Check for updates

Access to housing credit and marriage: evidence from redlining maps

Margarida Madaleno[1]

Abstract

This study evaluates the impact of a policy popularly known as "redlining" on marriage. This policy led to the creation of a series of maps that guided banks in their lending, where some areas were favored and others were discriminated against. Given the quasi-randomness of mortgage discrimination, this policy allows us to make inferences regarding whether housing credit constraints affect marriage. Furthermore, it also provides insight into whether unequal access to housing credit played a role in the contemporary racial marriage gap. This policy allows us to make these inferences due to the fact that neighborhood blocks that were more heavily discriminated against had higher proportions of Black residents. The study uses a spatial differences in discontinuities design to show that the maps led to a reduction in marriage in discriminated areas. These effects are shown to not be due to sorting. They can also be ascribed to the housing credit mechanism per se, rather than competing second-order mechanisms that result from individuals being denied mortgages. These second-order effects can include, for instance, neighborhood decline effects found in other redlining work.

Keywords Marriage · Housing credit · Redlining

1 Introduction

Broadly, this study contributes to understanding how home ownership affects marriage decisions. There is little evidence on whether obtaining housing credit, and purchasing a home, is an important determinant of marriage. As noted by Eriksen (2010), the classic Gary Becker framework for marriage decisions (Becker 1973) is ambiguous about the effects of housing wealth on marriage. Two opposing channels may lead to these ambiguous effects. On one hand, there are so-called

✉ Margarida Madaleno
  m.madaleno@lse.ac.uk

[1] Instituto Superior de Economia e Gestao (ISEG), University of Lisbon, Lisbon, Portugal

✷ Springer

"independence" effects whereby home-ownership renders single households less reliant on a spouse for support in the production of household goods. On the other hand, economies of scale in household production and housing consumption may lead to an a positive impact of home-ownership on marriage - i.e. owning a home (rather than renting) makes the household more productive, and therefore, marriage more attractive. Furthermore, it is also possible that, in the marriage market, home-ownership (and the wealth that this implies) makes individuals more "marriageable". In this vein, Lafortune & Low (2017) find that individuals with more assets are more likely to marry.

Empirical evidence on the impact of housing credit on house prices is extensive. For instance, Carozzi et al. (2024) consider the impact of a UK policy, "Help to Buy", on house prices. Authors find that this equity loan scheme is capitalized into house prices due to structural inelasticities in housing supply. Furthermore, research on the impact of home ownership on marriage seems to support the idea that home-ownership is conducive to marriage. This relationship holds in China (Hu & Wang 2019 and the US Eriksen 2010). In particular, Eriksen (2010) find that home-ownership subsidies increase the likelihood of marriage. Similarly, Ricks (2021) find that the VA Loan Program, which subsidized housing credit for veterans in the 1950s, increased marriage. Furthermore, Hu & Wang (2019) note that in China, this causal relationship is driven by the greater attractiveness of home-owners in the marriage market. Finally, Miller & Park (2018) consider whether marriage leads to more home-ownership, exploiting quasi-random legalization of same-sex marriage. They find that marriage does indeed lead to higher rates of home-ownership.

Furthermore, this study also provides insight into whether unequal home ownership contributed to the racial marriage gap. Black Americans are less likely to marry than their white counterparts - this is well-documented in the literature (see for instance Raley et al. 2015). This fact, however, is often ascribed to income or incarceration levels (see for instance Caucutt et al. 2016). In contrast, this work ascribes at least part of the racial marriage gap to historical discrimination in home ownership. Specifically, it finds that an area-based mortgage discrimination policy popularly known as "redlining" but henceforth referred to as "HOLC maps" (Home Owners Loan Corporation), which targeted predominantly Black neighborhoods blocks, led to a causal decline in marriage. To this end, the first part of the analysis documents racial gaps in marriage and home ownership. It also finds that the racial marriage gap narrows (although not fully) once home ownership is accounted for. The causal effects of the maps on marriage provide evidence in favor of the conclusion that this policy - which specifically targeted Black neighborhood blocks - at least partially explains the racial marriage gap.

As noted, this study exploits quasi-exogenous variation in housing credit terms afforded by the HOLC maps. HOLC, a federal Roosevelt-era agency, created color-coded maps for over 239 major US cities, which described the probable economic trajectory of each neighborhood block. In practice, discriminated neighborhood blocks had a higher proportion of Black residents, as this population was considered a "disamenity" likely to drive down house prices in the future. These maps were then covertly circulated to banks, and ultimately guided lending. As prior redlining impact evaluations have shown, these maps changed the course of neighborhood block trajectories (Krimmel 2018, Aaronson et al. 2017).

In terms of the methodology used to uncover the reduced form effects, which allow us to make inferences about housing credit and marriages, the policy is superficially amenable to a spatial regression discontinuity design (RDD) set-up. This is due to the fact that the maps represent sharp yet contiguous discontinuities in grades. However, it will be shown that these grades were not randomly assigned, and were in fact based on careful data collection. More crucially, in addition to being based on careful collection, the grade designations followed pre-existing discontinuities in the variables featured in HOLC's data surveys. This means that the basic RDD requisite of variable smoothness across the boundary does not hold.

To circumvent this, a spatial regression discontinuity difference in differences (RDD-DID) approach is deployed. In particular, this entails performing a spatial RDD, with border fixed effects and a running variable polynomial, whilst interacting the treatment variable with a before-and-after time variable. This method allows this study to overcome potential endogeneity issues inherent to a simple RDD.

The study finds that the policy decreased marriage in discriminated areas. However, several mechanisms could be driving these results, and the main regressions do not isolate the housing credit mechanism. Note, henceforth the term "housing credit mechanism" refers the the first-order impact of individuals being denied housing credit, in contrast to second-order mechanisms that are a result of the first-order impact of individuals being denied housing credit. An example of this would be general area-level decline, which is a second-order effect of e.g. local businesses being unable to secure loans. For instance, it is possible that the HOLC policies led to area-level decline, which in turn decreased marriage. In this case, the area-level effects previously outlined would not be due to the housing credit mechanism. In order to understand whether alternative mechanisms explain the main results, this study presents the main regressions with and without a variable that controls for area-level decline. In this case, area-level decline is summarized by a variable that captures house prices. It is held that if the results remain statistically significant in spite of the inclusion of the house price variable, then the results are explained by the housing credit mechanism per se, rather than area-level decline. Indeed, the results are robust to the inclusion of this variable, which suggests that the main results are driven by housing credit per se.

In order to understand whether the results simply reflect sorting, a city-level analysis is performed. It is possible that the housing credit discrimination simply caused never-married individuals to sort into discriminated areas. However, the city-level results show that cities that received the maps experience more marriage decline than cities that did not receive the maps. In sum, the results are not a feature of sorting, and instead suggest that poor access to housing credit indeed decreased marriage.

