The Distribution of Federally-Insured Mortgages: 1935-1975, Evidence from Local Land Records*

EXTENDED ABSTRACT

Omer Ali Jacob Faber Andreas Ferrara Jonathan Rose Thomas Storrs
Kate Thomas Randall Walsh Wenfei Xu†

December 12, 2025

Over half a century after the Civil Rights Era, racial disparities in homeownership, wealth, and neighborhood opportunity persist. Scholars have long argued that these disparities were influenced by federal housing policies that favored white suburban residential developments at the expense of housing in central cities occupied by low-income minority residents (Rothstein, 2017). They identify mortgage programs introduced by the Federal Housing Administration (FHA) and Veterans Administration (VA) as central contributors, but systematic data on mortgage activity from this period is scarce. Using newly digitized real estate records obtained from county recorders of deeds and state archives, this paper characterizes the distribution of FHA-insured and VA-guaranteed mortgages issued between 1935 to 1975 across a range of counties, providing the first address- and borrower-level demographic and spatial characterization of these loans beyond 1940.[1]

During the New Deal and post-World War II period, the FHA and VA programs accounted for up to 50% of all mortgages in the United States (J. D. Rose and Snowden, 2013). They predominantly served white, suburban borrowers, fueling suburbanization and white flight and congealing the segregated order in urban areas. Research also suggests that later programs intended to expand lending to lower-income and minority applicants, such as the Section 235 program, sometimes resulted in predatory inclusion, com-

*Support from the Department of Economics at the University of Pittsburgh, Russel Sage Foundation grant G-2310-45068, and a grant from NYU's Institute of Human Development and Social Change is gratefully acknowledged. This work benefited from comments received from audiences at the Western Economic Association International annual conference, the Diversity Initiative for Tenure in Economics (DITE) workshop at Duke University, the Social Sciences History Association annual conference, and the Linking Historical Data Sources for Small Populations conference at NBER. We thank the following students for excellent research assistance: Mardoché Ayitchehou, Nick Datto, Kylie Fuller, Risa Gelles-Watnick, Yeji Lee, Fizzah Malik, Uday Sapra, and Gabrielle Toborg. We thank the following students for assistance with data collection: Saniya Abylkassova, Leyla Ahmadova, Max Barrett, Kasey Brown, Samantha Candalla, Nick Cao, Martin Chen, Roan Chiarilli, Megan Duda, Chazz Ehlis, Claire Gatz, Stephanie Ijomor, Adam Kaufman, Emilia Lam, Sally Margalitachvili, Katie Power, Karthik Sivachandran, Luke Smarra, Isabel Spotz, Simon Wang, Ilan Ziv, and students at the BYU Record Linking Lab. The views expressed do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

†Ali: University of Pittsburgh; Faber: New York University; Ferrara: University of Pittsburgh & NBER; Rose: Federal Reserve Bank of Chicago; Storrs: University of Virginia; Thomas: New York University; Walsh: University of Pittsburgh & NBER; Xu: Cornell University.

[1]To our knowledge, Fishback, J. Rose, et al., 2022 is the first paper to describe these patterns at the address and borrower level, with records from three cities for the period 1935-40.

pounding economic vulnerabilities (Taylor, 2019; Marchiel, 2021). The dearth of granular historical data on these mortgages limits our ability to evaluate their effects and draw lessons for today.

In this paper, we digitize local land records across a range of counties, using optical character recognition and machine learning methods to identify FHA-insured and VA-guaranteed mortgages. Once tabulated, addresses are geo-coded, and the names of borrowers matched to full-count Census data from 1940 and 1950, providing insights on borrowers' demographic and socioeconomic backgrounds. This approach overcomes hurdles and data limitations that have prevented research on this important topic from advancing in the past, significantly enhancing our ability to understand how federal housing policy shaped American metropolitan areas and entrenched racial inequality.

Related literature

Scholars have long connected present-day racial inequalities in housing, wealth, and neighborhood outcomes to the federal government's interventions in the housing market during the New Deal, especially the Home Owners Loan Corporation (HOLC) and its "redlining" maps.[2] HOLC's maps, which rated neighborhood lending risk in almost 240 cities, explicitly penalized urban areas with poor, immigrant, or minority residents (Rothstein, 2017; Fishback, LaVoice, et al., 2023; Hynsjö and Perdoni, 2024). Since these maps were digitized, researchers have linked them to a wide range of contemporary outcomes, including lending bias, crime, and even health disparities (Markley, 2024; Mallach, 2024). Despite this wealth of research, some scholars question whether HOLC's maps directly caused long-term inequality, given that they were primarily intended as internal documents and were compiled after HOLC had largely completed its refinancing operations (Michney, 2021; Fishback, LaVoice, et al., 2023). Moreover, HOLC refinanced loans in central city neighborhoods and did not systematically exclude Black borrowers (Michney and Winling, 2019).

