit filed against the Conference in 1946, the U.S. Justice Department charged members with making block-level maps of racial change available to members who then "refrained from making mortgage loans on properties in such blocks" (Metzger 1999, 198). The Justice Department succeeded in securing a civil injunction, and the Mortgage Conference of New York disbanded in 1948 (Abrams 1955; Metzger 1999; Schisgall 1975).

Evidence exists, therefore, that federal housing agencies, local government agencies, and the mortgage lending industry all devised methods for guiding lending decisions based on location. These primary sources demonstrate that, during the 1930s and 1940s, lending was understood in geographic terms. Maps provided a means of systematically recording and communicating information within and across organizations about the risks of mortgage lending.

Other research has pointed to explicitly racial underwriting standards as evidence of redlining. This type of research is more concerned with determining whether lenders discriminated against certain types of areas—those home to African Americans—than specific locations. The Chicago Commission on Race Relations, writing in the wake of Chicago's 1919 race riot, was among the first to document concerns about this type of redlining. In addition to race-based lending discrimination, the Commission noted lenders' "blanket provision" not to make loans on "property in changing or depreciated districts" (Chicago Commission on Race Relations 1922, 221).

FHA was a leader in promoting neighborhood risk ratings during the post-Depression era. Its *Underwriting Manual* outlined a detailed rating system that encouraged appraisers to consider the stability of an area and its protection from "adverse influences," generally referring to African Americans and other racial and ethnic minorities (FHA 1935). This preoccupation with the role of racial minorities in neighborhood decline also appeared in the National Association of Realtors' code of ethics, which warned members against "introducing into a neighborhood a character of property or occupancy, members of any race and nationality, or any individual whose presence will clearly be detrimental to property values in the neighborhood" (Squires and O'Connor 2001, 4). The dozens of articles about neighborhood risk-rating systems published in the real estate and appraisal industry journals in the 1930s and 1940s further promoted disparate treatment of mortgage applications based on the location of a property (Hillier 2003; Mohl 1997).

All this research focusing on maps and underwriting criteria has sought to identify discrimination within the lending process. Unlike studies of contemporary process-based redlining, though, these efforts have not tested the effect this clearly discriminatory process had on individual mortgage applications. Similarly, historians have made few efforts to analyze aggregate data on mortgage outcomes. A few exceptions have taken advantage of aggregate data and have described in general terms where HOLC and FHA mortgages were extended (Cohen 1990; Jackson 1985; Metzger 1999). More common are historical narratives that associate racial composition with access to mortgages and neighborhood decline (Bissinger 1997; Massey and Denton 1993; Mohl 1987). The subsequent sections of this article consider what data are available for testing more directly for the categories of redlining identified in studies of contemporary mortgage lending.

Data for Studying Historical Redlining

The lack of empirical research on historical redlining primarily results from the lack of access to quantitative data on lending. Information about where lenders provided mortgages, the acceptance rate of mortgage applications, and the terms of mortgages is difficult to locate. Prior to the passage of HMDA, the federal government made no effort to collect lending data from lenders. FHA has denied maintaining records for anything below the county level on properties it insured during the 1930s, 1940s, and 1950s (Jackson 1985). The decennial census does contain information about racial composition and housing since 1940, including information about the number of owner-occupied homes with and without mortgages, population change, length of residence (starting in 1950), and housing values. WPA's real property surveys, conducted in the 1930s, produced similar information about racial composition and housing characteristics. However, neither the decennial census nor the WPA real property survey includes specific information about mortgage activity.

In the absence of federal records on lending, the search turns to local sources: municipal records departments, title companies, real estate publications, and lenders' own records. Lenders potentially could be a source of detailed information about their own mortgage activity. However, the literature on historical redlining provides no examples of this. There is little reason to believe that lenders would have maintained detailed records for decades, or that they would be willing to share such records. Title companies, which have maintained extensive property-specific files, also might be a source of information about mortgages. The author had no success using title company records in Philadelphia, however, because mortgage instruments were removed from the property files once they were satisfied.

Municipal mortgage records are another possibility. The mortgage records for Philadelphia have limited usefulness because they are indexed by date and by the names of the lender and homeowner, not by the location. To find address information, it is necessary to use the index to locate the actual mortgage instrument, which is very time consuming. The most accessible source of mortgage data in Philadelphia is the *Philadelphia Realty Directory and Service*, published annually between 1926 and 1958. The volumes include a list of all real estate transactions by date and address, including those that did not involve a mortgage, as well as a list of all real estate in the city with the date of the most recent transaction for each property.

The *Philadelphia Realty Directory and Service* includes property-level rather than aggregate data, so one needs to use a large sample of the data to build a data set of aggregate mortgage data. A more considerable limitation of lending data involves the lack of ideal dependent or control variables. The directory contains information about successful real estate transactions, not mortgage applications, so there is no information about rejected applications. The directory does include information about the terms of the mortgage, including the interest rate, amount of mortgage, and sale price, along with the name of the lender (which generally indicates the type of lender). The directory also includes information about mortgage defaults, but because the directory is organized by date, this information is contained in a separate section from the transaction information. The data available to serve as control variables also are limited. The directory lists the size of the property, its assessed value, and the address of the owner (if off property), but it includes no information about the race, income, or credit history of the homeowner.

These limitations are considerable, but the kind of information contained in the directory makes it possible to test for redlining in a number of ways, including spatial analysis, tests for statistical associations, and tests for process-based redlining. The data used to test for redlining in Philadelphia are described below; then the specific methods of analysis are introduced.

Description of Philadelphia Data Samples

The samples of mortgage data used in this article are from Philadelphia between 1940 and 1950. Although Philadelphia's history in the 20th century is distinct in many ways, it resembles the history of other large northeastern and midwestern cities that experienced significant decline. Philadelphia's population peaked midcentury at 2 million, but its transition to postindustrialism started well before then (Adams et al. 1991). The city's share of the region's jobs dropped considerably as manufacturers left the city, and part-time and temporary service jobs replaced stable industry jobs that paid a living wage. In the meantime, the size of the African-American population increased in the Philadelphia region—tenfold between 1850 and 1950—with the overwhelming majority located in the city. In subsequent decades, tens of thousands of residents, mostly white, left the city. The tens of thousands of abandoned row homes and factories that remain scattered across the city are a testament to the massive disinvestment that Philadelphia experienced over several decades.