2 Racial marriage and home ownership gaps

This section will motivate the rest of the study by showing that it is likely that racial gaps in home ownership are partially responsible for gaps in marriage. Graphs here will provide suggestive evidence that will justify the ensuing analysis, which establishes a causal relationship between discriminatory policies that affected Black
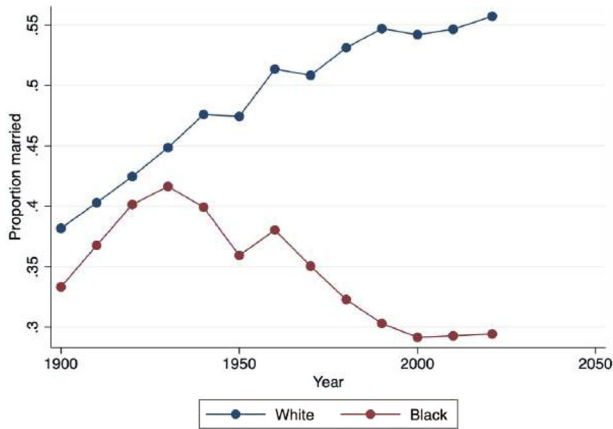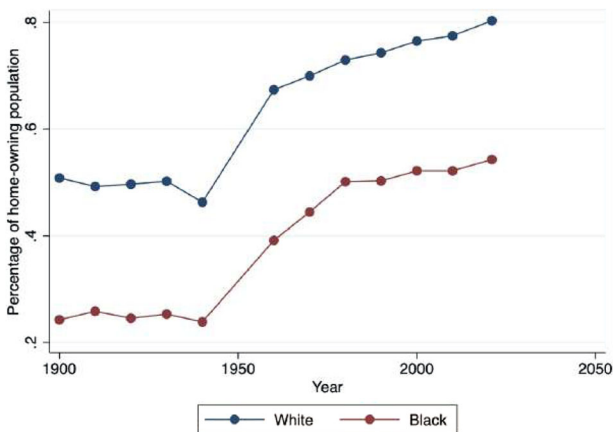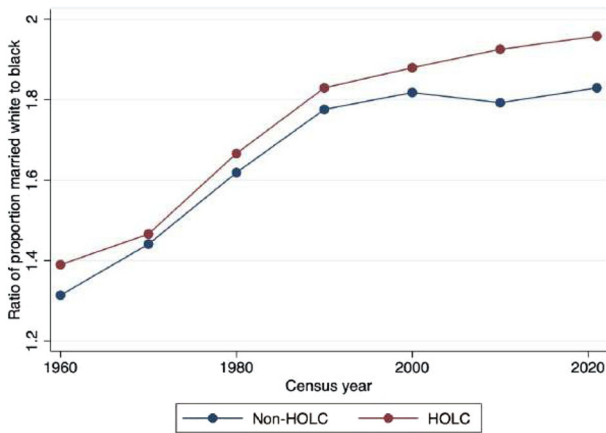
Fig. 1 Racial marriage gap

Fig. 2 Racial home ownership gap

home ownership, and consequently, marriage. Data used in this analysis was obtained via the IPUMS nation-wide individual census.

Figure 1 features the racial marriage gap. White and Black marriage gaps were relatively small up to 1930, at which point the gap widens. In 2020, the proportion of married white individuals is over 55%, while the proportion of married Black individuals is around 30%.

Figure 2 shows the racial home ownership gap. Consistently, white home ownership is around 30 percentage points higher than Black home ownership. Unlike with the patterns shown in Fig. 1, the gap does not diverge in 1930. The fact that that gaps here are more or less consistently parallel, and do not mimic the marriage gaps, suggests that home ownership - as will be shown - might be an important determinant, but does not, alone account for the divergence in marriage rates.

Furthermore, home ownership may be collinear with other variables that determine marriage - such as income. If this is the case, it may be irrelevant to causally

Fig. 3 Racial marriage gap across HOLC and non-HOLC cities

determine the impact of mortgage discrimination on marriage. To understand whether this is the case, Fig. 3 plots the ratio of white to Black marriage for cities that received the HOLC maps and cities that did not receive the HOLC maps. Although the gap is narrow, the ratio is higher for cities that received the maps than cities that did not. These figures therefore justify a closer look at the causal impact of the maps on marriage.

3 Policy overview

Narrowly, this study is interested in understanding whether access to housing credit influences marriage decisions. Broadly, it is interested in understanding whether the racially discriminatory redlining policy contributed to the racial marriage gap. Understanding the causal impact of credit access on marriage is difficult because housing credit access is endogenous to family outcomes. For instance, poverty may prevent someone from obtaining housing credit, while simultaneously make them less attractive marriage partners. This section will describe the redlining maps, and how they are an exogenous source of housing credit discrimination. It will also describe the fact that the maps targeted predominantly Black neighborhoods, thereby providing insight into whether the maps contributed to the racial marriage gap. This section outlines some policy detail.

As noted by Hillier (2005), the HOLC was created by the federal government to slow the inevitable home foreclosures caused by the Great Depression. This institution was initially conceived in order to provide more favorable housing credit and loans to struggling families, and it did so between 1933 and 1936. In 1935, the Federal Home Loan Bank Board (FHLBB) used HOLC staff to conduct surveys of the desirability of neighborhood blocks within 239 cities. The objective of this work was to understand the trajectories of different neighborhood blocks in order to better understand the viability of the aforementioned housing credit payments. Throughout this City Survey Program, HOLC surveyors collected data and conducted qualitative
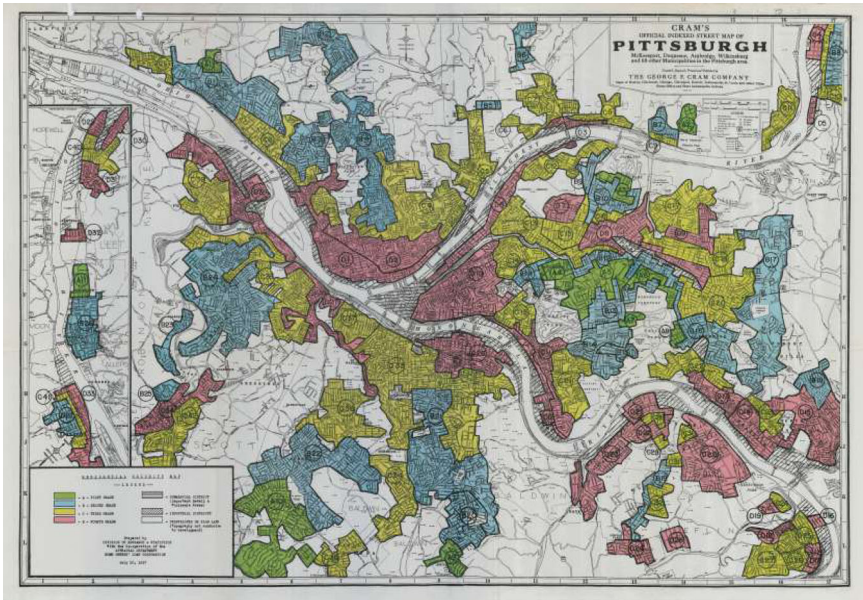
Fig. 4 HOLC map for Pittsburgh, Pennsylvania

analyses for each neighborhood block. This data collection culminated in the drafting of maps for chosen cities. In particular, the maps outlined which parts of the cities were to be lent to and which parts of the city were considered dangerous to loan to. In practice, each block was assigned a grade from "A" (most desirable) to "D" (least desirable), and the maps are color-coded to reflect this. Green designated "A" neighborhood blocks, blue designated "B" neighborhood blocks, yellow designated "C" neighborhood blocks, and finally red designated "D" neighborhood blocks. Although the maps were arguably made for internal consumption by HOLC, the maps were then covertly (in the sense that the general public did not know about them until much later on) circulated to lending institutions (Hillier 2003). Given that these maps were designed to predict the neighborhood block trajectories, it is not obvious that differences between neighborhood blocks after the treatment period are due to the maps per se. However, recent causal evidence shows that the maps did indeed catalyze area-level decline in D-graded areas (Krimmel 2018, Aaronson et al. 2017). Furthermore, in an analysis of housing credit disbursal in Philadelphia, Hillier (2003) shows that it was only in D-graded areas that the maps were binding, in the sense that lenders offered more stringent housing credit terms than they would have otherwise. Here, the author shows that the maps did not just follow pre-existing discrimination, but further compounded it. Figure 4 shows the HOLC map for Pittsburgh.

The varying HOLC grades represent differing treatment intensities. Recall that these maps were covertly circulated to banks, and it was on the basis of this legend that banks were guided regarding how to approach different neighborhood blocks when evaluating the loan-to-value ratio (LTV). In A-graded neighborhood blocks, banks are told to lend liberally as these are considered "hot spots". A LTV of 75-80

Fig. 5 Block-level survey for Pittsburgh, Pennsylvania

percent is suggested. Next, "B" areas are described as "still good" but not "hot spots", and a 65 percent LTV is proposed. On the other hand, "C" areas are described as being "infiltrated by 'lower grade' population", and lenders are told to be "conservative". Finally, D-graded areas are describes as "hazardous" and lenders are told to refuse loans altogether.