By contrast, the FHA and VA programs are believed to have had larger, more persistent impacts (Jackson, 1985; Freund, 2007; Fishback, J. Rose, et al., 2022; Ali, 2022). Starting in 1934, the FHA insured mortgages to revive the housing market and stimulate employment, yet it did so in a manner that excluded most Black borrowers (Fishback, J. Rose, et al., 2022). Later, FHA's Section 235 program (of the 1968 HUD Act) attempted to expand access to low-income households but often resulted in "predatory inclusion," where units in disrepair were repeatedly sold to low-income buyers, fostering foreclosures and neighborhood disinvestment (Taylor, 2019).

Although many FHA records were lost, scholars have gleaned insights from government correspondence, legislative documents, and contemporaneous critiques (Jackson, 1985; Freund, 2007). A notable study examining 4,334 FHA loans in three U.S. cities between 1935 and 1940 found that fewer than 1% of these mortgages were issued to African American borrowers—far more discriminatory than HOLC (Fishback, J. Rose, et al., 2022). However, no comparable analyses exist for the VA. The GI Bill's VA mortgage guarantee program aided nearly nine million veterans by 1975, but early evidence indicates it discriminated against nonwhite veterans (Veteran Affairs, 1975; Woods II, 2013; Agbai, 2022).

Aggregated Census data corroborate these patterns. In 1950, white homeowners received significantly more FHA and VA mortgages than nonwhites, and although these gaps narrowed by 1960, persistent racial

[2]See, for example, Hirsch (1983), Jackson (1985), Massey and Denton (1993), Sugrue (1996), Freund (2007), Connolly (2014), Appel and Nickerson (2016), Rothstein (2017), Faber (2020), Aaronson, Hartley, and Mazumder (2021), and Hynsjö and Perdoni (2024).

disparities remained (Census, 1952; Census, 1963). Still, the absence of granular mortgage- and borrower-level data has prevented scholars from fully investigating where, how, and for whom these loans were made, limiting our understanding of the FHA's and VA's roles in perpetuating segregation and widening wealth gaps.

Data construction

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, most of which remain in paper form. The few that have digitized their archives may not allow access that facilitates the collection of large numbers of documents.[3]

Building on the pioneering work of Fishback, J. Rose, et al. (2022) that digitized borrower-level mortgage data for three counties to study FHA insurance from 1935 to 1940, this paper expands the scope to additional counties and extends the time period beyond 1940 to examine both FHA and VA mortgages. In selected counties, we aim to gather the universe of FHA-insured mortgages from 1935-40 and FHA-insured and VA-guaranteed mortgages in 1950, 1960, 1965, 1970, and 1975. Our selected years enable repeated observations of federally-underwritten mortgage activity up to the passage in 1975 of the Home Mortgage Disclosure Act (HMDA) with more intense focus on the New Deal and Civil Rights Eras.

Counties were selected for (1) availability of digitized mortgage records, (2) geographic representativeness, and (3) varied trajectories over the twentieth century. While important urban centers like New York, Los Angeles, and Detroit remain out of reach due to data inaccessibility or expense, we have obtained access to mortgage records from Maricopa, AZ (Phoenix), Peoria, IL, Marion, IN (Indianapolis), Hampden, MA (Springfield), Baltimore City and County, MD, Hillsborough, NH (Manchester), Durham, NC, Guilford, NC (Greensboro), Mecklenburg, NC (Charlotte), Summit, OH (Akron), Tulsa, OK, Allegheny, PA (Pittsburgh), Berks, PA (Reading), Philadelphia, PA, Bexar, TX (San Antonio), and Galveston, TX.[4]

Once records are obtained, we combine optical character recognition (OCR) technology, large language models, and human validation to extract and tabulate key information efficiently, while minimizing error. For the records in each county, we follow these five steps:

1. OCR Processing and Identification: Historical mortgages stored as image files are digitized and processed using OCR, typically with a paid service such as AWS Textract, or an open-source tool like Tesseract. This step generates text data from scanned images, preserving the structure of the original documents where possible. At this stage, we identify FHA-insured mortgages and, from 1950 onward, VA-guaranteed mortgages by searching for certain key terms. These mortgages were always printed on standardized forms (see Figure 2 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.