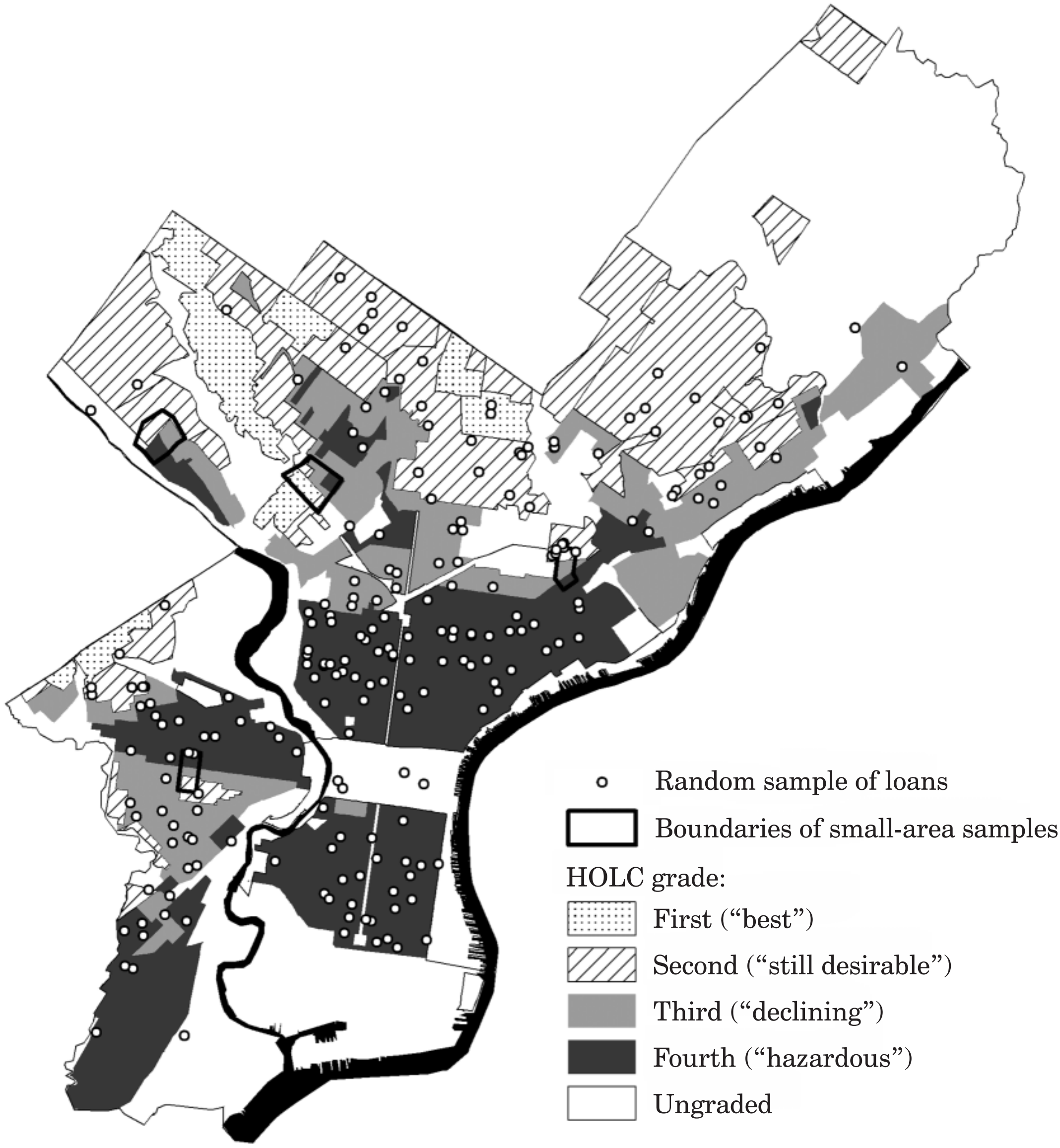
Data about Neighborhood Characteristics

Data about neighborhood conditions in Philadelphia were collected from maps and tables produced by HOLC, WPA, and the U.S. Bureau of the Census. The 1937 residential security map HOLC created for Philadelphia was digitized with GIS software. A photograph of the original HOLC map was used to identify the streets that served as boundaries between graded areas. Then, a digital 1990 census block map was used to create the digital HOLC map by dissolving the boundaries between the blocks that made up the graded areas on the HOLC maps. The same process was used to digitize the boundaries of 1930 and 1940 census tracts for Philadelphia. Attribute data from the 1934 WPA real property survey and the 1940 census were appended to the digitized census tract boundary files.

Random Sample of Loans

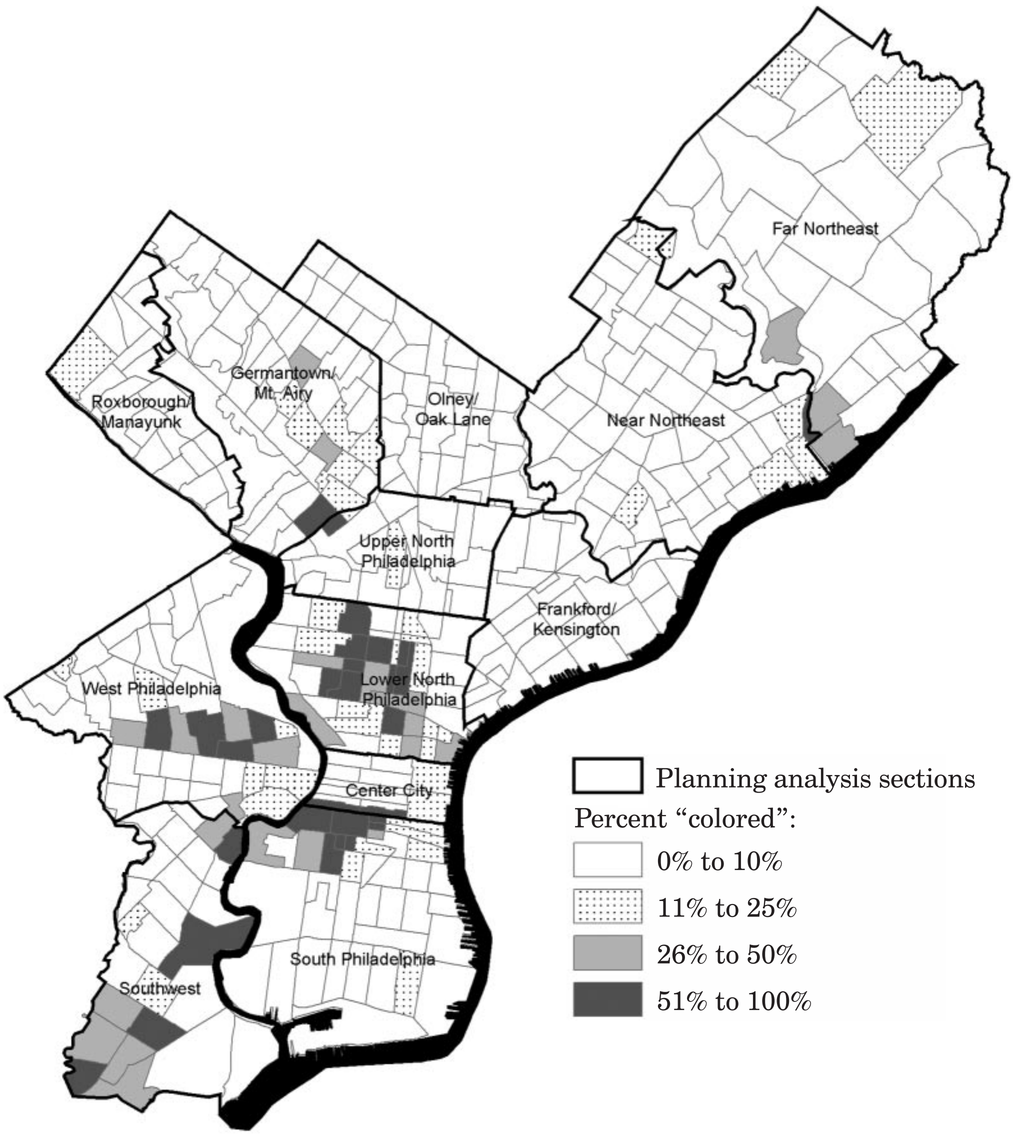
Mortgage data were collected from the *Philadelphia Realty Directory and Service* for a random sample of 500 property transactions in Philadelphia between 1938 and 1950. Random numbers were generated to select pages, columns, and item numbers in the directory. For each item, the name of the lender, amount of the mortgage, interest rate, assessed value of the property, sales price, size of the property, and address of the owner were recorded (see figure 1 for the location of properties included in the random sample). Census tract–level data from the 1934 WPA real property survey were joined to the specific properties; the data include the percent of "colored" families (figure 2), the percent of overcrowded housing, the percent of multifamily structures, the percent of owner-occupied housing, the percent of housing needing major repairs, and the median value of housing (tables 1 and 2). The HOLC grade for each property was determined using the digitized version of HOLC's 1937 residential security map of Philadelphia. Only transactions that involved a mortgage and have complete information for the interest rate, sales price, and amount of the mortgage are included, yielding a sample of 186 mortgages.

Figure 1. **Home Owners' Loan Corporation Residential Security Map for Philadelphia, 1937**



Sources: Data from Federal Home Loan Bank Board (1933–1940) and *Philadelphia Realty Directory and Service* (1938–1950).
Note: The Home Owners' Loan Corporation (HOLC) created this residential security map for Philadelphia in 1937, assessing the risk for real estate investment using four grades. Mortgage data were collected on the random citywide sample (points on the map) and for all properties in the small areas (areas bounded by thick black lines).

Figure 2. **Percent of "Colored" Households in Philadelphia by Census Tract, 1940**



Source: Data from Works Progress Administration (1934).
Note: Most "colored" (nonwhite) families lived near the central part of Philadelphia in 1934. The census tract–level data make it possible to analyze the effect of racial composition on mortgage outcomes.