The grades, however, were not randomly assigned. In fact, the masses of color in Fig. 4 represent contiguous sets of neighborhood blocks that were meticulously analysed by HOLC. For each block, surveyors would fill out a form that would depict its characteristics. Figure 5 presents one of these block-level surveys for Pittsburgh, Pennsylvania. This survey was collected in 1937, and the image was provided by the Mapping Inequality project. From this image, it is clear that HOLC was looking for specific things, and they were rather uniform in their data collection. The insights collected by the surveyors ranged from qualitative to quantitative. For instance, with respect to the nationality of foreign-born individuals, this particular block is described as a "mixture" where 40% of individuals are non-native. Average family income at the block level is deemed between $1200 and $2500. There is a binary indicator for whether there are Black Americans in the block (here a "yes") as well as a percentage for the proportion of the population that is Black (10% in this case).

Crucially, Black residents were considered a "disamenity", which led to neighborhoods that housed a large number of Black Americans to receive worse grades.

Table 1 Share of black population for each grade pre-treatment

| | Proportion black | | | |
	A	B	C	D
Mean	0.012	0.012	0.021	0.15
SD	(0.02)	(0.004)	(0.09)	(0.29)
Obs	117	366	1115	1266

Note: This table features descriptive statistics regarding share of black residents for each grade, for the pre-treatment year of 1930. Each column represents a grade, and the first row is the mean proportion of black residents. This is calculated as the total number of black residents over the total number of residents. The second row is the standard deviation, and the final row is the number of observations
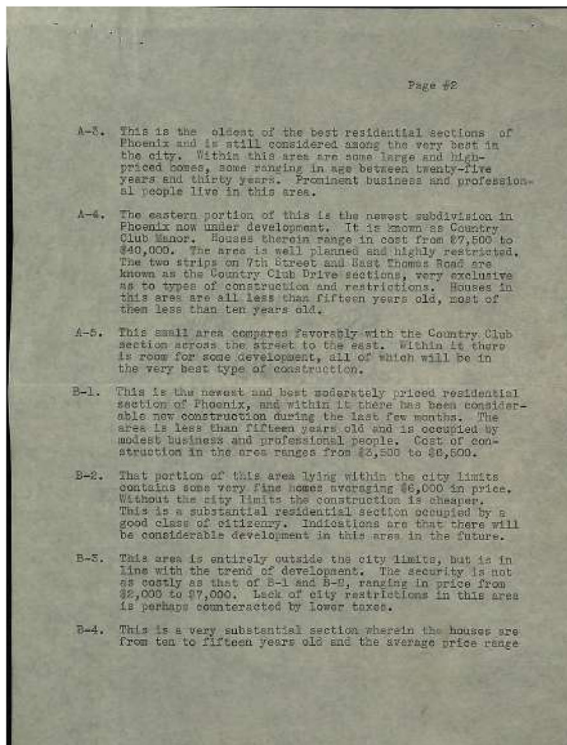
Fig. 6 Block-level survey for Phoenix, Arizona

Table 1 shows the proportion of Black residents by grade, for the pre-treatment year of 1930. As the last column of the table shows, D-graded areas were on average 15 percent Black, compared to A areas that were 1.2% Black.

As an accessory research question, it is worth understanding whether at what level this discrimination occurred. In particular, there were several iterations of the maps. Earlier versions were based on qualitative assessments by HOLC surveyors. Figure 6

Table 2 Impact of share of black population on D designation across map versions

	D designation	
	(1) Old	(2) New
Proportion black	4.458	0.710[*]
	(0.80)	(4.22)
Observations	19	159

Note: This table features the OLS correlation between proportion black residents of an observation and the likelihood of D designation. Column 1 restricts the sample to cities that received an early version of the maps, while column 2 restricts the sample to cities that received later versions of the maps. Standard errors are clustered at the city level

t statistics in parentheses

[+]p < 0.10, [*]p < 0.05

presents such a survey for Phoenix, Arizona. Conversely, later versions were based on quantitative assessments, as was shown in Fig. 3.

It is held that the earlier, qualitatively-based, versions of these maps reflect the personal biases of the surveyors more accurately than the later, quantitatively-based versions of the maps. The latter are more likely to reflect institutional biases. Table 2 features the correlation between the proportion of Black residents in an area and D-grade designation. Column 1 looks at effects for old (qualitative) maps, while column 2 looks at effects for new (quantitative) maps. While the latter is significant and the former is not, it is worth noting that the number of cities that only received earlier versions is much lower. Comparing the estimates yields a *p*-value of 0.345, suggesting that the coefficients are not significantly different. However, the sample size, particularly for the qualitative maps, is small and does not allow us to make precise inferences. The relatively magnitudes of these coefficients suggest that earlier versions were more prone to bias. Still, these are just suggestive estimates, as it is also possible that there is endogeneity related to which cities received which versions.

This study considers years 1950, 1960, and 1970 as post-treatment years. The final post-treatment year is 1970, largely because the Fair Housing Act of 1968 effectively rendered redlining illegal. However, as noted by Massey (2015), while the Fair Housing Act curtailed discrimination, it did not fully end it. In practice, it was very difficult to legally contest housing discrimination, and those found guilty of it faced few repercussions. Still, it would be expected that the policy's impact reaches a maximum in 1970 and declines thereafter.

4 Data

This study exploits the Mapping Inequality project's digitization of the HOLC maps. In particular, this data includes block-level information on HOLC grades for 202 cities across the USA. Note that only 202 of 239 maps were digitized, due to unavailability of maps at the National Archives.

Fig. 7 Schematic representation of observations. Panel **A** represents HOLC neighbourhood blocks, where green and blue represent different grades. Panel **B** represents the dissolution of the blocks in (**A**), to yield a contiguous set of same-graded polygons where green and blue again represent different grades. Panel **C** represents the intersection of the observations in Panel **B** with census tract polygons

However, the outcome of interest, provided by Integrated Public Use Microdata Series' National Historical GIS (Manson et al. 2019), is at the census tract level. In sum, the independent variable of interest (i.e. grade) was attributed at a finer spatial scale than the outcome (i.e. marriage). In order to bridge the grades at block-level with marriage at census tract level, the following steps were undertaken. Recall that the HOLC data was collected at the neighborhood block level. Figure 7 presents a schematic diagram of the units of observations. In first instance, we have neighborhood blocks as shown in Fig. 7A. Each square of Fig. 7A represents a neighborhood block for which data was collected, with green representing a higher grade than blue. In second instance, contiguous sets of these same-graded blocks were dissolved into larger polygons. Figure 7B shows these two same-graded polygons. Finally, the polygons in Fig. 7C were then intersected with the census tracts. These are our units of observation. Each number represents a different census tract which was overlayed with the polygons in Fig. 7B to yield a unit of observation. Census tracts 1 to 3 represent unique, A-graded observations. Census tracts 4 to 7 represent unique B-graded observations. It is worth noting that since the blocks were dissolved into larger polygons, as shown in Fig. 7C, a census tract can span several HOLC blocks. Census tract 8, on the other hand, spans both A and B grades. This census tract, and others like it, was dropped and is therefore represented by a line-patterned fill. In general, census tracts that spanned several grades were dropped, as the RDD set-up would yield effects that mechanically tended towards zero otherwise (i.e. differences across borders would be nonexistent due to same outcome variable).

Table 3 features pre-treatment descriptive statistics for each grade, where odd columns are statistics for included census tracts, and odd columns are statistics for excluded census tracts. With respect to the number of observations that are dropped in this process, favored grades ("A" and "B") have a larger number of dropped observations than discriminated grades ("C" and "D"). Furthermore, broadly, tracts that are excluded for favored grades tend to fare better in terms of variables such as home value. The opposite is true to discriminated grades. With respect to the share of Black residents, there are no notable differences in included versus excluded tracts, exception being for "D" grades where included tracts feature an average of 15% Black residents while excluded tracts feature only 4%. Taken together, these statistics suggest that for favored grades, excluded tracts - by virtue of spanning less favored grades - are generally worse off. The opposite is true for discriminated grades.