2. Text Extraction and Conversion to Tabular Data: The extracted OCR text is next sent to OpenAI's GPT API. A custom prompt guides the model to identify and structure specific information, including the names of parties, property descriptions, and transaction details. The prompt takes advantage of the regular

[3]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.

[4]We sometimes refer to counties by their principal city. Also, at the time of writing, we are actively pursuing records from a number of additional counties.

format of mortgages, prompting the model to search for key phrases within the text. The model outputs the parsed data in a tabular format.

3. Human Review and Final Validation: The tabular data is linked to the original mortgage images. Research assistants conduct a review of the tabular data against the original image to ensure accuracy. The full set of columns in the tabular data are: Type of mortgage (FHA/VA); Date of issue; Book number; Page number; Borrower last name; Borrower first name; Second borrower first name; Lender bank; Amount borrowed; Interest rate; Monthly rate; Start date; Due/end date; Legal description; Modern address; and Lat/Long.

4. Geolocating Properties: We can geolocate properties based on the legal description included in the mortgage. Some counties - Baltimore, for example - include the address within the property description. In other counties, such as Tulsa, OK or Charlotte, NC, nearly all properties are uniquely identified by a subdivision name or corresponding plat map, as well as a lot and block number. These can often be linked directly with Regrid parcel data that includes the modern street address, as well as latitude and longitude. Some counties are not as straightforward. These may contain a mixture of addresses, subdivisions, and metes and bounds property descriptions or might not have available parcel data for geolocating. In those cases, properties can still usually be geolocated manually, by cross-referencing current and historical maps.

5. Linking borrowers to the Census: To match borrowers to the Census, we use a combination of automated techniques and manual verification to ensure reliable matches and minimize false positives. The process begins by cleaning and standardizing names in both datasets using the Abramitzky, Boustan, and Eriksson (ABE) method (Abramitzky et al., 2019), which harmonizes names to account for spelling variations and transcription errors. Street names are standardized to reduce formatting and abbreviation discrepancies, ensuring better alignment between datasets.

Once data cleaning is complete, records are matched in a series of increasingly flexible steps. Initial matches are based on exact agreements between first name, last name, and street name. Subsequent steps incorporate additional information, such as spouse names and middle names, with manual checks to confirm accuracy. For cases requiring greater flexibility, we employ the Jaro-Winkler string similarity measure to identify potential matches with minor discrepancies. At each stage, manual review ensures the removal of false positives and evaluates alternative matches for accuracy. This combination of automation and verification ensures linkages between borrowers and Census records, providing a reliable foundation for subsequent analyses.

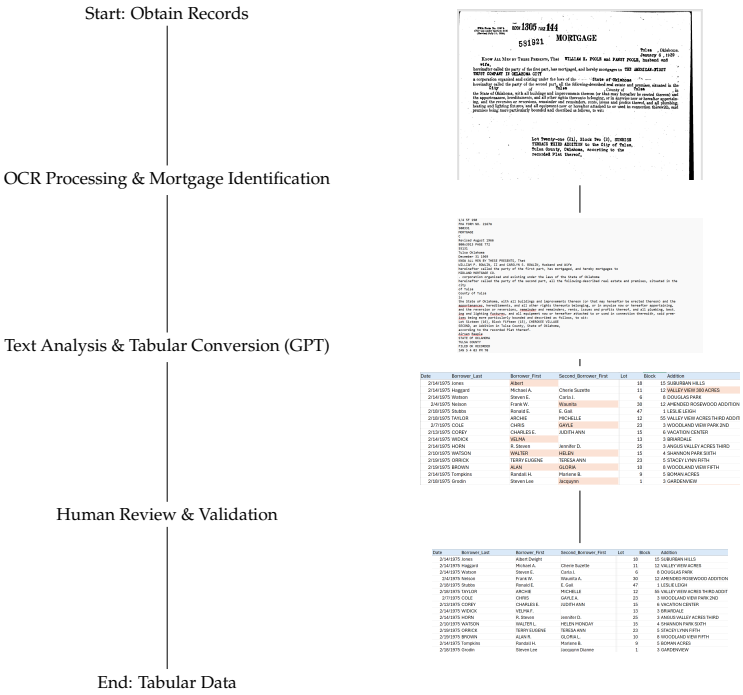
In Mecklenburg county, after removing false positives, we achieve a matching rate of 63.8%.[5] Accurate matches are particularly important in our case since one of our main research questions concerns the extent to which Black home-owners are under-represented among FHA borrowers. In Mecklenburg county, a total of 1,503 borrowers are linked, only 22 of whom are Black.