Table 1. **Variables Included in Random Citywide Sample**

Variable	Description
Interest rate	Interest rate on the mortgage, as recorded in the *Philadelphia Realty Directory and Service*
Mortgage amount	Amount of the mortgage in thousands of dollars, as recorded in the *Philadelphia Realty Directory and Service*
Sales price	Amount paid for the house in thousands of dollars, as recorded in the *Philadelphia Realty Directory and Service*
Size of house	Square footage of the house in thousands, as recorded in the *Philadelphia Realty Directory and Service*
Owner occupied	Dummy variable, where 1 = owner occupied, 0 = other, as recorded in the *Philadelphia Realty Directory and Service*
Year of mortgage	Continuous variable, where 1 = 1938, 13 = 1950
Median value	Median value of housing (in thousands of dollars) in the census tract, according to the 1934 WPA real property survey
Percent owner occupied	Percent of housing units in the census tract that were owner occupied, according to the 1934 WPA real property survey
Median age	Median age of housing in the census tract, according to the 1940 census
Percent overcrowded	Percent of housing units in the census tract that were overcrowded, according to the 1934 WPA real property survey
Percent needing major repairs	Percent of housing units needing major repairs in the census tract, according to the 1934 WPA real property survey
Percent "colored" families	Percent of "colored" families in the census tract, according to the 1934 WPA real property survey
Percent multifamily structures	Percent of housing units in multifamily buildings in the census tract, according to the 1934 WPA real property survey

Note: WPA = Works Progress Administration.

Table 2. **Summary Statistics for Random Citywide Sample**

Variable	Minimum	Maximum	Mean	Standard Deviation
Interest rate (percent)	4.0	6.0	5.3	0.7
Mortgage amount ($, in thousands)	0.350	15.000	3.602	2.359
Sales price ($, in thousands)	0.500	22.500	5.220	3.359
Size of house (feet^2, in thousands)	0.421	12.800	1.457	1.225
Owner occupied	0	1	0.8	0.4
Year of mortgage	3	13	8.9	2.8
Median value ($, in thousands)	0	8.659	3.724	1.526
Percent owner occupied	0	80.5	43.1	15.1
Median age (years)	0	80	31.3	21.3
Percent overcrowded	0	99.0	12.7	9.8
Percent needing major repairs	0	50.4	5.2	7.4
Percent "colored" families	0	93.6	11.8	20.3
Percent multifamily structures	0	21.6	1.9	3.7

Note: N = 186.

Small-Area Loan Samples

Mortgage data also were collected from four small areas in Philadelphia where three or more HOLC grades come together (see figure 1 for the location of the small areas). The areas were chosen to represent parts of the city where there were different real estate patterns, racial and household income compositions, and levels of industrial activity. The areas include parts of West Philadelphia, Roxborough/Manayunk, East Falls/Germantown, and Kensington/Frankford. All of South Philadelphia and most of North Philadelphia were excluded from consideration because they were assigned the same grade (the fourth grade, "hazardous") by HOLC. Each of the four small areas selected, located in the west, northwest, and lower northeast sections of the city, covers roughly 20 to 25 blocks and includes between 300 and 900 properties. Details of all the transactions between 1925 and 1950 for all of the properties in these small areas were collected from the *Philadelphia Realty Directory and Service*. This includes the assessed value (in 1950), the size of the house, the amount of each mortgage, and the interest rate.

Only two of the small areas—East Falls/Germantown and Kensington/Frankford—have enough racial diversity to test for the effect of racial composition on mortgage outcomes. Census block–level data on housing and racial composition from the 1939 WPA real property survey were assigned to the properties in these two small areas. The variables include the percent of nonwhite residents, median rent, percent of substandard housing, percent of owner-occupied housing, and percent of overcrowded housing (tables 3 through 5). The only data available from the 1939 survey are maps with ordered categories corresponding to the actual ranges in the characteristics (Philadelphia Housing Authority and WPA of Pennsylvania 1939).

Table 3. **Variables Included in Analysis of Small-Area Samples**

Variable	Description
Interest rate	Interest rate on the mortgage, as recorded in the *Philadelphia Realty Directory and Service*
Mortgage amount	Amount of the mortgage in thousands of dollars, as recorded in the *Philadelphia Realty Directory and Service*
Sales price	Amount paid for the house in thousands of dollars, as recorded in the *Philadelphia Realty Directory and Service*
Size of house	Square footage of the house in thousands, as recorded in the *Philadelphia Realty Directory and Service*
Owner occupied	Dummy variable, where 1 = owner occupied, 0 = other, as recorded in the *Philadelphia Realty Directory and Service*
Year of mortgage	Continuous variable, where 1 = 1938, 13 = 1950
Percent nonwhite	Percent of nonwhite residents in the census block, according to the 1939 WPA real property survey, by ordered category (0 = 0 percent, 1 = 1 to 9.9 percent, 2 = 10 to 19.9 percent, 3 = 20 to 39.9 percent, 4 = 40 to 59.9 percent, 5 = 60 to 79.9 percent, 6 = 80 to 100 percent)
Median rent	Median rent in the census block, according to the 1939 WPA real property survey, by ordered category (0 = less than $10, 1 = $10 to $14.99, 2 = $15 to $19.99, 3 = $20 to $24.99, 4 = $25 to $29.99, 5 = $30 to $49.99, 6 = $50 or more)
Percent substandard	Percent of substandard housing units in the census block, according to the 1939 WPA real property survey, by ordered category (0 = 0 percent, 1 = 1 to 14.9 percent, 2 = 15 to 29.9 percent, 3 = 30 to 49.9 percent, 4 = 50 to 69.9 percent, 5 = 70 to 84.9 percent, 6 = 85 to 100 percent)
Percent owner occupied	Percent of housing units in the census block that were owner occupied, according to the 1939 WPA real property survey, by ordered category (0 = 0 percent, 1 = 1 to 9.9 percent, 2 = 10 to 19.9 percent, 3 = 20 to 29.9 percent, 4 = 30 to 49.9 percent, 5 = 50 to 74.9 percent, 6 = 75 to 100 percent)
Percent overcrowded	Percent of housing units in the census block with more than 1.5 persons per room, according to the 1939 WPA real property survey, by ordered category (0 = 0 percent, 1 = 1 to 9.9 percent, 2 = 10 to 14.9 percent, 3 = 15 to 19.9 percent, 4 = 20 to 29.9 percent, 5 = 30 to 49.9 percent, 6 = 50 percent or more)

Note: WPA = Works Progress Administration.

Table 4. **Summary Statistics for East Falls/Germantown Small-Area Sample**

Variable	Minimum	Maximum	Mean	Standard Deviation
Interest rate (percent)	4	6	5.4	0.7
Mortgage amount ($, in thousands)	0.300	28.000	4.348	3.427
Sales price ($, in thousands)	0.774	45.000	6.838	6.532
Size of house (feet^2, in thousands)	0.576	52.886	3.317	5.620
Owner occupied	0	1	0.7	0.4
Year of mortgage	3	13	8.0	2.9
Percent nonwhite	1	7	4.0	2.6
Average rent	2	7	5.6	1.2
Percent substandard	1	5	2.5	1.1
Percent owner occupied	1	7	4.4	1.4
Percent overcrowded	1	2	1.4	0.5

Note: N = 332.

Table 5. **Summary Statistics for Kensington/Frankford Small-Area Sample**

Variable	Minimum	Maximum	Mean	Standard Deviation
Interest rate (percent)	4.0	6.0	5.3	0.6
Mortgage amount ($, in thousands)	0.600	11.000	3.796	1.794
Sales price ($, in thousands)	1.050	9.500	4.886	1.952
Size of house (feet^2, in thousands)	0.161	33.440	1.207	1.559
Owner occupied	0	1	0.9	0.3
Year of mortgage	2	13	9.4	3.3
Percent nonwhite	0	2	0.8	0.5
Average rent	0	6	5.0	1.9
Percent substandard	0	5	1.3	1.0
Percent owner occupied	0	7	3.9	2.3
Percent overcrowded	0	3	0.9	0.6

Note: N = 456.

Statistical Tests for Historical Redlining

Using the data from the HOLC map, census, and WPA in conjunction with the random citywide and small-area samples, it is possible to conduct several exploratory analyses to look for evidence of redlining in Philadelphia. The first two tests use ordinary least squares (OLS) and SAR models to test for the effect of racial composition and HOLC grade on the amount of the mortgage and the interest rate for the random and small-area samples. These tests are designed to identify process-based, rather than outcome-based, redlining because they use transaction-level data. They also are designed to identify statistical associations, rather than spatial relationships, in lending patterns. The third test, using a local K function to conduct hot spot analysis on the random citywide sample, aims to identify areas that may have been significantly overserved or underserved by mortgage lenders. The final test uses Kriging to interpolate a surface of interest rates, based on the random citywide sample, to determine if there are contiguous areas that paid higher interest rates. These final two methods are designed to test for outcome-based redlining, because they essentially aggregate the transaction-level data. They also are designed to identify spatial, rather than just statistical, patterns within the data. Each method is described below, and the results from all four tests are compared and discussed in the following section.