Table 3 Pre-treatment descriptive statistics for included and excluded census tracts for each grade

	A		B		C		D	
	Included	Excluded	Included	Excluded	Included	Excluded	Included	Excluded
Home value	6584	6323	5213	5333	3625	4235	2093	3687
	(4226)	(3555)	(2843)	(3173)	(2493)	(2696)	(1950)	(2462)
Population	4070	6323	5213	5333	3625	4235	2093	4696
	(2325)	(3555)	(2843)	(3173)	(2493)	(2696)	(1950)	(2820)
Age	35	35	34	34	33	34	33	33
	(2)	(2)	(3)	(2)	(3)	(2)	(3)	(3)
Female	0.53	0.53	0.52	0.52	0.51	0.51	0.49	0.051
	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)	(0.03)
Rent	38	37	34	34	29	30	21	28
	(20)	(18)	(15)	(15)	(13)	(13)	(10)	13
Owner-occupied	0.51	0.53	0.47	0.47	0.38	0.41	0.26	0.38
	(0.23)	(0.21)	(0.19)	(0.20)	(0.19)	(0.20)	(0..17)	(0.19)
Foreign-born white	0.11	0.11	0.12	0.12	0.17	0.15	0.20	0.16
	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.09)	(0.08)
Black	0.01	0.01	0.013	0.01	0.02	0.03	0.15	0.04
	(0.01)	(0.02)	(0.04)	(0.04)	(0.09)	(0.11)	(0.29)	(0.12)
Observations	120	188	351	570	1138	883	1268	577

Note: This table features descriptive statistics regarding pre-treatment variables for each grade, featuring both included (odd columns) and excluded (even columns) census tracts. Each odd row features a mean for the variable in question, and each even row features the standard deviation in parentheses

The main outcome variable is the proportion of married individuals at the census tract level. This statistic is calculated as the number of married individuals, over the total number of married and never-married. The dates that will be used in the post-treatment analysis include: 1950, 1960, and 1970. It is held that the maps were less binding after the Fair Housing Act of 1968, which rendered area-based housing credit discrimination illegal.

Table 4 presents the proportion of married individuals for each grade by year. Aligned with the progressive decline of marriage, for each grade, the marriage rate declines from 1930 to 1970. Other notable trends are that D-graded areas have less marriage that other grades. Furthermore, there seems to be a decrescendo in marriage as the areas become less desirable. In general, it seems that marriage is positively correlated with area desirability.

5 Estimation

This section will describe the estimation strategy that will be deployed in order to understand the causal impacts of the HOLC maps on marriage. In particular, it will start by outlining the basic estimation strategy - the spatial RDD. It will discuss the

Table 4 Proportion married across grades and years

	Marriage			
	1930	1950	1960	1970
A	0.61	0.75	0.74	0.69
	(0.14)	(0.09)	(0.10)	(0.09)
Obs	55	550	1003	1159
B	0.60	0.74	0.74	0.69
	(0.12)	(0.08)	(0.08)	(0.09)
Obs	165	1937	2902	3378
C	0.61	0.74	0.74	0.68
	(0.09)	(0.07)	(0.08)	(0.09)
Obs	317	4430	6139	6985
D	0.59	0.72	0.70	0.63
	(0.09)	(0.08)	(0.10)	(0.11)
Obs	425	3908	5144	5537

Note: This table features descriptive statistics regarding proportion of married individuals for each grade, across the post-treatment years of 1950, 1960, 1970. Each column represents a year, and each set of three rows represents a grade. Within each of these sets, the first row is the mean proportion married, the second row is the standard deviation, and the third row is the number of observations. Proportion married is calculated as the total number of married residents over the total number of married and never-married residents

estimating assumptions of this method, and how they are not satisfied. Subsequently, the regression discontinuity difference-in-differences (RDD-DID) method is introduced as a way to overcome these estimation obstacles.
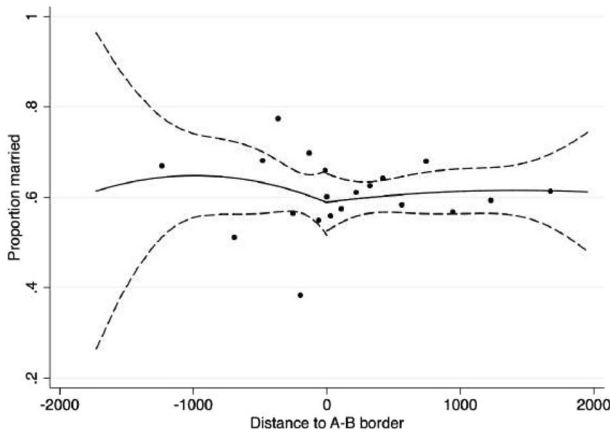
As is clear from Fig. 4, there are sharp spatial discontinuities in treatment. Given this set-up, the natural estimation strategy would be to perform a spatial RDD, following Black (1999). As with any RDD, the basic intuition is that there is are adjacent observations that are similar in all respects except for their exposure to a policy. This policy is assigned according to a particular running variable, which in this case would be distance to a border. The policy effects are given by:

$$E[Y_1 - Y_2|X = c] = E[Y_1|X = c] - E[Y_2|X = c]$$
$$= \lim_{x \downarrow c} E[Y|X = x] - \lim_{x \uparrow c} E[Y|X = x] \qquad (1)$$

Where Y_1 is the outcome for the treated group, Y_2 is the outcome for the untreated group, X is the running variable (i.e. distance to the D-A/D-B/D-C border), and c is the cut-off that marks the discontinuity in treatment (i.e. zero). Essentially the idea is to compare observations that approach the cut-off from either side.

5.1 RDD-DID

Plans to perform a standard RDD are disrupted by the fact that these borders were not randomly drawn. With respect to the standard RDD method, the identifying assumption here is that there are no discontinuities across the border for any variable
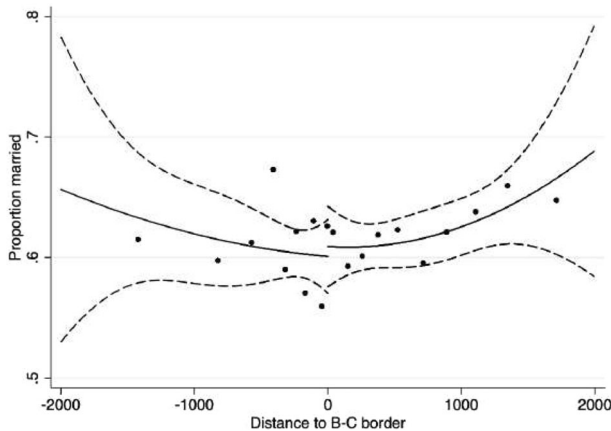
Fig. 8 Pre-treatment marriage cross-border comparisons: A-B. *Note*: Only observations within 2000 meters of the border are used. Negative distances refer to A-sided areas while positive distances refer to B-sided areas. Observations were grouped into ten bins (for each side of the border) according to distance to the border. Accordingly, each dot on the graph represents the mean proportion married for each distance-to-border decile. The solid Black lines represent the predicted values of the regression of proportion married on a second-order polynomial of distance to border, where each border side has unique parameters. The dashed lines represent the 95% confidence intervals of the fitted polynomial function

except treatment. Formally this amounts to arguing that $E[Y_1|X]$ and $E[Y_2|X]$ are continuous at $X = c$. In practice this amounts to arguing that neighbors across the border are the same except for their exposure to policy. This would be the case if, for instance, the borders had been drawn randomly. As was discussed, the borders were based on careful block level surveys, which implies that this is likely not the case. Given the granularity of the surveys it is indeed unlikely.

