PROPOSAL

The Distribution of Federally-Insured Mortgages:
Evidence from Land Records in Mecklenburg County, NC*

Omer Ali (University of Pittsburgh)
Andreas Ferrara (University of Pittsburgh)
Jonathan Rose (Federal Reserve Bank of Chicago)
Randall Walsh (University of Pittsburgh)

November 21, 2024

Introduction

Systemic racism in the United States is perhaps best exemplified by the stratified urban environments where Black and white households live apart (Gordon and Bruch, 2019; Ali et al., 2022; Bayer, Charles, and Park, 2019). Recent debates about the legacy of federal redlining have emphasized the role of public policies in congealing segregation and precipitating inequalities in homeownership and wealth. While acknowledging the role of individual decision making in residential choices as well as broader disadvantages faced by African Americans, authors and scholars have also made the case that federal housing policies supported the development of segregated neighborhoods and suburbs and funneled subsidized mortgages to white households

Coates links the unequal distribution of federally-insured mortgages beginning in the 1930s directly to the case for reparations (Coates, 2014). In a widely read book, Rothstein pushes in the same direction, taking aim at legal precedent that views segregation as a de-facto outcome of many individual decisions made by households that, unfortunately, accrete into stratified outcomes (Rothstein, 2017). Instead, he proposes a different mechanism: throughout the twentieth century, federal, state and local governments used public policies to engineer segregated metropolitan areas by directing federally insured mortgages towards white households, developing homo-

*Support from Russel Sage Foundation grant G-2310-45068 is gratefully acknowledged. We thank the following students for excellent research assistance: Mardoché Ayitchehou, Gabrielle Toborg, Leyla Ahmadova, Max Barrett, Kasey Brown, Samantha Candalla, Nick Cao, Roan Chiarilli, Megan Duda, Chazz Ehlis, Claire Gatz, Stephanie Ijomor, Adam Kaufman, Sally Margalitachvili, Katie Power, Karthik Sivachandran, Luke Smarra, Isabel Spotz, Simon Wang, and Ilan Ziv. The views expressed do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

geneous suburbs, using Federal Highway Act projects to build physical barriers between white and Black neighborhoods and using urban renewal as a tool for removing Black residents from land deemed too valuable for their occupancy.

Some of these claims have received empirical support elsewhere (LaVoice, 2024; Valenzuela-Casasempere, 2024). This project, however, is concerned with the central claim in the redlining debate that pertains to the distribution of federally-insured mortgages: who received them and where were they issued? Given the importance of this topic for contemporary debates about the persistence of racial inequality, it is remarkable that these two questions remain unanswered.

Background

The three federal agencies implicated in this history are the Home Owners Loan Corporation (HOLC), the Federal Housing Administration (FHA), and the Veterans Administration (VA). Briefly, the HOLC was introduced during the housing crisis of the Great Depression in order to save home-owners who were at risk of default. The agency went on to refinance roughly a million mortgages and prevent their foreclosure. Agency records show that Black neighborhoods were meaningfully served by the emergency refinancing program (Michney and Winling, 2019). Once the program ended, HOLC staff were left with a portfolio of mortgage holdings that needed to be managed. To aid in their decision-making, the Mortgage Rehabilitation Department began drafting a series of now-infamous residential security maps rating neighborhoods on a 4-letter scale in roughly 240 urban areas in the United States (Michney, 2021; Hynsjö and Perdoni, 2024). Black neighborhoods were uniformly rated with the worst grade, D (red).

The FHA was founded a year later and was primarily involved in insuring home improvement and new residential home loans against the risk of default. Lenders were reluctant to issue new mortgages at this time as the United States was in the middle of a deep economic recession. To encourage the flow of credit, the federal government offered to insure lenders against default risk as long as mortgages met certain criteria that made repayment easier for borrowers. Thus, FHA-insured mortgages had lower down-payments than conventional mortgages, lower interest rates, and longer terms. These features made them popular among borrowers and starting from a relatively small number of insured mortgages in 1935 (23,397), the FHA went on to insure roughly 150,000 new mortgages annually by the end of the decade. By 1942, it served one fourth of the entire mortgage market (Freund, 2007). White households accounted for 97.7% of insured single family mortgages by 1950 (Table SS, 1950 Census of Housing, (Bureau, 1950)).

Since the FHA shifted the risk of default from the lender onto itself, it sought to develop a systematic evaluation of mortgage risk in order to ensure that losses from defaults were met by mortgage insurance payments. They developed underwriting manuals that instructed staff on best practices in the adjudication of applications for mortgage insurance. Four broad categories of risks were highlighted: borrower, structure, neighborhood, and mortgage market. Applications

were more likely to be approved if neighborhoods were desirable and, in particular, had stable and uniform racial compositions. The following quotes are from the FHA's 1936 underwriting manual.