Testing Effect of Racial Composition Using Spatial Regression

Because they incorporate information about the racial characteristics of the areas in which mortgages were made with information about the individual properties and mortgage transactions, the random citywide and small-area samples can be analyzed using regression models to determine if the racial composition of an area affected mortgage outcomes. In the absence of information about mortgage applications and loan rejections and approvals, the amount of the mortgage and the interest rate on the mortgage are the best available dependent variables. Efforts were made to calculate a loan-to-value ratio, but neither assessed value (as determined for property tax purposes and listed in the *Philadelphia Realty Directory and Service*) nor sales price serves as an appropriate proxy for appraised value. The percent of nonwhite population serves as an independent variable, with characteristics of the property, transaction, and area (census tract for the citywide sample and census block for the small area samples) as control variables.

Tests for spatial autocorrelation were performed because properties near each other are more likely to have similar outcomes, violating the assumption of independence assumed in OLS regression and potentially inflating the significance of variables (Bailey and Gatrell 1995). Moran's I tests show significant autocorrelation in the small-area samples but not in the random citywide sample. As a result, the random citywide sample is analyzed with OLS, and an SAR model is used in addition to OLS to analyze the small-area samples. The SAR model is defined as:

$$y = \mathbf{X}\beta + u, \tag{1}$$

where $u = \rho \mathbf{W} u + \varepsilon$,

where y is the dependent variable, $\mathbf{X}$ is a vector of independent and control variables, β is the coefficient for the independent and control variables, ρ is an autoregressive parameter, $\mathbf{W}$ is a weight matrix (incorporating values of the dependent variable for nearby observations), and ε represents a general error term. A weight matrix that considers properties on the same block face with mortgages within the same three-year period was constructed.[1]

The amount of the mortgage and the interest rate were regressed in separate equations against the neighborhood variables in addition to several property-level variables. Results show that the percent of nonwhite residents is not significantly related to the amount of the mortgage, but it is a significant factor for interest rates in all three samples (tables 6 through 11). For the random and East Falls/Germantown small samples, higher percentages of

[1] The SAR model was calculated using the program "sar.m" written by Tony E. Smith for Matlab.

nonwhites predict higher interest rates, whereas in Kensington/Frankford, the relationship is the opposite, with lower percentages of nonwhites predicting higher interest rates. Other neighborhood characteristics, including the condition of the housing and the percent that is owner occupied, have a more consistent significant effect. There are few differences between the OLS and SAR results.

Table 6. **Ordinary Least Squares Estimates Predicting Amount of Mortgage for Random Sample**

Variable	Beta	Standard Error
Constant	0.9838	0.5381
House-level variables		
Sales price	0.6192	0.0304***
Size of house	0.1421	0.0738*
Owner occupied	0.3707	0.2108*
Year of mortgage	0.0231	0.0322
Census tract–level variables		
Percent "colored" families	0.0007	0.0053
Median age	–0.0076	0.0051
Percent overcrowded	0.0005	0.0095
Percent multifamily structures	0.0276	0.0268
Percent owner occupied	–0.0150	0.0070**
Percent needing major repairs	0.0029	0.0133
Median value	–0.1403	0.0664**

Note: Adjusted R^2 = 0.7916.
*$p < 0.1$. **$p < 0.05$. ***$p < 0.01$.

Table 7. **Ordinary Least Squares Estimates Predicting Interest Rates for Random Sample**

Variable	Beta	Standard Error
Constant	5.3272	0.3071
House-level variables		
Mortgage amount	–0.1084	0.0234***
Size of house	0.0440	0.0423
Owner occupied	–0.0323	0.1184
Year of mortgage	–0.0236	0.0180
Census tract–level variables		
Percent "colored" families	0.0058	0.0030*
Median age	0.0088	0.0029***
Percent overcrowded	–0.0038	0.0054
Percent multifamily structures	0.0253	0.0152*
Percent owner occupied	0.0099	0.0040**
Percent needing major repairs	0.0144	0.0075*
Median value	–0.0721	0.0000**

Note: Adjusted R^2 = 0.3204.
*$p < 0.1$. **$p < 0.05$. ***$p < 0.01$.

Table 8. **Regression Estimates Predicting Amount of Mortgage for East Falls/Germantown**

	OLS Results		SAR Results	
Variable	Beta	Standard Error	Beta	Standard Error
Constant	−1.7794	1.2863	−2.3398	1.3570
House-level variables				
Sales price	0.4102	0.0191***	0.4064	0.0195***
Size of house	0.0585	0.0189***	0.0463	0.0187**
Owner occupied	0.1969	0.2527	0.1132	0.2488
Year of mortgage	0.1777	0.0368***	0.1741	0.0414***
Census block–level variables				
Percent nonwhite	0.0082	0.0799	0.0226	0.0872
Average rent	0.2674	0.1708	0.4215	0.1840**
Percent substandard	−0.0219	0.1253	0.0132	0.1336
Percent owner occupied	−0.0337	0.0897	−0.1112	0.1015
Percent overcrowded	0.1595	0.2269	0.1819	0.2433

Note: Adjusted R^2 = 0.7309. Pseudo R^2 = 0.7147. The census block–level variables represent ordered categories because this is the only way block-level data from the Works Progress Administration are available. The value ranges change for each block-level variable and for each category, complicating interpretation of the beta coefficients. The beta coefficients can be interpreted as indicating what size change in the dependent variable would occur from a jump from one category to the next for each of these independent or control variables. For example, in a census block, an increase in the average rent from the second highest category ($30 to $49.99) to the highest category ($50 or more) brings an approximate increase of $267.40 to a mortgage in the East Falls/Germantown area. See table 3 for more detailed explanation of how the categories are defined for the block-level variables. OLS = ordinary least squares, SAR = spatial autoregressive.

$^{*}p < 0.1$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Table 9. **Regression Estimates for Predicting Amount of Mortgage for Kensington/Frankford**

	OLS Results		SAR Results	
Variable	Beta	Standard Error	Beta	Standard Error
Constant	0.6057	0.2857	0.5332	0.2985
House-level variables				
Sales price	0.4145	0.0374***	0.4035	0.0379***
Size of house	−0.0329	0.0297	−0.0253	0.0295
Owner occupied	0.2851	0.1455*	0.2523	0.1434*
Year of mortgage	0.1631	0.0208***	0.1707	0.0221***
Census block–level variables				
Percent nonwhite	−0.0897	0.2271	−0.1180	0.2373
Average rent	0.0029	0.0297	0.0128	0.0328
Percent substandard	−0.3088	0.0926***	−0.3465	0.0974***
Percent owner occupied	−0.0835	0.0390**	−0.0764	0.0422*
Percent overcrowded	0.2228	0.1795	0.3004	0.1887

Note: Adjusted R^2 = 0.7103. Pseudo R^2 = 0.6627. The census block–level variables represent ordered categories. The beta coefficients can be interpreted as indicating what size change in the dependent variable would occur from a jump from one category to the next for each of these independent or control variables. For example, in a census block, an increase in the percent owner occupied from the second highest category (50 to 74.9 percent) to the highest category (75 percent and higher) brings an approximate decrease of $83.50 to a mortgage in the Kensington/Frankford area. See table 3 for more detailed explanation of how the categories are defined for the block-level variables. OLS = ordinary least squares, SAR = spatial autoregressive.

$^{*}p < 0.1$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Table 10. **Regression Estimates Predicting Interest Rates for East Falls/Germantown**

	OLS Results		SAR Results	
Variable	Beta	Standard Error	Beta	Standard Error
Constant	5.7384	0.4226	5.6675	0.4364
House-level variables				
Mortgage amount	−0.0224	0.0120*	−0.0230	0.0120*
Size of house	−0.0057	0.0065	−0.0052	0.0064
Owner occupied	−0.1304	0.0821	−0.1133	0.0820
Year of mortgage	−0.0660	0.0128***	−0.0625	0.0135***
Census block–level variables				
Percent nonwhite	0.0698	0.0268***	0.0743	0.0281***
Average rent	0.0173	0.0575	0.0299	0.0599
Percent substandard	0.1393	0.0418***	0.1210	0.0433***
Percent owner occupied	−0.0266	0.0297	−0.0336	0.0318
Percent overcrowded	−0.1496	0.0763*	−0.1366	0.0793*

Note: Adjusted R^2 = 0.2579. Pseudo R^2 = 0.2663. The census block–level variables represent ordered categories. The beta coefficients can be interpreted as indicating what size change in the dependent variable would occur from a jump from one category to the next for each of these independent or control variables. For example, in a census block, an increase in the percent nonwhite from the second highest category (60 to 79.9 percent) to the highest category (80 percent and higher) brings an approximate increase of 0.07 to the interest rate on mortgages in the East Falls/Germantown area. See table 3 for more detailed explanation of how the categories are defined for the block-level variables. OLS = ordinary least squares, SAR = spatial autoregressive.