Preliminary Results:

By collecting the near-universe of FHA-insured and VA-guaranteed mortgages, we can describe some general trends in mortgage activity. Mecklenburg County in North Carolina, which contains the state's largest

[5] Fishback, J. Rose, et al. (2022) achieve match rates with the addition of manual matching of 66-78% across the three cities they study.

Figure 1: Data extraction pipeline



city, Charlotte, is one of the few counties that allows free access to their historic real estate records 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. 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.

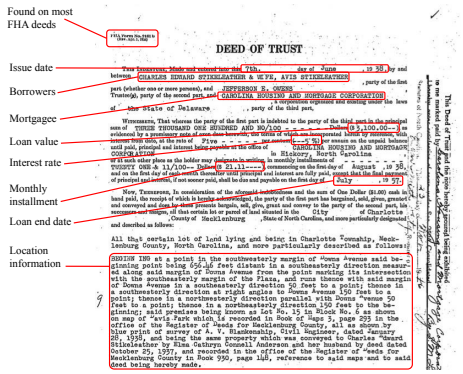
Next, we can examine the characteristics of FHA borrowers who were successfully matched with respondents to the 1940 Census. In line with existing work, we find that Black borrowers make up a small share of those who received FHA-insured mortgages (1.5%) despite accounting for more than a quarter of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

Turning to the geographic distribution of mortgages, we can examine where recipients lived within the city. We begin with the HOLC map of Charlotte, which includes neighborhoods rated by their perceived residential security. To examine the racial composition of neighborhoods where FHA-insured mortgages

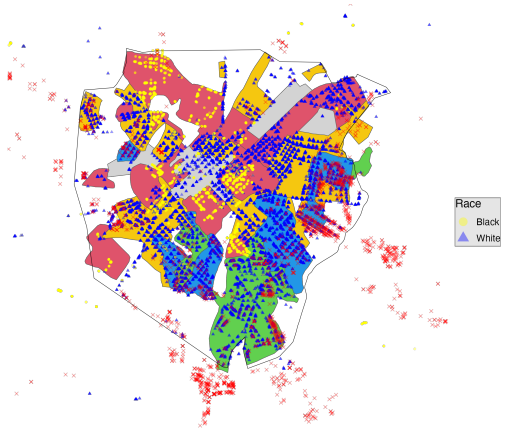
were issued, we also geo-coded the addresses of all 1940 Census respondents in the city and included these on the same map as well as the FHA-insured mortgages (see Figure 2).

Figure 2

(a) FHA-insured mortgage



(b) HOLC map and FHA mortgages



Note: Panel (a) shows a scanned image of an FHA-insured mortgage (called "deed of trust" in some states) recorded in Mecklenburg County, annotated with the location of relevant information extracted in the automated data entry stage. Panel (b) shows the HOLC map of Charlotte, NC, city boundary (black line), location of households by race in 1940 (yellow circles and blue triangles), and location of FHA mortgages issued in 1935-40 (red crosses). The polygons correspond to neighborhoods rated in the HOLC map for Charlotte. Color-coded maps rated neighborhoods as best (green), still desirable (blue), definitely declining (yellow) and hazardous (red).

Figure 2 shows that the city of Charlotte was rigidly segregated in 1940. All neighborhoods with Black residents were assigned the worst rating (D - hazardous). This confirms the association between neighborhood racial composition and ratings assigned by the HOLC.[6] However, there is also a redlined neighborhood in the Northeast of the city that is composed 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 Mecklenburg County, very few enumeration districts with a significant Black share of residents received any 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.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder (2021). "The Effects of the 1930s HOLC 'Redlining' Maps". In: *American Economic Journal: Economic Policy* 13.4, pp. 355–92 (cit. on p. 2).

[6] Fishback, LaVoice, et al. (2023) however, argue that socio-economic conditions and other characteristics were also important for the ratings assigned to neighborhoods and the vast majority of homeowners in redlined neighborhoods in their sample were white.