> "the Valuator should investigate areas surrounding the location to determine whether or not incompatible racial and social groups are present."

> "If a neighborhood is to retain stability it is necessary that properties shall continue to be occupied by the same social and racial classes. A change in social or racial occupancy generally leads to instability and a reduction in values." (Section 233 of the 1936 FHA Underwriting Manual)

Together with the criteria governing the construction quality and age of structures and the socio-economic characteristics of neighborhoods, the FHA primarily served white households seeking to purchase properties in moderate and high income neighborhoods and newly constructed suburbs. Indeed, the agency encouraged the adoption of racially restrictive covenants since these were seen as a way to ensure (legally) that racial composition would remain unchanged.

The absence of individual-level data on the identity and location of borrowers who received insured mortgages makes it difficult to characterize the extent of racial discrimination in insurance provision and to effectively study the effects of these policies. With the exception of Fishback, Rose, et al., 2022, data on FHA and VA mortgages are only available at the state, metro area or county level, which obscures variation within these larger geographical areas. We build on the work in Fishback, Rose, et al. (2022), who collect mortgage records from Baltimore, MD, Greensboro, NC, and Peoria, IL between 1935 and 1940, by studying mortgages in Charlotte, NC between 1935 and 1975. In that study, the authors find that less than 1% of FHA borrowers were Black in each of the cities they examine. Our results are similar: we find that 1.5% of FHA borrowers in Mecklenburg County were Black, although Black households constitute 27.6% of the county.

The third federal agency implicated in the history of redlining is the Veterans' Administration. Following the end of World War II, the Servicemen's Readjustment Act of 1944 (G.I. Bill) introduced the Home Loan Guarantee program, which allowed returning veterans to purchase homes using federally guaranteed mortgages. The program was modeled on the FHA's mortgage insurance program and included more generous provisions such as lower interest rates and down-payments. Scholars have argued that these benefits were distributed in a way that disfavored Black veterans relative to their white counterparts (Katznelson, 2006). To date, however, no study has provided evidence for these claims using records of issued mortgages as we aim to do in this paper. Those results are pending the linking of mortgage records from 1950 to the newly available 1950 de-anonymized decennial census.

Data

Data on historic mortgages in the United States is difficult to obtain because each county is responsible for recording and maintaining their own real estate records. Although recent mortgages are more likely to be available in digitized form, the vast majority of counties still maintain paper copies of older records, including mortgages. The few that have digitized their archives may not allow access that facilitates the collection of large numbers of documents, which is required for this study.[1]

Mecklenburg County in North Carolina, which contains the state's largest city, Charlotte, is one of the few counties that has digitized their historic mortgage records and allows cost-free access through a navigable website. We leveraged this opportunity to scrape all mortgages issued in the county in the following years: 1935-40, 1950, 1960, 1970, and 1975.

Once obtained, we identified FHA-insured mortgages and, from 1950 onward, VA-guaranteed mortgages by searching for certain key terms. This was done by recognizing that these mortgages were always printed on standardized forms (see Figure 4 for an example of an FHA-insured mortgage). As a result, FHA and VA mortgages contain specific terms that can be matched in the text output from an OCR processor. Documents identified in this way are then manually checked to remove false positives. In a sample of mortgage books that were manually searched for FHA mortgages, we find that the false negative rate of this procedure is less than 0.5% making us confident that we are able to identify the near-universe of all FHA-insured and VA-guaranteed mortgages.

Having identified the subset of federally-insured mortgages among the trove of other real estate records, we digitized their contents by collecting borrower, mortgage, and location information. We collect the name, and spouse name, mortgage amount, interest rate and lender, as well as the legal description of the property, among other details. The legal description can be linked to a (modern) street address using the county's GIS system or, alternatively, by looking up the referenced plat map in the county's records and comparing it with a modern map with the same streets.

Finally, to facilitate linking to the census, we sometimes need to change the street names since some streets undergo a name change between the 1930s and the present. For this step, we leverage the Sanborn Fire Insurance Company's maps of urban areas to match historic street names with their modern analogues.