Figures 8–9 show pre-treatment (1930) graphs for differences across different border types. Indeed, there seem to be problematic discontinuities for B-C border pairs. Again, this is not surprising given that block grades were attributed on the basis of careful area-level data collection. It worth noting that there is a concern with respect to these graphs. This is due to the specific data set-up. In particular, recall that the observation is the census tract. Census tracts are defined as a given constant population, with differing physical area. This means that densely-populated places will have physically smaller census tracts. Recall also that distances to the border are calculated from the census tract centroid to the closest part of the border, meaning that physically smaller (i.e. more densely populated) census tracts will by construction be closer to the border. The graphs do not account for these data specificities, which means that in an extreme case, observations in the bins surrounding the border may belong to one densely populated city (e.g. New York), while bins farther away may correspond to a sparsely populated city. This is problematic if spatial trends differ across cities.

To account for this, this study performs an analogous pre-treatment analysis that is regression-based rather than visual. In particular, it deploys a pre-treatment RDD to see if there are important discontinuities in the variable of interest.

In particular, the baseline estimation strategy that is used is a spatial RDD with border fixed effects, and linear distance to border. In order for the linear distance

Fig. 9 Pre-treatment marriage cross-border comparisons: B-C. *Note*: Only observations within 2000 meters of the border are used. Negative distances refer to B-sided areas while positive distances refer to C-sided areas. Observations were grouped into ten bins (for each side of the border) according to distance to the border. Accordingly, each dot on the graph represents the mean proportion married for each distance-to-border decile. The solid Black lines represent the predicted values of the regression of proportion married on a second-order polynomial of distance to border, where each border side has unique parameters. The dashed lines represent the 95% confidence intervals of the fitted polynomial function

function to successfully account for spatial trends, a restricted sample of observations around the border are considered, following methods outlined in Calonico et al. (2014). The estimating equation in this case amounts to:

$$Y_{ib} = \omega_b + \beta_1 Grade_i + \beta 2 Dist_i + \epsilon_{ib} \tag{2}$$

Where Y_{ib} is the outcome for census tract i, closest to border b; ω_b is the border fixed effect; $Grade_i$ is a dummy for the grade of census tract i; $Dist_i$ is the distance from the centroid of census tract i to the closest border. Again, here we assume the running variable takes a linear form, and restrict observations on either side of the border to those for which this is the case.

Here, as for other specifications, standard errors are clustered at the city level. The justification for this clustering follows Abadie et al. (2022), where authors note that standard errors should be clustered with respect to the groups inherent to the sampling process and assignment mechanisms. In this case, the treatment is A to D assignment at the block level. With respect to the sampling process, we observe blocks that - as will be discussed later on - belonged to cities that received the maps, and more specifically, had populations of over 40,000. Thus, inclusion in our sample is determined at the city level. With respect to the assignment mechanism, blocks received "A" to "D" designation depending on the city they belonged to. For instance, a block with given average house price in New York City may received a "B" designation, but due to differing average house prices, may receive an "A" designation in Pittsburgh.

Table 5 features the RDD effects described in Figs. 8–10, accounting for the data problems outlined in the previous paragraph.
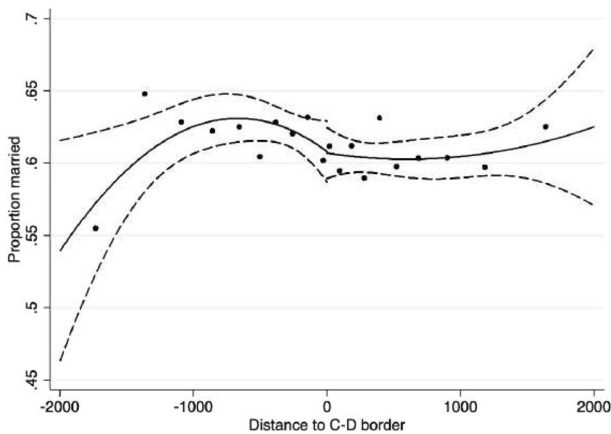
As Table 5 shows, there are no apparent RDD discontinuities for all grade pairs. This is a positive sign for the method that will later be introduced (RDD-DID).

Table 5 Pre-treatment RDD of maps on marriage

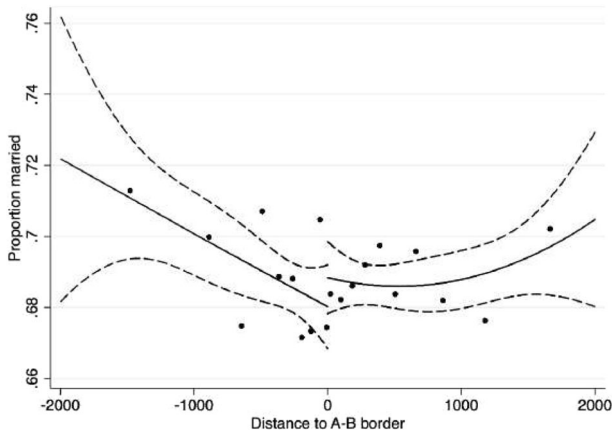|	Proportion married		
	(1)	(2)	(3)
B (versus A)	−0.00179		
	(−0.15)		
C (versus B)		0.00508	
		(0.80)	
D (versus C)			0.00402
			(0.47)
Observations	142	385	565

Note: This table features simple spatial RDD effects for different border pairs. The outcome varies according to the column. Here, border fixed effects and a distance-to-border first degree polynomial function are used. Observations are limited to a distance bandwidth around which a linear polynomial is justified. Different border pairs are considered in each row. Standard errors clustered at the city level in parentheses

$^+p<0.10$, $^*p<0.05$, $^{***}p<0.01$

Fig. 10 Pre-treatment marriage cross-border comparisons: C-D. *Note*: Only observations within 2000 meters of the border are used. Negative distances refer to C-sided areas while positive distances refer to D-sided areas. Observations were grouped into ten bins (for each side of the border) according to distance to the border. Accordingly, each dot on the graph represents the mean proportion married for each distance-to-border decile. The solid Black lines represent the predicted values of the regression of proportion married on a second-order polynomial of distance to border, where each border side has unique parameters. The dashed lines represent the 95% confidence intervals of the fitted polynomial function

However, as will be shown in robustness checks, is not necessarily a sufficient condition for the robustness of the method. This study deploys the method used in another redlining study (Krimmel 2018). In this study, authors deploy the RDD-DID method. This stands in contrast to another redlining impact evaluation (Aaronson et al. 2017), where authors augment the spatial RDD method by attributing more weight to border-pairs that are similar prior to the treatment. To implement the RDD-DID,

Fig. 11 Post-treatment (1970) marriage cross-border comparisons: A-B. *Note*: Only observations within 2000 meters of the border are used. Negative distances refer to A-sided areas while positive distances refer to B-sided areas. Observations were grouped into ten bins (for each side of the border) according to distance to the border. Accordingly, each dot on the graph represents the mean proportion married for each distance-to-border decile. The solid Black lines represent the predicted values of the regression of proportion married on a second-order polynomial of distance to border, where each border side has unique parameters. The dashed lines represent the 95% confidence intervals of the fitted polynomial function

this study essentially calculates a DID method with treatment interacted with the time variable, and also includes a border fixed effect, as well as a linear form of the running variable (distance to border). This analysis restricts the sample so that linear distance holds. The regression equation amounts to:
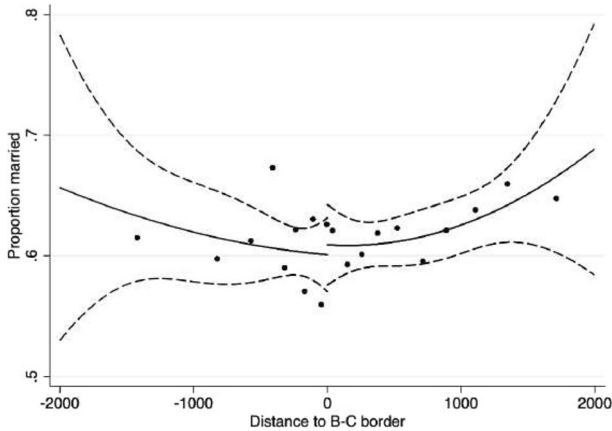
$$Y_{ibt} = \omega_b + \beta_1 Grade_i \times Post_t + \beta_2 Grade_i + \beta_3 Post_t + \beta4 Dist_i + \epsilon_{ibt} \qquad (3)$$

Where Y_{ibt} is the outcome for census tract i, closest to border b, at time t; ω_b is the border fixed effect; $Grade_i$ is a dummy for the grade of census tract i; $Post_t$ is a dummy that takes the value of one if the year is post-treatment (i.e. 1950, 1960, or 1970), and zero if the year is pre-treatment (i.e. 1930); $Dist_i$ is the distance from the centroid of census tract i to the closest border. Again, here we assume the running variable takes a linear form, and restrict observations on either side of the border to those for which this is the case.