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. 4).

Agbai, Chinyere O. (Sept. 2022). *Wealth Begins at Home: The Housing Benefits of the 1944 GI Bill and the Reproduction of the Black-White Wealth Gap in Homeownership and Home Value, 1940-1960* (cit. on p. 2).

Ali, Omer (May 2022). "On the Impact of Federal Housing Policies on Racial Inequality". In: *AEA Papers and Proceedings* 112, pp. 229–233 (cit. on p. 2).

Appel, Ian and Jordan Nickerson (2016). "Pockets of poverty: The long-term effects of redlining". In: *Available at SSRN 2852856* (cit. on p. 2).

Census, Bureau of the (1952). *1950 Census of Housing Volume IV. Residential Financing: Mortgaged Nonfarm Properties*. Tech. rep. Washington, D.C.: United States Department of Commerce (cit. on p. 3).

— (1963). *1960 Census of Housing Volume V. Residential Finance: Homeowner Properties*. Tech. rep. Washington, D.C.: United States Department of Commerce (cit. on p. 3).

Connolly, Nathan D.B. (2014). *A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida*. Chicago, IL: University of Chicago Press (cit. on p. 2).

Faber, Jacob William (2020). "We Built This: Consequences of New Deal Era Intervention in America's Racial Geography". In: *American Sociological Review* 85.5, pp. 739–75 (cit. on p. 2).

Fishback, Price V., Jessica LaVoice, Allison Shertzer, and Randall P. Walsh (Dec. 2023). "The HOLC Maps: How Race and Poverty Influenced Real Estate Professionals' Evaluation of Lending Risk in the 1930s". In: *The Journal of Economic History* 83.4, pp. 1019–1056 (cit. on pp. 2, 6).

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* (cit. on pp. 1–4).

Freund, David M.P. (2007). *Colored Property*. Chicago, IL: University of Chicago Press (cit. on p. 2).

Hirsch, Arnold R. (1983). *Making the Second Ghetto: Race and Housing in Chicago 1940-1960*. Chicago, IL: University of Chicago Press (cit. on p. 2).

Hynsjö, Disa M. and Luca Perdoni (2024). *The Effects of Federal "Redlining" Maps: A Novel Estimation Strategy*. Tech. rep., p. 76 (cit. on p. 2).

Jackson, Kenneth T. (1985). *Crabgrass Frontier: The Suburbanization of the United States*. Oxford, United Kingdom: Oxford University Press (cit. on p. 2).

Mallach, Alan (2024). "Shifting the Redlining Paradigm: The Home Owners' Loan Corporation Maps and the Construction of Urban Racial Inequality". In: *Housing Policy Debate* 0.0, pp. 1–27 (cit. on p. 2).

Marchiel, Rebecca K. (Sept. 2021). *After Redlining: The Urban Reinvestment Movement in the Era of Financial Deregulation*. Historical Studies of Urban America. Chicago, IL: University of Chicago Press (cit. on p. 2).

Markley, Scott (Feb. 2024). "Federal 'redlining' maps: A critical reappraisal". In: *Urban Studies* 61.2, pp. 195–213 (cit. on p. 2).

Massey, Douglas S. and Nancy A. Denton (1993). *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press (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).

Rose, Jonathan D. and Kenneth A. Snowden (2013). "The New Deal and the origins of the modern American real estate loan contract". In: *Explorations in Economic History* 50.4, pp. 548–566 (cit. on p. 1).

Rothstein, Richard (May 2017). *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liveright Publishing (cit. on pp. 1, 2).

Sugrue, Thomas J. (1996). *The Origins of the Urban Crisis: Race and Inequality in Postwar Detroit: Race and Inequality in Postwar Detroit*. Princeton, NJ: Princeton University Press (cit. on p. 2).

Taylor, Keeanga-Yamahtta (2019). *Race for Profit: How Banks and the Real Estate Industry Undermined Black Homeownership*. Chapel Hill, NC: UNC Press Books (cit. on pp. 2, 5).

Veteran Affairs, Administrator of (1975). *1975 Annual Report*. Tech. rep. Washington, D.C. (cit. on p. 2).
Woods II, Louis Lee (2013). "Almost "no negro veteran … could get a loan": African Americans, the GI Bill, and the NAACP Campaign Againts Residential Segregation, 1917–1960". In: *The Journal of African American History* 98.3, pp. 392–417 (cit. on p. 2).