Linking borrowers to the census

To match FHA borrowers to the census, we employ a process that combines automated techniques and manual verification to remove false positives and ensure that matches are reliable. First, we clean and standardize the names in both datasets. To do this, we use the Abramitzky, Boustan,

[1] For example, some counties charge users by the document, which makes data collection infeasible due to the costs associated with the acquisition of many records.

and Eriksson (ABE) method, developed over a sequence of papers (Abramitzky et al., 2019). This fully automated approach is designed for linking historical datasets, such as complete-count censuses, by harmonizing names and accounting for variations in spelling, transcription errors, and other inconsistencies. This step ensures that the same names in the two datasets being linked will have the same spelling. In particular, we run the $clean_names$ routine in the R package $abeR$ for both FHA borrowers as well as all census respondents in Mecklenburg County in the 1940 census (Ruggles et al., 2024).

Similarly, we clean and standardize street names in both datasets. Address information in the census and the mortgage records often contains variations in formatting, spelling, and abbreviations. By standardizing these entries, we reduce discrepancies and improve the likelihood of accurate matches in the subsequent steps. There is no standard routine for implementing this step, so we do so iteratively, examining the output after each round of standardization. The goal of this step is to reach a set of street names that are simple in order to reduce the opportunities for mismatches when matching exactly, but do not result in false positives. For example, we remove suffixes such as road, street, boulevard etc., but keep and harmonize directional prefixes such as North, South, East and West. The idea behind this choice is that while suffixes may prevent a true match, prefixes likely distinguish between streets that are in practice different. Removing them would likely result in false positives.

Once the data are cleaned and standardized, we proceed to matching the records. The matching process is carried out in a series of steps with increasingly flexible criteria. We begin with 2,129 individuals, who represent the near universe of FHA borrowers.[2] Initially, we match records exactly based on the borrower's first name, last name, and street name. This step alone achieved a matching rate of 12%. We do not use the street number field for any of the matching steps. This provides the most restrictive matching condition. We then proceed to match using the spouse's first name, last name, and the street name, which contributes another 12%. For matches involving the spouse, we manually review the results to ensure that the spouse being matched is the correct one. Borrowers who are successfully matched are removed from the pool at each step, once we establish that the matches are not false positives. The third step matches using the individual's first name, their spouse's first name, and their last name, adding another 10%.

After completing these three steps, we introduce middle names to add additional pieces of information in the matching process. Using middle names, we proceed in the following two ways: matching by the individual's first name, middle name, and last name; and matching by the spouse's first name, middle name, and last name, which also requires manual checks to verify that the spouse matches are accurate. Together, these five steps collectively achieved a cumulative matching rate of 42%. All matched borrowers are removed.

Up to this point, we insist on exact matches. Next, we employ the Jaro-Winkler (JW) string sim-

[2]We compare the number of FHA-insured mortgages we found in the Mecklenburg County land records with official statistics reported by the FHA in agency reports for the Charlotte, NC metropolitan area.

ilarity algorithm to relax this criteria. The Jaro-Winkler method allows us to identify matches even when minor discrepancies in spelling or formatting are present. For each of the five matching criteria (first name, last name, street name, spouse name and middle name), we gradually increase the JW distance threshold to capture more potential matches. Along with the best match, we also retrieve the second-best match for every record to provide an alternative. This step generates matches for an additional 35% of borrowers.

After completing the automated matching, we manually review the matches to ensure accuracy. During this step, we carefully verify the records to remove false positives where the matched records clearly do not correspond to the same individual. Additionally, we assess whether the second-best match is a more accurate fit than the initial match. After removing the false positives, the matching rate declines to 71%. Exact matching accounts for 60% of the matched sample, while the JW algorithm accounts for the remaining 40%.

In summary, our methodology combines automated techniques, including the ABE method and Jaro-Winkler matching, with manual validation to produce a dataset that achieves relatively high match rates without false positives.[3] The removal of false positives is particularly important in our case since one of our main research questions concerns the extent to which Black home-owners are under-represented among FHA borrowers. In total, 1,503 borrowers are linked, 22 of whom are Black.