6 Results

Figures 11–13 show raw discontinuities in post-treatment marriage across the border, for each border pair. In particular, Fig. 11 shows that B-graded areas had more marriage when compared to analogous areas in A-graded areas. Similar results hold for B-C comparisons. On the other hand, there do not seem to be pre-treatment discontinuities for C-D areas. While these results may seem counter-intuitive, it may be the case that these raw discontinuities, which do not account for certain biases. This could be the case if, for example, residents of C-graded areas are effectively barred (due to e.g. historical racial zoning) from living in better-graded areas. In this case, individuals with more financial security may choose to live closer to better-

Fig. 12 Post-treatment (1970) marriage cross-border comparisons: B-C. *Note*: Only observations within 2000 meters of the border are used. Negative distances refer to B-sided areas while positive distances refer to C-sided areas. Observations were grouped into ten bins (for each side of the border) according to distance to the border. Accordingly, each dot on the graph represents the mean proportion married for each distance-to-border decile. The solid Black lines represent the predicted values of the regression of proportion married on a second-order polynomial of distance to border, where each border side has unique parameters. The dashed lines represent the 95% confidence intervals of the fitted polynomial function

Fig. 13 Post-treatment (1970) marriage cross-border comparisons: C-D. *Note*: Only observations within 2000 meters of the border are used. Negative distances refer to C-sided areas while positive distances refer to D-sided areas. Observations were grouped into ten bins (for each side of the border) according to distance to the border. Accordingly, each dot on the graph represents the mean proportion married for each distance-to-border decile. The solid Black lines represent the predicted values of the regression of proportion married on a second-order polynomial of distance to border, where each border side has unique parameters. The dashed lines represent the 95% confidence intervals of the fitted polynomial function

graded areas. If this were the case, then the patterns outlined in Figs. 11–13 may arise.

It is therefore necessary to account for these biases by using the aforementioned RDD-DID method. This remainder of this section shows the main results for the

Table 6 RDD-DID impact of maps on marriage

	Proportion married		
	(1) 1950	(2) 1960	(3) 1970
B (versus A) × Post	0.00240	0.000165	−0.0134
	(0.21)	(0.01)	(−1.14)
B (versus A)	0.00529	−0.000292	0.00959
	(0.46)	(−0.03)	(0.85)
Post	0.133*	0.110*	0.0733*
	(6.15)	(7.80)	(3.40)
Observations	2082	2898	3349
C (versus B) × Post	−0.0103$^+$	−0.00735	−0.0149*
	(−1.95)	(−1.48)	(−2.81)
C (versus B)	0.00291	0.00580	0.0101*
	(0.65)	(1.16)	(2.06)
Post	0.114*	0.104*	0.0416*
	(10.51)	(9.08)	(3.77)
Observations	6867	8674	9528
D (versus C) × Post	−0.00653	−0.0109	−0.0208*
	(−1.11)	(−1.59)	(−3.46)
D (versus C)	0.00235	0.00799	0.00685
	(0.35)	(1.18)	(1.04)
Post	0.102*	0.0879*	0.0156*
	(15.49)	(10.39)	(2.21)
Observations	9685	11536	12285

Note: This table features RDD-DID effects for different border pairs for post-treatment years of 1950, 1960, 1970. Each column represents a different year, while the three sections of the table represent different border pairs. The outcome is the proportion of married individuals in the area, which is calculated as the total number of married individuals over the number of never-married and married individuals. Standard errors clustered at the city level in parentheses

$^+p < 0.10$, $^*p < 0.05$, $^{***}p < 0.01$

impact of redlining on marriage in 1950, 1960, and 1970. In particular, Table 6 shows these results, with each column describing a post-treatment year, and each set of three rows describing a particular grade pair. The first row of each set gives us the treatment effect. The first set, that is the impact of B versus A grades, does not seem to show any effects for any years. This may be due to the fact that both areas were in fact favored. The second row, C versus B, shows effects for 1950 and 1970. These effects are increasing between 1950 and 1960/1970, and show evidence of "compounding" policy effects. In particular, in 1950, C designation versus B designation decreased marriage by 1 percentage point. In 1970, this effect increases to 1.7 percentage points. Finally, the last set of rows, which shows effects for D versus C areas, shows effects for 1970. In this year, D designation decreased marriage by 2

Table 7 Pre-treatment RDD of maps on other outcomes

	House prices (1)	Population (2)	Age (3)	Female (4)
B (versus A)	540.8	−91.23	−0.548	0.000664
	(0.68)	(−0.19)	(−1.17)	(0.12)
Observations	320	320	320	320
C (versus B)	−736.0[*]	−27.21	−0.418[*]	−0.00714[*]
	(−2.48)	(−0.10)	(−2.14)	(−2.46)
Observations	1126	1126	1123	1123
D (versus C)	−298.4[+]	382.4[*]	−0.442[*]	−0.00673[*]
	(−1.73)	(2.00)	(−2.35)	(−2.51)
Observations	1947	1947	1934	1934

Note: This table features simple spatial RDD effects for different border pairs for 1930 (pre-treatment). The outcome varies according to the column. Here, border fixed effects and a distance-to-border linear polynomial function is used. Observations are limited to a distance bandwidth around which a linear polynomial is justified. Different border pairs are considered in each row. Standard errors clustered at the city level in parentheses

[+]$p < 0.10$, [*]$p < 0.05$, [***]$p < 0.01$

percentage points. It is worth noting that differences between B-C and C-D borders, where the latter featured effects of a larger magnitude, are to be expected given that D-graded areas were more heavily discriminated against.

It is worth noting that there are several other forces, in addition to the maps, that may have affected marriage during the during the time period in question. For instance, female labor participation increased markedly (Cebula & Coombs 2008), and a fundamental shift in expectations regarding marriage occurred (Coontz 2007). It is important to note that these trends may have affected residents of different grades to a different extent. However, the estimation strategy - which aims to account for a priori demographic differences - should account for these.

6.1 Parallel trends

The RDD-DID method hinges on parallel trends across time, conditional on border fixed effects and a distance function. This assumption may not hold, if, for instance, there are other variables for which pre-treatment gaps are found. These other variables may in fact interact with exogenous shocks (e.g. wider cultural shifts in approaches to marriage), to yield "unparallel" trends between grades. Accordingly, Tables 7 and 8 feature the same pre-treatment RDD analysis as in Table 4, featuring other relevant outcomes that may interact with shocks as previously described. As Tables 7 and 8 show, there are significant differences between border sides with respect to almost all the featured survey variables.

These pre-treatment gaps could be problematic, provided such shocks exist. However, in Table 9, the outcome is directly tested for parallel trends. In particular,

Table 8 Pre-treatment RDD of maps on other outcomes (cont.)