$^{*}p < 0.1$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Table 11. **Regression Estimates Predicting Interest Rates for Kensington/Frankford**

	OLS Results		SAR Results	
Variable	Beta	Standard Error	Beta	Standard Error
Constant	6.2439	0.1648	6.3030	0.1480
House-level variables				
Mortgage amount	−0.0928	0.0245***	−0.0948	0.0232***
Size of house	−0.0140	0.0173	−0.0081	0.0161
Owner occupied	−0.0051	0.0852	−0.0171	0.0828
Year of mortgage	−0.0477	0.0122***	−0.0499	0.0107***
Census block–level variables				
Percent nonwhite	−0.3223	0.1325**	−0.3638	0.1179***
Average rent	−0.0275	0.0172	−0.0301	0.0142**
Percent substandard	0.1423	0.0545***	0.1601	0.0477***
Percent owner occupied	0.0364	0.0227	0.0348	0.0193*
Percent overcrowded	−0.0765	0.1040	−0.0772	0.0916

Note: Adjusted R^2 = 0.2191. Pseudo R^2 = 0.3908. The census block–level variables represent ordered categories. The beta coefficients can be interpreted as indicating what size change in the dependent variable would occur from a jump from one category to the next for each of these independent or control variables. For example, in a census block, an increase in the percent substandard from the second highest category (70 to 84.9 percent) to the highest category (85 percent and higher) brings an approximate increase of 0.14 to the interest rate on mortgages in the Kensington/Frankford area. See table 3 for more detailed explanation of how the categories are defined for the block-level variables. OLS = ordinary least squares, SAR = spatial autoregressive.

$^{*}p < 0.1$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Testing Effect of HOLC Grades Using Spatial Regression

Hillier (2003), using OLS and spatial lag models, analyzed the effect of HOLC's 1937 residential security map for Philadelphia on the random citywide and small-area samples. The number of mortgages made on a property, the amount of the mortgage, and the interest rate were used as dependent variables. HOLC grade and the distance from an area colored red (fourth grade, or "hazardous") by HOLC served as independent variables. The same property, transaction, and area variables used as control variables to test the effect of racial composition in the previous section were used. Results indicated that HOLC grade has no significant effect on the amount of the mortgage, but mortgages in areas colored red have significantly higher interest rates.

Testing for Significant Sparseness of Mortgages with Hot Spot Analysis

Hot spot analyses use a range of methods to identify significant clustering within point patterns. Ripley's local K function is an especially powerful test, determining where and at what scale significant clustering exists within a point pattern, relative to what is expected under complete spatial randomness (Bailey and Gatrell 1995; Cressie 1993). This differs from the global K function, which determines whether significant clustering exists within a point pattern at a particular scale but does not identify where the significant clustering takes place. K functions generally are used to identify significant clustering—in the context of mortgage lending, areas that were potentially overserved. The K function also can be used to identify significant sparseness—in the context of mortgage lending, areas that were potentially underserved.

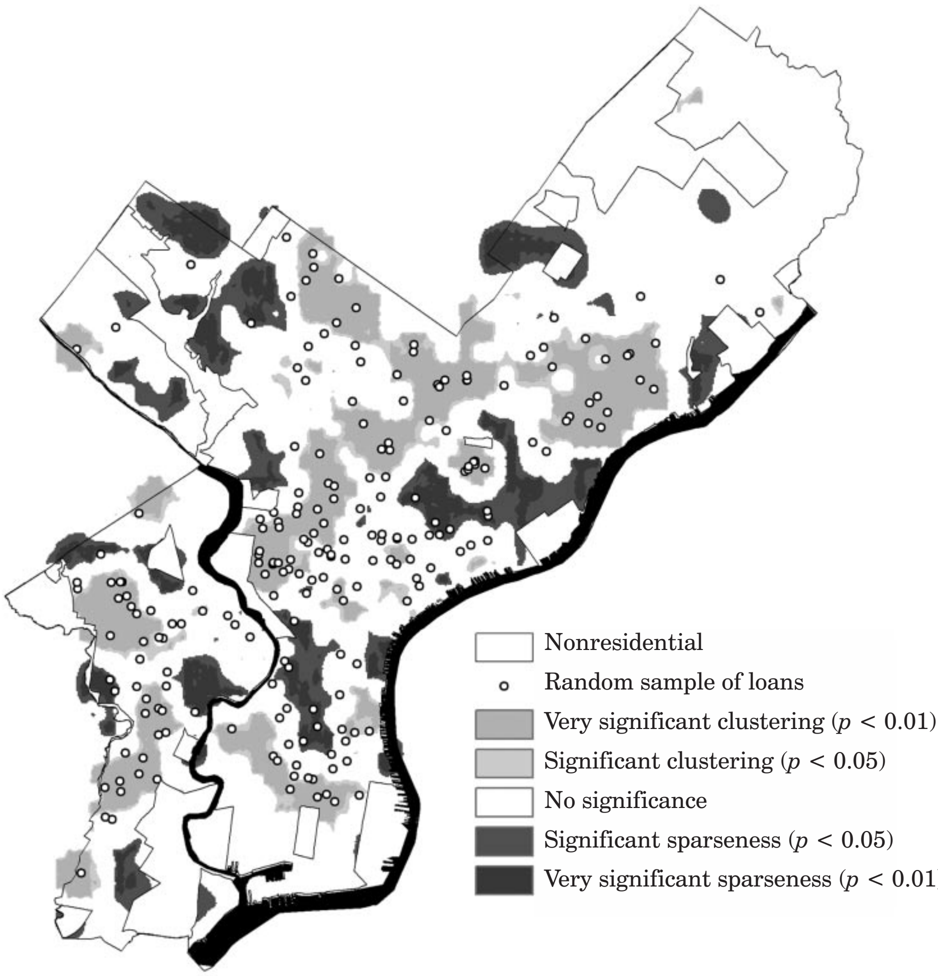
Using a local K function, the random citywide sample of mortgages for Philadelphia (N = 336, because unlike the regression analyses, a property does not need complete data to be included) was compared with 999 randomly generated patterns to determine whether there are significantly more—or significantly fewer—mortgages in particular areas than expected. In the absence of information about mortgage applications, the number of occupied housing units by census tract according to the 1940 census is used as a "backcloth" to represent the expected number of mortgages. A bandwidth of 0.6 miles was chosen, meaning that the local K function counts the number of mortgages within 0.6 mile of a regular grid of points that extends across Philadelphia. This fairly small scale generates more conservative results than bandwidths of a mile or more, which designate most parts of the city as either significantly clustered or sparse. The K function calculates the significance of clustering or sparseness for each grid point. These values were then used to create contours and interpolate a surface through ordinary Kriging (using a search radius of 12 points).[2]

The result is figure 3, which shows areas with statistically significant clustering and sparseness, given the number of occupied housing units. Areas with significantly fewer mortgages than expected include a stretch of the central part of the city, extending from Lower North Philadelphia through Center City and into South Philadelphia; small areas in West and

[2] The local K function was computed using "k_count_loc" written for Matlab by Tony E. Smith.

Southwest Philadelphia; a fairly large part of the Kensington/Frankford area; and in the Northwest, parts of lower Roxborough/Manayunk and Mount Airy/Chestnut Hill.

Figure 3. **Significant Clustering and Sparseness of Mortgages in Philadelphia, 1940 to 1950**



Note: The results of Ripley's local K function, analyzing the random citywide sample of mortgages, show areas of significantly more and significantly fewer mortgages than expected, given the distribution of occupied housing units in Philadelphia.