Preliminary Results

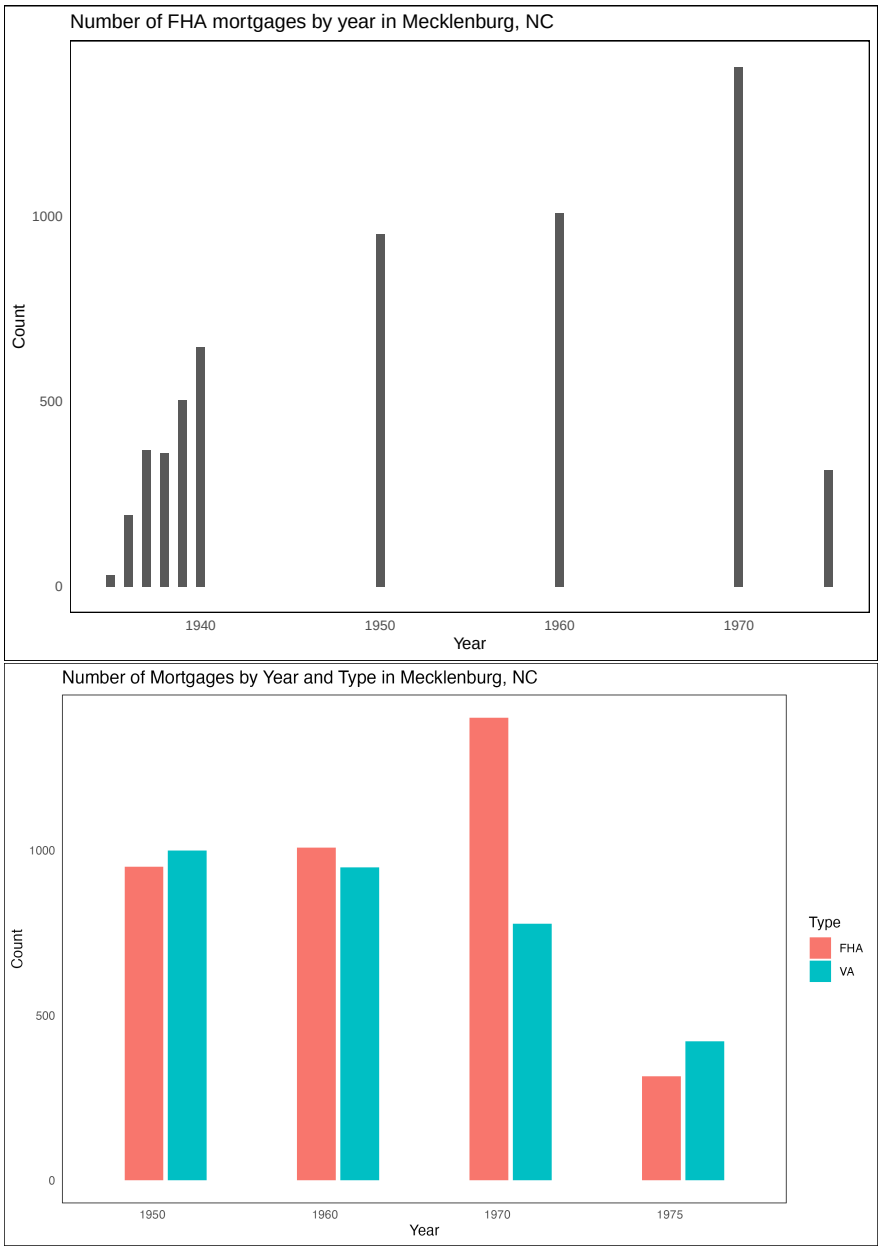
By collecting the near-universe of FHA-insured and VA-guaranteed mortgages, we can begin to describe some general trends in mortgage activity in the county. Figure 1 shows that FHA activity is rather modest in the mid-1930s and expands dramatically over the decade. It remains elevated in the immediate postwar period before declining by 1975. This pattern is consistent with historical accounts, which suggest that the FHA program underwent significant changes in the early 1970s that affected the volume of insured mortgages, as well as the underwriting criteria (Taylor, 2019). This is an area we plan to investigate further with the data collected for this project.

VA-guaranteed mortgages exhibit a different pattern. They peak in 1950 and gradually decline in number over time, likely due to the decline in the number of veterans eligible for the program. Nevertheless, in 1975 more VA-guaranteed mortgages were issued in Mecklenburg County than FHA-insured mortgages.

Next, we can examine the characteristics of FHA borrowers who were successfully matched with respondents to the 1940 census. In line with existing work, we find that Black borrowers make up a small share of those who received FHA-insured mortgages (1.5%) despite accounting for more than a quarter of all households in the county. Linking borrowers to the census also reveals that FHA mortgage recipients are younger, have higher occupational scores and incomes

[3] Fishback, Rose, et al. (2022) achieve match rates of 66-78% across the three cities they study.