	Rent	Owner-occupier	Foreign-born white	Black
	(1)	(2)	(3)	(4)
B (versus A)	−0.0354	0.0319	−0.0108	−0.0248
	(−0.01)	(1.20)	(−1.22)	(−1.44)
Observations	1126	1120	1122	1122
C (versus B)	−3.526*	0.00531	−0.000955	0.0103$^+$
	(−2.73)	(0.38)	(−0.19)	(1.79)
Observations	1126	1120	1122	1122
D	−1.482	−0.0187*	−0.00134	0.0265$^+$
	(−1.50)	(−2.11)	(−0.28)	(1.75)
Observations	1947	1926	1934	1934

Note: This table features simple spatial RDD effects for different border pairs for 1930 (pre-treatment). The outcome varies according to the column. Here, border fixed effects and a distance-to-border linear polynomial function is used. Observations are limited to a distance bandwidth around which a linear polynomial is justified. Different border pairs are considered in each row. Standard errors clustered at the city level in parentheses

$^+p < 0.10$, $^*p < 0.05$, $^{***}p < 0.01$

marriage for 1934 is compared to 1930. As is shown in the table, there do not seem to be diverging trends, and the gaps shown in Tables 7 and 8 do not seem to matter.

6.2 Housing credit mechanism

This section verifies whether the housing credit mechanism per se is driving the previous results. In particular, there are two mechanisms potentially yielding the results in Table 6. The first is the housing credit mechanism, which we are interested in isolating. This is defined as a first-order impact of the redlining maps. For instance, if a couple is unable to secure housing credit, they may not get married. The second mechanism is that of area-level decline. These are defined as second-order effects resulting from the first-order channel of individuals being unable to access housing credit. For instance, local businesses may struggle to gain access to credit, leading to a loss of jobs at the area level, and an associated decrease in incomes. In order to understand whether the results are at least partially driven by housing credit per se, Table 10 includes contemporaneous house prices in the matching process. The idea is to control for the second-order effects of individuals being denied housing credit, namely a decrease in desirable amenities (e.g. rising crime rates). If the first-order housing credit were irrelevant, the coefficients in Table 6 would drop to null significance upon including a control for house prices. It is clear from the table, however, that the coefficients remain similar in magnitude and significance. It is therefore likely that the results are driven by housing credit.

Note, the magnitude of the treatment coefficients, conditional on house prices, is informative insofar as it remains significant. It suggests that the treatment effect, and therefore the policy, is at least partially driven by the aforementioned first-order

Table 9 Robustness check: parallel trends

	Proportion married		
	(1)	(2)	(3)
B (versus A) × Post	0.00404		
	(0.14)		
B (versus A)	−0.000424		
	(−0.04)		
C (versus B) × Post		0.00635	
		(0.53)	
C (versus B)		0.0159*	
		(2.31)	
D (versus C) × Post			0.00450
			(0.47)
D (versus C)			0.00643
			(0.87)
Post	0.0459	0.0508	0.168*
	(1.01)	(1.64)	(25.46)
Observations	300	1129	1837

Note: This table features RDD-DID effects for different border pairs for two pre-treatment years of 1934 and 1930. This estimation serves as a placebo test for parallel trends. The outcome is the proportion of married individuals in the area, which is calculated as the total number of married individuals over the number of never-married and married individuals. Standard errors clustered at the city level in parentheses
$^+p < 0.10$, $^*p < 0.05$, $^{***}p < 0.01$

effects. However, the coefficients are likely lower bound estimates of this first-order channel - and not simply due to "bad control" issues. In particular, discriminatory mortgage disbursal has a direct effect on housing demand, depressing house prices. By controlling for house prices, the analysis is also partially controlling for the first-order impacts the study is interested in isolating.

Table 11 provides an additional check for whether the credit mechanism is responsible for the area-level effects. The intuition is that credit-constrained areas are more likely to be affected by the policy, if the credit mechanism is driving results. To understand whether this is the case, the analysis assumes that D-graded observations farther from the C-D border are more credit constrained, as the tendency would be for areas to become wealthier farther from the D centroid and into higher-graded areas. Accordingly, the analysis compares wealthier and less wealthy D-graded areas. In this case, the "placebo" variable takes the value of one if the D-graded observation is closer to the D centroid (i.e. "less wealthy") and takes the value of zero if the observation is farther away from the D-graded centroid (i.e. "wealthier"). If the hypothesis holds, then the coefficent for the RDD-DID interaction term would be significantly smaller than zero. Indeed, this seems to be the case, confirming that the effects may indeed be due to the credit mechanism.

Table 10 RDD-DID impact of maps on marriage controlling for house prices

	Proportion married		
	(1)	(2)	(3)
B (versus A) × Post	0.00302	0.00182	−0.0152
	(0.17)	(0.14)	(−1.03)
B (versus A)	0.00689	−0.00302	0.0152
	(0.43)	(−0.25)	(1.08)
Post	0.0901*	0.0763*	0.0280+
	(4.98)	(4.78)	(1.76)
House prices	0.00000124	0.000000100	0.000000540
	(1.17)	(0.16)	(1.46)
Observations	1781	2579	2971
C (versus B) × Post	−0.00541	−0.00640	−0.0126*
	(−1.03)	(−1.22)	(−2.42)
C (versus B)	0.000427	0.00252	0.00769
	(0.09)	(0.49)	(1.62)
Post	0.0870*	0.0871*	0.0175*
	(13.13)	(11.83)	(2.28)
House prices	0.000000522	−0.00000119*	−0.000000134
	(1.42)	(−3.08)	(−0.53)
Observations	6185	7917	8668
D (versus C) × Post	−0.00807	−0.0107+	−0.0186*
	(−1.42)	(−1.77)	(−3.15)
D (versus C)	0.00537	0.00822	0.00730
	(0.94)	(1.49)	(1.24)
Post	0.0929*	0.0794*	0.00930
	(11.54)	(10.56)	(1.27)
House prices	0.00000124	0.000000100	0.000000540
	(1.17)	(0.16)	(1.46)
Observations	8636	10,342	10,818

Note: This table features RDD-DID effects for different border pairs for post-treatment years of 1950, 1960, 1970, controlling for house prices. Each column represents a different year, while the three sections of the table represent different border pairs. The outcome is the proportion of married individuals in the area, which is calculated as the total number of married individuals over the number of never-married and married individuals. Standard errors clustered at the city level in parentheses

+$p < 0.10$, *$p < 0.05$, ***$p < 0.01$

6.3 Sorting

As is shown, the policy led to a decrease in marriage, particularly in 1960 for A-B pairs, 1950 and 1970 for B-C pairs, and all years for C-D pairs. However, it is worth

Table 11 Credit mechanism placebo

	Proportion married		
	(1) 1950	(2) 1960	(3) 1970
Placebo × Post	−0.00484[+]	−0.00592[+]	−0.00212
	(−1.68)	(−1.84)	(−0.62)
Placebo	0.00452[*]	0.00450[*]	0.00431[+]
	(2.02)	(2.06)	(1.83)
Post	0.100[*]	0.0839[*]	0.00476
	(16.09)	(8.83)	(0.69)
Observations	9685	11,536	12,276

Note: This table features RDD-DID effects for a placebo test, for post-treatment years of 1950, 1960, 1970. The variable "placebo" takes the value of one if the observation is closer to the D-grade centroid, and takes the value of zero if the observation is farther from the D-grade centroid. Each column represents a different year. The outcome is the proportion of married individuals in the area, which is calculated as the total number of married individuals over the number of never-married and married individuals. Standard errors clustered at the city level in parentheses

[+]$p < 0.10$, [*]$p < 0.05$, [***]$p < 0.01$

noting that these are area-level results, and therefore potentially the result of sorting. It is possible that the higher proportion of unmarried individuals in B-graded areas simply reflects the reshuffling of individuals across grades, rather than pointing towards the effect of housing credit on marriages. This could have happened if, for instance, unmarried individuals are more likely to rent, and B-graded areas, due to housing credit discrimination, had higher renter-occupier rates. In order to understand whether this is the case, a city-level analysis is performed. It is held that if receiving a redlining map led to a decrease in city-level marriage, then the area-level results do not merely reflect the re-shuffling of individuals. In other words, if the main results led to a rise in marriagelessness for indigenous residents, then marriage at the city level should be lower. If the area-level results simply reflect the fact that the maps led unmarried people to move to discriminated areas, there would be no effects at the city level.