Testing for Areas of High Interest Rates Using Surface Interpolation

Surface interpolation provides another exploratory tool for spatial data analysis of lending data. Methods such as Kriging and inverse distance weighting are used extensively in the natural sciences to estimates values for spatially continuous phenomena such as precipitation, elevation, and soil composition based on values at sample locations (Bailey and Gatrell 1995; Cressie 1993). Whereas lending to a particular property is a discrete spatial event, one might conceptualize the availability of mortgages or the terms of the mortgages as spatially continuous. Interest rates for the random citywide sample provide the best data for surface interpolation. Figure 4 shows the results of ordinary Kriging using the five nearest mortgages to estimate the interest rate for areas where there are no sample mortgages.[3] The map indicates that the central part of Lower North Philadelphia has the highest interest rates, followed by much of the surrounding area, into West and South Philadelphia.

Comparing Results across Methods

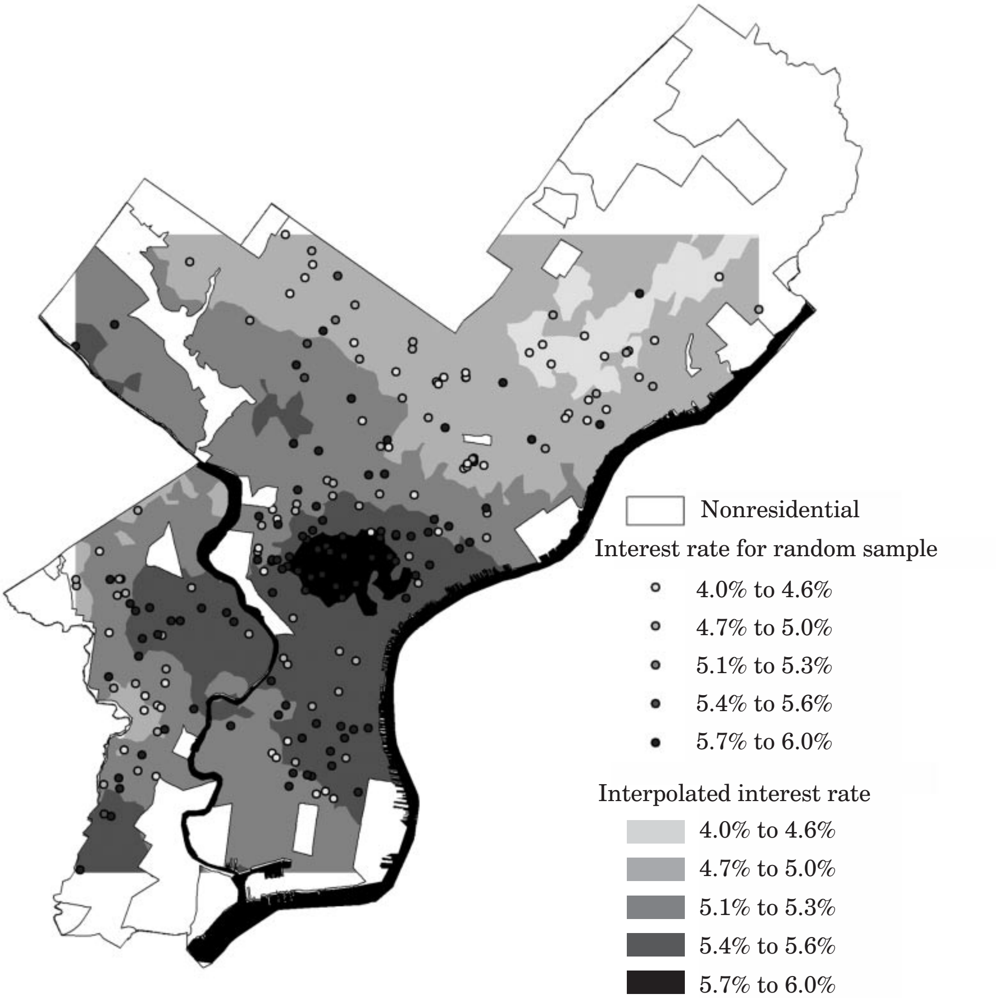
Collectively, the results from these four analyses suggest that several parts of the city were at a distinct disadvantage in the mortgage lending process. No parts of the city meet all four of the standards that these methods adopt—a high concentration of African Americans, colored red by HOLC, significantly sparse mortgages, and high interpolated interest rates. However, many parts meet two or three of the standards. Results across the four methods are shown as overlapping map layers in figure 5. The numbered areas indicate places that meet multiple criteria for potential redlining.

Area 1, located in the eastern part of West Philadelphia across the Schuylkill River from Center City, is made up of the neighborhoods now known as Mantua, Powelton, and University City. This section had more than 25 percent nonwhite population, was graded “hazardous” by HOLC, and paid relatively high interest rates. Results of the hot spot analysis do not indicate that this area received fewer mortgages than expected, given the number of owner-occupied housing units. The northern part of this area, Mantua, has become one of the poorest in the city and is home to a public housing development. This area is a primary target of demolition and redevelopment efforts through Philadelphia’s Neighborhood Transformation Initiative, launched in 2002. The southern portion of this area, on the other hand, has witnessed substantial gentrification around Drexel University and the University of Pennsylvania.

Area 2, located in the western part of Lower North Philadelphia, is made up of the North Central, Sharswood, and Brewerytown neighborhoods. This section had more than 25 percent nonwhite population, was graded “hazardous” by HOLC, and paid relatively high interest rates. A very small part of this area had significantly fewer mortgages than expected. This area has lost a significant amount of population and is one of the poorest parts of the city. Four public housing developments were built in this general vicinity, one of which was imploded in the mid-1990s. The area is currently the target of much of the city’s subsidized housing construction program.

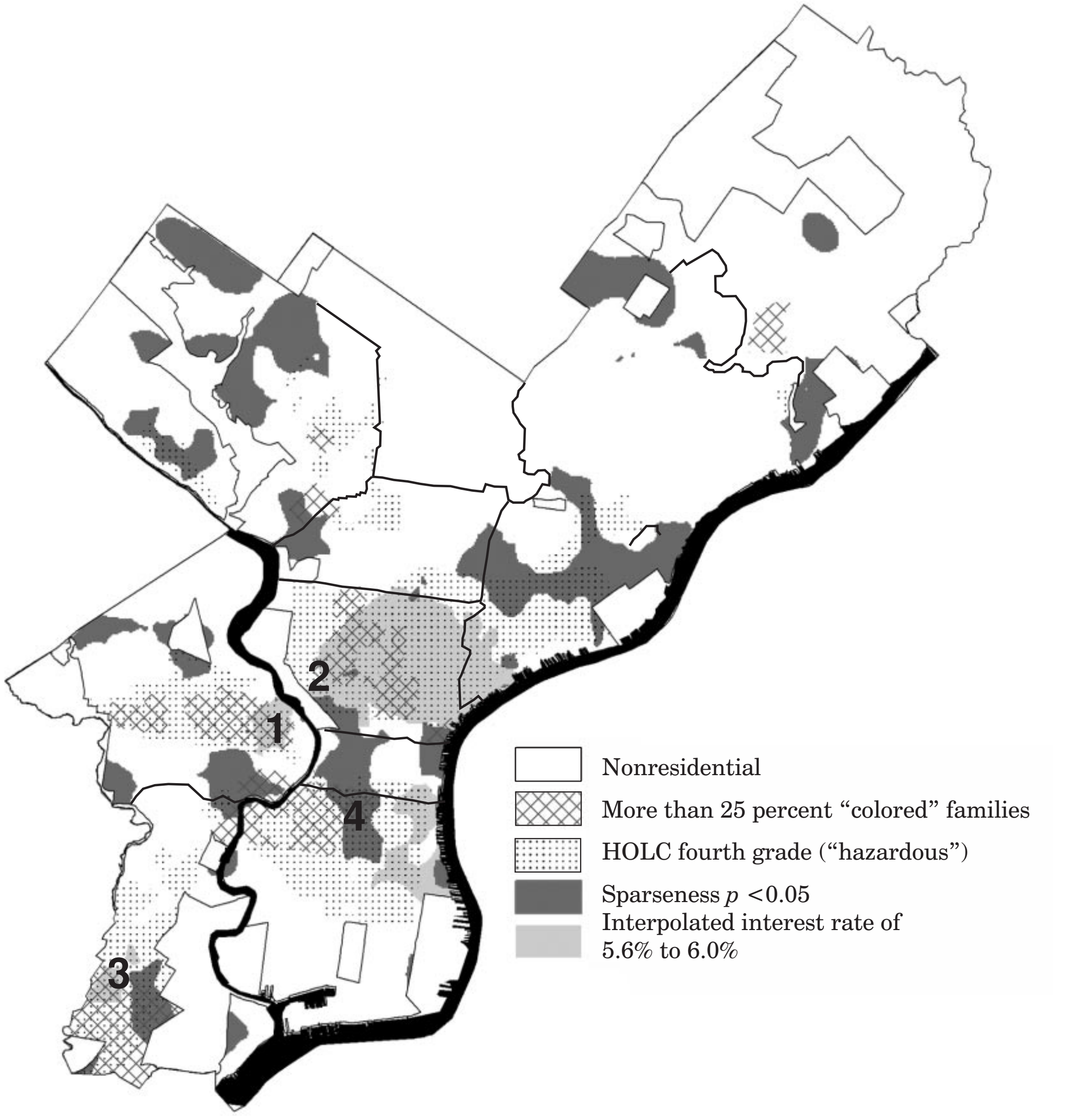
[3] The ordinary Kriging was computed using the Geostatistical Analyst extension for ArcView 8.1.