Figure 1: Number of FHA and VA mortgages issued in Mecklenburg County, NC



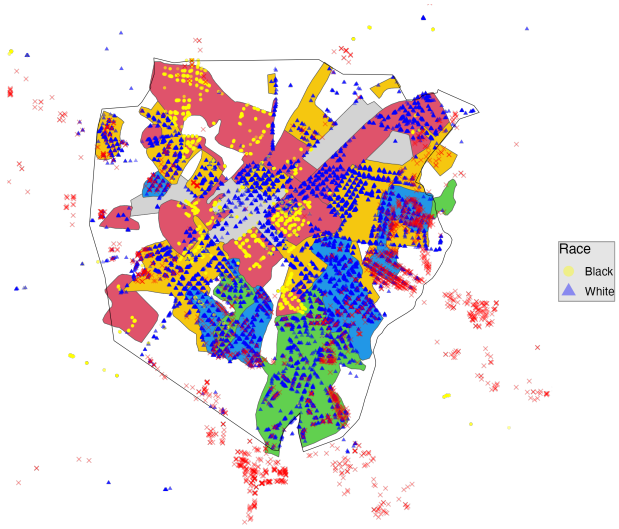
and live in homes with higher home values.

Turning to the geographic distribution of mortgages, we can examine where recipients lived within the city. To do so, we use the modern street address derived from the legal description to geo-code addresses of properties that received an FHA-insured mortgage. We begin with the HOLC map of Charlotte, which includes neighborhoods rated by their perceived residential security. Figure 5 reproduces the map along with its neighborhood ratings. To examine the racial

Table 1: Descriptive characteristics of linked sample

	FHA borrowers	County
Home value ($)	5,612	4,311
Income ($)	1,840	938
Occscore	31.32	22.13
Age	37.99	42.74
Black (%)	1.46	27.6
Home-owner (%)	79.91	33.1
N	1,503	36,636

Figure 2: HOLC map, race of households, and FHA mortgages (red crosses)



composition of neighborhoods where FHA-insured mortgages were issued, we also geo-coded the addresses of all census respondents in the city and included these on the same map as the FHA-insured mortgages (see Figure 2).

Figure 2 shows that the city of Charlotte was rigidly segregated in 1940. The colored polygons are neighborhoods identified by HOLC staff. The black outline is the municipal city boundary in 1940. Yellow circles depict individual Black households and blue triangles depict white households. All neighborhoods with Black residents were assigned the worst rating (D - red). This confirms the association between neighborhood racial composition and ratings assigned by the HOLC.[4] However, there is also a redlined neighborhood in the Northeast of the city that is com-

[4] Fishback, LaVoice, et al. (2020) however, argue that socio-economic conditions and other characteristics were also important for the ratings assigned to neighborhoods since the vast majority of redlined neighborhoods were composed entirely of white residents.

Table 2: Correlates of FHA-insured mortgages

	Dependent variable:	
	Total FHA mortgages	
	(1)	(2)
Occscore		1.723***
		(0.487)
Home value	0.002***	0.001*
	(0.0004)	(0.0005)
Rent	−0.313*	−0.412**
	(0.162)	(0.158)
Black share	−11.345***	−3.379
	(3.193)	(3.806)
Total pop	0.015***	0.012***
	(0.005)	(0.005)
Constant	0.497	−20.929***
	(3.265)	(6.825)
Observations	142	142
R^2	0.247	0.311
Adjusted R^2	0.225	0.285
Residual Std. Error	11.323 (df = 137)	10.876 (df = 136)
F Statistic	11.261*** (df = 4; 137)	12.263*** (df = 5; 136)
Note:		*p<0.1; **p<0.05; ***p<0.01

posed entirely of white households. Properties in this neighborhood do not appear to attract many FHA-insured mortgages. Green-lined and blue-lined neighborhoods, on the other hand, make up the bulk of locations with insured mortgages. There is also a significant number of insured mortgages outside the municipal boundaries of Charlotte, which were likely to have been new suburban developments.