Figure 4. **Interpolated Surface of Interest Rates in Philadelphia, 1940 to 1950**



Note: The map shows the interest rates on a random citywide sample of mortgages against a surface of interest rates predicted from the sample points using Kriging. Areas where both the point and surface data correspond (where both are the same shade of gray or black) show where the prediction is most consistent with the actual data. Within the areas displayed as black, the points that represent 5.7 to 6.0 percent are shown as dark gray.

Figure 5. **Comparison of Results across Methods**



Note: Four different parts of the city meet three of the criteria for redlining used in the spatial regression models, hot spot analysis, and surface interpolation, indicating areas that may have been redlined.

Area 3 is located in deep Southwest Philadelphia in what is now known as Eastwick. This section had more than 25 percent nonwhite population and was graded "hazardous" by HOLC. Part of this area paid higher interest rates, and a larger part had fewer mortgages than expected. This area was the primary focus of Philadelphia's urban renewal efforts in the 1950s, which completely transformed the area.

Area 4 is located just south of Center City in what is now known as Point Breeze, Southwest Center City, Hawthorne, and a commercial corridor known as the "Avenue of the Arts." This area was graded "hazardous" by HOLC and had fewer mortgages than expected. A smaller part of the area had more than 25 percent nonwhite population, but none of the area appears to have paid higher interest rates. The Martin Luther King public housing development was built on the eastern side of this area. The high-rise towers that made up this development were imploded in the late 1990s, making room for a new HOPE VI mixed-income housing development.

Discussion

The results of the four different methods suggest that several areas in Philadelphia—all areas that have suffered from disinvestment and have been the target of public redevelopment efforts—were at a disadvantage in securing mortgages. However, the methodological implications of this research are more far-reaching. The methods presented offer a way of analyzing the limited lending data available for studies of historical redlining, highlight the importance of scale in investigations of redlining, and emphasize the importance of spatial analysis to understanding redlining—in both historical and contemporary research.

The data analyzed for this article have significant limitations, most notably the lack of information about rejected mortgage applications. Future research is unlikely to uncover historical data about mortgage applications and credit histories, so researchers need to make the most of the data available, just as they did before the 1989 HMDA amendments. Given the limitations of data sets such as the one constructed from the *Philadelphia Realty Directory and Service*, it is advisable that researchers conceptualize redlining broadly—as involving process or outcome, spatial or statistical relationships. Using multiple methods to identify patterns that can be analyzed in map layers avoids looking for definitive evidence of redlining from any one test and increases the likelihood of identifying lending patterns.

The results from the SAR model and local K function highlight the importance of scale in investigations of redlining. In the tests of the effect of racial composition, the analysis of small areas using block-level data finds race to be very significant in explaining interest rates ($p < 0.05$), whereas the citywide analysis using tract-level data finds race to be only marginally significant. Although Philadelphia was an intensely segregated city throughout the second half of the 20th century, racial and housing characteristics often still vary from block to block. The literature on the modifiable area unit problem focuses on the difficulty researchers across disciplines have in capturing aggregate data at an appropriate scale (Cressie 1993; Longley and Batty 1996).

The local K function results are also highly dependent on scale. The selection of a bandwidth, the search radius for counting the number of points from each regular grid point, has a significant effect on determining which areas are identified as having significant clustering and sparseness. Scale also is relevant to the selection of a backcloth, which is used to represent the expected spatial distribution. The number of occupied units by census tract is used as the backcloth in this analysis. A backcloth represented by smaller aggregated areas (such as census blocks or block groups), or even a continuous surface, allows for a much smaller bandwidth, making it possible to identify areas of significant clustering and sparseness at a much smaller scale.

These four tests for historical redlining—assessing the effect of racial composition, assessing the effect of red lines, identifying areas of significant sparseness, and identifying areas with worse terms—hold promise for future research. The local K function and Kriging likely would perform better with larger random samples, possibly comparing results across years. The local K function analysis might be used to compare the distribution of transactions in which owners received a mortgage with the distribution of all property transactions, as well as to compare the distribution of mortgage defaults with the distribution of all mortgages.

Conclusion

This work demonstrates the integral relationship between conceptualization of redlining and methods of analysis. Reviewing the range of definitions of redlining used in research on contemporary redlining opens up possibilities for testing for historical redlining, even in the face of severe data limitations. This research also underscores how difficult it is to prove that redlining occurred. Ruling out alternative explanations to discrimination is even more difficult in an historical context, given the limited data available about mortgage lending. Without the needed data, researchers of both contemporary and historical redlining have found themselves at what George Galster has called an “investigative cul-de-sac” (Galster 1993, 299).

Just as researchers of contemporary redlining have persevered, the need for empirical research of historical redlining should not be dismissed because the ideal multivariate model is out of reach. Just as researchers of the Boston Fed studies used a variety of statistical models to determine if their results were robust (Munnell et al. 1996; Yinger 1995), researchers of historical lending patterns should try multiple tests to determine whether—and where—redlining took place. These conceptualizations and tests of historical redlining are not in competition with one another, and their greatest value may be when they are considered together. The research presented here is intended more as a methodological exploration than a definitive test of redlining in Philadelphia, but these preliminary results do show signs of differential access to loans. Redlining involves a pattern of lending discrimination, so multiple tests over different time periods in different cities are needed to determine its extent and effect.

Although this research supports a broad conceptualization of redlining and encourages the use of a variety of methods for identifying historical as well as contemporary redlining, it serves primarily as a call for thinking about redlining as a spatial phenomenon. Redlining is place-based discrimination, and it must be identified through spatial analyses. Spatial relationships can be considered in regression analysis through the continued inclusion of area (census tract) data, as well as through spatial weight matrices. They also can be studied more directly through hot spot and surface interpolation methods in conjunction with GISs. An emphasis on the spatial aspects of lending discrimination is critical to investigations of historical redlining that seek to explain why, how, and where disinvestment occurred in cities such as Philadelphia.

This emphasis on spatial relationships is equally relevant to studies of contemporary redlining. Researchers' elusive search for the ideal statistical model to test for redlining takes attention away from efforts to identify actual neighborhoods—not just neighborhood profiles—where redlining occurs. Address-level mortgage information, which is much more accessible for studies of contemporary redlining, can be mapped and analyzed using the fairly simple spatial methods identified in this article. By creating map layers that identify underserved communities, researchers and fair housing advocates can target efforts to address discrimination and create new opportunities for homeownership, wealth accumulation, and community development.

References

Abrams, Charles. 1955. *Forbidden Neighbors: A Study of Prejudice in Housing.* New York: Harper.

Adams, Carolyn, David Bartelt, David Elesh, Ira Goldstein, Nancy Kleniewski, and William Yancey. 1991. *Neighborhoods, Division, and Conflict in a Postindustrial City.* Philadelphia: Temple University Press.

Avery, Robert B., Patricia E. Beeson, and Mark S. Sniderman. 1999. Neighborhood Information and Home Mortgage Lending. *Journal of Urban Economics* 45:287–310.

Bailey, Trevor C., and Anthony C. Gatrell. 1995. *Interactive Spatial Data Analysis.* Essex, England: Longman Scientific and Technical.

Bartelt, David, David Elesh, Ira Goldstein, George Leon, and William Yancey. 1987. Islands in the Stream: Neighborhoods and the Political Economy of the City. In *Neighborhoods and Community Environments*, eds. Irwin Altman and Abraham Wandersman, 163–89. New York: Plenum.

Bissinger, Buzz. 1997. *A Prayer for the City.* New York: Random House.