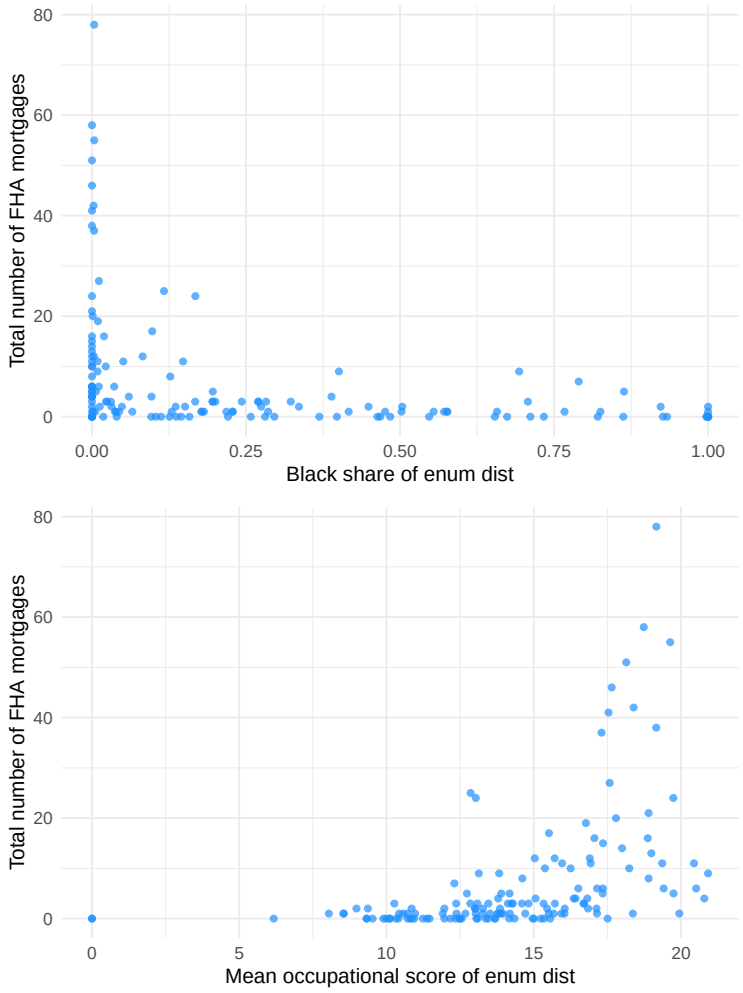
Leveraging the linked sample, we can examine the characteristics of enumeration districts that correlate with the incidence of FHA-insured mortgages. Typical enumeration districts contained roughly 400 households. In Figure 3, we show that very few enumeration districts with a significant Black share of residents received FHA-insured mortgages. Indeed, the variation in the incidence of mortgages is primarily within majority white districts. Among those neighborhoods, the mean occupational score is an important predictor of where insured mortgages are issued. This pattern is confirmed in a regression of the total number of FHA-insured mortgages against enumeration district characteristics in Table 2.

Summary

This project seeks to shed new light on the history of redlining by collecting borrower- and address-level data on insured mortgages in Charlotte, NC and linking these to the censuses of 1940 and 1950. Preliminary results confirm the view that Black households were unlikely to have received insured mortgages, and Black neighborhoods were disfavored in the distribution of FHA mortgages in the 1930s. We also find that the incidence of FHA-insured mortgages is correlated with occupational

score, an indicator of socio-economic advantage. Next steps for this project include linking borrowers to the 1950 census, examining how the distribution of mortgages changes (if at all) over the course of the twentieth century, and characterizing the spatial patterns of VA mortgages.

Figure 3: FHA-insured mortgages vs. Black share and Occupational score of enumeration districts



References

Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J Feigenbaum, and Santiago Pérez (May 2019). *Automated Linking of Historical Data*. Working Paper 25825. National Bureau of Economic Research (cit. on p. 5).

Ali, Omer, Nicholas Datto, Pei Yi Zhuo, Clinton Boyd Jr., and William A. Darity Jr. (2022). *Race and home values in Durham, North Carolina: 1940-2020* (cit. on p. 1).

Bayer, Patrick, Kerwin Kofi Charles, and JoonYup Park (2019). "Separate and Unequal: Race and the Geography of the American Housing Market". In: (cit. on p. 1).

Bureau, US Census (1950). *1950 Census of Housing: Volume 4. Residential Financing: Mortgaged Nonfarm Properties*. Census.gov. URL: https://www.census.gov/library/publications/1952/dec/housing-vol-04.html (visited on 11/22/2024) (cit. on p. 2).

Coates, Ta-Nehisi (May 2014). "The Case for Reparations". In: *The Atlantic* JUNE 2014 (cit. on p. 1).

Fishback, Price V., Jessica LaVoice, Allison Shertzer, and Randall Walsh (Nov. 2020). *The HOLC Maps: How Race and Poverty Influenced Real Estate Professionals' Evaluation of Lending Risk in the 1930s*. Working Paper 28146. National Bureau of Economic Research (cit. on p. 8).

Fishback, Price V., Jonathan Rose, Ken Snowden, and Thomas Storrs (May 2022). "New Evidence on Redlining by Federal Housing Programs in the 1930s". In: *Journal of Urban Economics*, p. 103462 (cit. on pp. 3, 6).

Freund, David M. P. (Aug. 2007). *Colored Property: State Policy and White Racial Politics in Suburban America*. University of Chicago Press (cit. on p. 2).

Gordon, Colin and Sarah K. Bruch (Sept. 2019). "Home inequity: race, wealth, and housing in St. Louis since 1940". In: *Housing Studies* 0.0, pp. 1–24 (cit. on p. 1).

Hynsjö, Disa M. and Luca Perdoni (June 2024). *Mapping Out Institutional Discrimination: The Economic Effects of Federal "Redlining"*. Tech. rep., p. 76 (cit. on p. 2).

Katznelson, Ira (Aug. 2006). *When Affirmative Action Was White: AN UNTOLD HISTORY OF RACIAL INEQUALITY IN TWENTIETH-CENTURY AMERICA*. W. W. Norton & Company (cit. on p. 3).

LaVoice, Jessica (Aug. 2024). "The long-run implications of slum clearance: A neighborhood analysis". In: *Journal of Public Economics* 236, p. 105153 (cit. on p. 2).

Michney, Todd M. (May 2021). "How the City Survey's Redlining Maps Were Made: A Closer Look at HOLC's Mortgagee Rehabilitation Division". In: *Journal of Planning History* Online FIrst, pp. 1–29 (cit. on p. 2).

Michney, Todd M. and LaDale Winling (Jan. 2019). "New Perspectives on New Deal Housing Policy: Explicating and Mapping HOLC Loans to African Americans". In: *Journal of Urban History*, p. 0096144218819429 (cit. on p. 2).

Rothstein, Richard (May 2017). *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liveright Publishing (cit. on p. 1).

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren (2024). *IPUMS Ancestry Full Count Data: Version 4.0* [*dataset*] (cit. on p. 5).

Taylor, Keeanga-Yamahtta (Sept. 2019). *Race for Profit: How Banks and the Real Estate Industry Undermined Black Homeownership*. UNC Press Books (cit. on p. 6).

Valenzuela-Casasempere, Pablo (2024). "Displacement and Infrastructure Provision: Evidence from the Interstate Highway System". In: (cit. on p. 2).