Black, Harold A. 1999. Is There Discrimination in Mortgage Lending? What Does the Research Tell Us? *Review of Political Economy* 27(1):23–30.

Bluestone, Barry, and Bennett Harrison. 1982. *The Deindustrialization of America: Plant Closings, Community Abandonment, and the Dismantling of Basic Industry*. New York: Basic Books.

Bogdon, Amy S., and Carol A. Bell. 1999. Making Fair Lending a Reality in the New Millennium. Proceedings of the Fannie Mae Foundation research roundtable, Washington, DC, June 30.

Bradbury, Katherine L., Karl E. Case, and Constance R. Dunham. 1989. Geographic Patterns of Mortgage Lending in Boston, 1982–1987. *New England Economic Review* September/October: 3–29.

Brewer, J. M. 1934. Survey of Philadelphia. Map collection. Philadelphia: Free Library of Philadelphia.

Calem, Paul S. 1996. Mortgage Credit Availability in Low- and Moderate-Income Neighborhoods: Are Information Externalities Critical? *Journal of Real Estate Finance and Economics* 13(1):71–89.

Canner, Glenn B. 1982. *Redlining: Research and Federal Legislative Response.* Washington, DC: Board of Governors of the Federal Reserve System.

Chicago Commission on Race Relations. 1922. *The Negro in Chicago: A Study of Race Relations and a Race Riot.* Chicago: University of Chicago Press.

Cohen, Lizabeth. 1990. *Making a New Deal: Industrial Workers in Chicago, 1919–1939.* New York: Cambridge University Press.

Cressie, Noel C. 1993. *Statistics for Spatial Data.* New York: Wiley.

Ellen, Ingrid Gould, and Margery Austin Turner. 1997. Does Neighborhood Matter? *Assessing Recent Evidence. Housing Policy Debate* 8(4):833–66.

Federal Home Loan Bank Board. 1933–1940. Records of the City Survey Program. Record group 195, box 450. College Park, MD: National Archives II.

Federal Housing Administration. 1935. *Underwriting Manual.* Washington, DC.

Galster, George C. 1993. Use of Testers in Investigating Discrimination in Mortgage Lending and Insurance. In *Clear and Convincing Evidence: Measurement of Discrimination in America,* eds. Michael Fix and Raymond J. Struyk, 287–334. Washington, DC: Urban Institute Press.

Harrison, David M. 2001. The Importance of Lender Heterogeneity in Mortgage Lending. *Journal of Urban Economics* 49(2):285–309.

Harvey, David. 1985. *The Urbanization of Capital.* Baltimore, MD: Johns Hopkins University Press.

Hillier, Amy E. 2003. Redlining and the Home Owners' Loan Corporation. *Journal of Urban History* 29(4):394–420.

Holloway, Steven R. 1998. Exploring the Neighborhood Contingency of Race Discrimination in Mortgage Lending in Columbus, Ohio. *Annals of the Association of American Geographers* 88(2):252–76.

Holloway, Steven R., and Elvin K. Wyly. 2001. "The Color of Money" Expanded: Geographically Contingent Mortgage Lending in Atlanta. *Journal of Housing Research* 12(1):55–90.

Home Mortgage Disclosure Act. 1975. 12 U.S.C. Section 2801.

Hoyt, Homer. 1939. *The Structure and Growth of Residential Neighborhoods in American Cities.* Washington, DC: Federal Housing Administration.

Jackson, Kenneth T. 1980. Race, Ethnicity, and Real Estate Appraisal: The Home Owners' Loan Corporation and the Federal Housing Administration. *Journal of Urban History* 6(4):419–52.

Jackson, Kenneth T. 1985. *Crabgrass Frontier: The Suburbanization of the United States.* New York: Oxford University Press.

Lang, William W., and Leonard I. Nakamura. 1993. A Model of Redlining. *Journal of Urban Economics* 33(2):223–34.

Ling, David C., and Susan M. Wachter. 1998. Information Externalities and Home Mortgage Underwriting. *Journal of Urban Economics* 44(3):317–32.

Longley, Paul, and Michael Batty. 1996. *Spatial Analysis: Modeling in a GIS Environment.* Cambridge, England: GeoInformation International.

Marcuse, Peter. 1979. The Deceptive Consensus on Redlining: Definitions Do Matter. *Journal of the American Planning Association* 45(4):549–56.

Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid.* Cambridge, MA: Harvard University Press.

Metzger, John T. 1999. Social Capitalism in American Cities. Unpublished paper. Columbia University.

Mohl, Raymond A. 1987. Trouble in Paradise: Race and Housing in Miami during the New Deal Era. *Prologue* 19:7–21.

Mohl, Raymond A. 1997. The Second Ghetto and the "Infiltration Theory" in Urban Real Estate, 1940–1960. In *Urban Planning and the African American Community: In the Shadows,* eds. June Manning Thomas and Marsha Ritzdorf, 58–74. Thousand Oaks, CA: Sage.

Mohl, Raymond A. 1999. H-Urban posting, April 14. World Wide Web page <http://www2.h-net.msu.edu/~urban/> (accessed June 7, 2003).

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney. 1996. Lending in Boston: Interpreting HMDA Data. *American Economic Review* 86(March):25–53.

National Commission on Urban Problems. 1969. *Building the American City: Report of the National Commission on Urban Problems.* New York: Praeger.

Philadelphia Housing Authority and Works Projects Administration of Pennsylvania. 1939. Real Property Survey. Philadelphia.

Philadelphia Realty Directory and Service. 1938–1950. Philadelphia: Philadelphia Real Estate Directory Incorporated.

Phillips-Patrick, Fred J., and Clifford V. Rossi. 1996. Statistical Evidence of Mortgage Redlining? A Cautionary Tale. *Journal of Real Estate Research* 11(1):13–23.

Pogge, Jean. 1992. Reinvestment in Chicago Neighborhoods: A Twenty-Year Struggle. In *From Redlining to Reinvestment,* ed. Gregory Squires, 133–148. Philadelphia: Temple University Press.

President's National Advisory Panel on Insurance in Riot-Affected Areas. 1968. *Meeting the Insurance Crisis of Our Cities.* Washington, DC: U.S. Government Printing Office.

Reibel, Michael. 2000. Geographic Variation in Mortgage Discrimination: Evidence from Los Angeles. *Urban Geography* 21(1):45–60.

Ross, Stephen L., and John Yinger. 2002. *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement.* Cambridge, MA: MIT Press.

Schill, Michael H., and Susan M. Wachter. 1993. A Tale of Two Cities: Racial and Ethnic Geographic Disparities in Home Mortgage Lending in Boston and Philadelphia. *Journal of Housing Research* 4(2):245–75.

Schisgall, Oscar. 1975. *Out of One Small Chest: A Social and Financial History of the Bowery Savings Bank.* New York: Amacom.

Shlay, Anne B. 1989. Financing Community: Methods for Assessing Residential Credit, Disparities, Market Barriers, and Institutional Reinvestment Performance in the Metropolis. *Journal of Urban Affairs* 11(3):201–23.

Smith, L. Elden. 1938. Measuring the Neighborhood Risk. *Insured Mortgage Portfolio* 2(8):9–11, 22–23.

Smith, Niel, Paul Caris, and Elvin Wyly. 2001. The "Camden Syndrome" and the Menace of Suburban Decline: Residential Disinvestment and Its Discontents in Camden County, New Jersey. *Urban Affairs Review* 36(4):497–531.

Squires, Gregory. 1994. *Capital and Communities in Black and White.* Albany: State University of New York Press.

Squires, Gregory D., and Sally O'Connor. 2001. *Color and Money.* Albany: State University of New York Press.

Tootell, Geoffrey M. B. 1996. Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods? *Quarterly Journal of Economics* 111(4):1049–79.

Turner, Margery Austin, and Felicity Skidmore. 1999. *Mortgage Lending Discrimination: A Review of Existing Evidence.* Washington, DC: Urban Institute Press.

Works Progress Administration. 1934. Philadelphia Real Estate Property Survey. Philadelphia.

Wyly, Elvin K. 2002. Mortgaged Metropolis: Evolving Urban Geographies of Residential Lending. *Urban Geography* 23(1):3–30.

Yinger, John. 1995. *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination.* New York: Russell Sage Foundation.

Zenou, Yves, and Nicolas Boccard. 2000. Racial Discrimination and Redlining in Cities. *Journal of Urban Economics* 48:260–85.