A Figures

Figure 4: Example FHA-insured mortgage

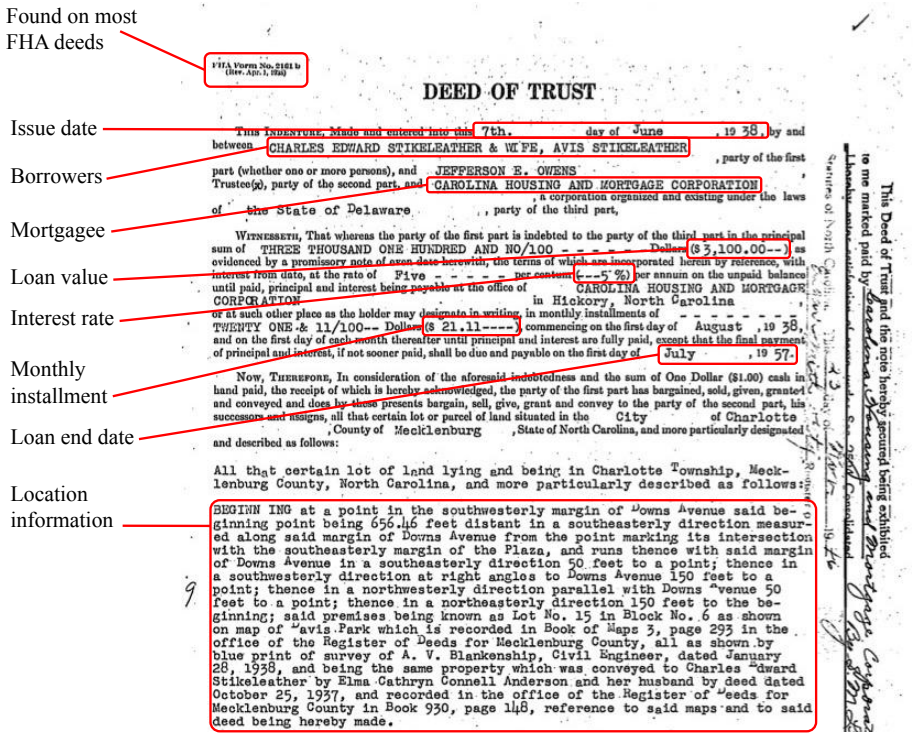


Figure 5: HOLC map of Charlotte, NC

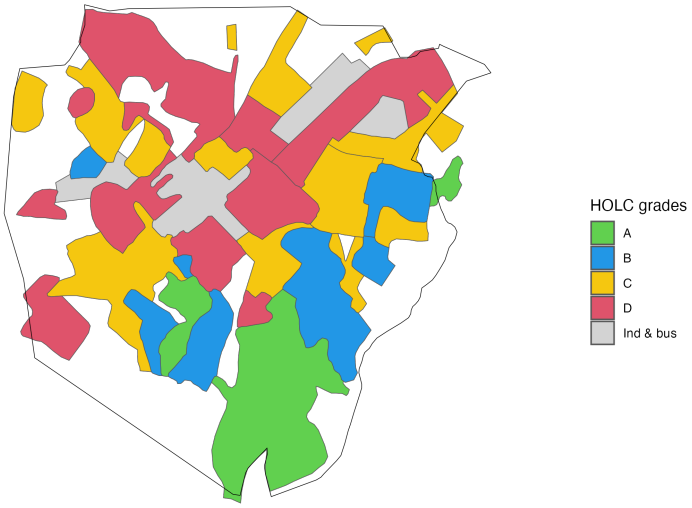


Figure 6: Location of FHA and VA mortgages in 1950 and 1960

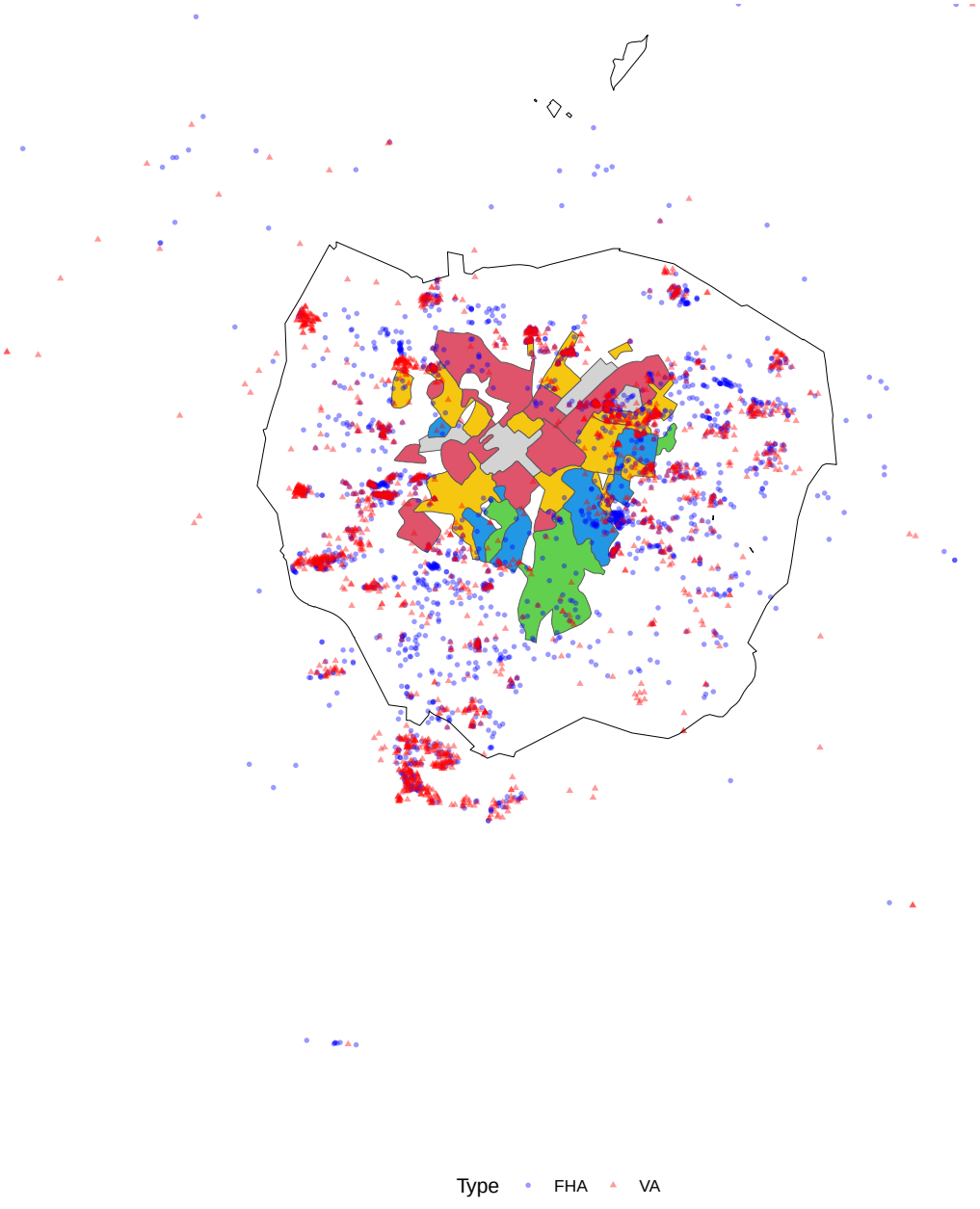


Figure 7: FHA mortgages in 1935-40 and 1970