It is worth noting that in theory, the maps could have led to more marriage in A-graded areas (as these were actively favored), and, simultaneously, less marriage in D-graded areas (as these were actively discriminated against). In the case where both these propositions hold, the city-level effects would be ambiguous. On one hand, we might see more city-level marriage driven by A-graded areas. On the other, we would observe less city-level marriage driven by D-graded areas. However, as is shown in Table 5, the area-level estimates suggest that positive discrimination in A areas did not decrease area-level marriage. In contrast, discrimination in C and D areas did decrease marriage. In this case, we would expect city-level aggregate estimates to reflect the effects of negative discrimination. In other words, the effects we observe at the city level are not masked by competing effects that may have resulted from positive discrimination. It is worth noting, however, that these city-level results do not account for inter-city sorting. It is possible that in light of

Table 12 City-level impacts of redlining on marriage

| | Proportion married | | |
| | (1) | (2) | (3) |
	1950	1960	1970
1930s population cut-off	−0.00265	−0.00290[+]	−0.00207[+]
	(0.00170)	(0.00158)	(0.00121)
1930s population (millions)	−0.00441	−0.00762	−0.00679
	(0.00921)	(0.00852)	(0.00655)
1930s population (millions) squared	0.00140	0.00241	0.00201
	(0.00335)	(0.00310)	(0.00238)
Constant	0.0133[*]	0.0119[*]	0.0133[*]
	(21.62)	(21.02)	(30.74)
Observations	2997	3083	3088

Note: This table features city-level observations for 1930. It includes results for the city-level RDD where the HOLC population cut-off is the relevant discontinuity. Standard errors clustered at the city level in parentheses

[+]$p < 0.10$, [*]$p < 0.05$, [***]$p < 0.01$

discriminatory policy, families moved to cities that did not receive the maps. In this case, city-level results would be tainted by the fact that, for instance, individuals less likely to marry did not move away. This is unlikely, however, given that the maps were covertly circulated and the general public was unaware of their existence.

In order to perform this analysis, county-level historical data on marriage and population is used. This data is provided by the Integrated Public Use Microdata Series' National Historical GIS (Manson et al. 2019). Data on county-level marriage and population is used for 1950, 1960, and 1970. Furthermore, data for 1930s population is used.

In order to understand whether the area-level results are due to sorting, this study exploits quasi-randomness in treatment assignment at the city level. In particular, it exploits the fact that every city with a 1930s population above 40,000 received a map (Hillier 2005). This clean, but arguably random, cut-off, lends itself to a standard RDD set-up. Thus, the estimating equation amounts to:

$$Y_i = \alpha + \beta_1 Cut_i + \beta_2 Pop1930_i + \beta_3 Pop1930Sq_i + \epsilon_i \qquad (4)$$

Where Y_i is the number of marriages per capita in county i, Cut_i is a dummy that takes the value of one if county i had a population over 40,000 in 1930, 0 otherwise; $Pop1930_i$ and $Pop1930Sq_i$ are the 1930s population polynomial.

Table 12 features the city-level results. The first line of the table features the effect of interest - a dummy variable that takes the value of one if the city received a map. The remaining two lines feature the second-order polynomial control function of the population variable. Column 1 features results for 1950, column 2 for 1960, and column 3 for 1970. Indeed, it seems that for 1960 and 1970, the HOLC maps led to less overall marriage per capita. It is worth noting that the 1950 coefficient is close to the required significance cut-off. In sum, there there were city-level changes in

marriage due to the maps. Taking the area-level and city-level results together, evidence strongly suggests that access to housing credit affects the decision to marry.

7 Conclusion

This study analysed the extent to which obtaining housing credit affects the decision to get married. Given that it evaluated the impact of an area-based housing credit discrimination policy that was heavily punitive towards Black Americans, it also considered the extent to which state-enforced inequality in home ownership contributed to the racial marriage gap. The findings suggest that the HOLC maps did indeed negatively impact marriage in discriminated areas. Furthermore, given that the residents of these areas were predominantly Black, these results point to the fact that redlining maps increased the racial marriage gap. Furthermore, in finding that the policy's negative effects on marriage were not due to sorting or area-level decline, there are some avenues for potential further work.

The most obvious stream of ensuing research would be to understand the intergenerational effects of this discriminatory policy on Black skill acquisition or income more generally. Other works could also focus on further quantifying the income effects of racial segregation in US cities in the beginning of the 20th century. Other streams of research that could follow from this work would be to understand the extent to which the rise in cohabitation (versus marriage) is due to rising house prices. Understanding the reasons for the rise of cohabitation are policy-relevant, as children raised in married, versus cohabiting, households, have better outcomes (Doepke et al. 2022). Married parents invest more in their children, and are less likely to break up and consequently cause disruptions in skill acquisition. While this study found that the inability to access housing credit deters individuals from marriage, the link between trends towards cohabitation and house prices has not been established. Another, related, potential stream of work could be to understand the extent to which access to housing credit could incentivize marriage, as other benefits to marriage decrease. As education gaps between men and women narrow, so does the wage differential, and, consequently, the benefits to household specialization.

With respect to other evidence on the impact of redlining maps, this study sought to understand whether the observed effects on marriage were due to the housing credit mechanism per se, or whether due to the documented effect on area-level decline (see e.g. Krimmel 2018 or Aaronson et al. 2017). It showed that the effects are due to the mortgages per se. Given that the housing credit mechanism is a first-order consequence of the redlining maps, and given that this mechanism affected marriage causally, it can therefore be held that the policy's final effects on outcomes like e.g. income, can at least partially be assumed to be mediated by marriage effects.

Author contributions Margarida Madaleno wrote the entire paper.

Compliance with Ethical Standards

Conflict of interests The authors declare no conflicts of interest.

References

Aaronson, D., Hartley, D., & Mazumder, B. (2017). The effects of the 1930s HOLC "redlining" maps. Tech. Rep., Working Paper.

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2022). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, *138*, 1–35.

Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political Economy*, *81*, 813–846.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, *114*, 577–599.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*, 2295–2326.

Carozzi, F., Hilber, C. A. L., & Yu, X. et al. (2024). On the economic impacts of mortgage credit expansion policies: evidence from help to buy. *Journal of Urban Economics*, *139*, 103611.

Caucutt, E. M., Guner, N., & Rauh, C. (2016). Is marriage for White people? Incarceration and the racial marriage divide. *Unpublished paper, Department of Economics, Western University*.

Cebula, R. J., & Coombs, C. K. (2008). Recent evidence on factors influencing the female labor force participation rate. *Journal of Labor Research*, *29*, 272–284.

Coontz, S. (2007). The origins of modern divorce. *Family Process*, *46*, 7–16.

Eriksen, M. D. (2010). Homeownership subsidies and the marriage decisions of low-income households. *Regional Science and Urban Economics*, *40*, 490–497.

Hillier, A. E. (2003). Redlining and the home owners' loan corporation. *Journal of Urban History*, *29*, 394–420.

Hillier, A. E. (2005). Residential security maps and neighborhood appraisals: The home owners' loan corporation and the case of Philadelphia. *Social Science History*, *29*, 207–233.

Hu, M., & Wang, X. (2019). Homeownership and household formation: No homeownership, no marriage? *Journal of Housing and the Built Environment*, *35*, 1–19.

Krimmel, J. (2018). Persistence of prejudice: Estimating the long term effects of redlining.

Lafortune, J., & Low, C. (2017). Tying the double-knot: The role of assets in marriage commitment. *American Economic Review*, *107*, 163–167.

Manson, S., Schroeder, J., Van Riper, D., & Ruggles, S. (2019). Ipums national historicla geographical information system: Version 14.0 [database]. ipums. *Institute for Social Research and Data Innovation. University of Minnesota*.

Massey, D. S. (2015). The legacy of the 1968 Fair Housing Act. In *Sociological Forum*, (vol. 30, 571–588). Wiley Online Library.

Miller, J. J., & Park, K. A. (2018). Same-sex marriage laws and demand for mortgage credit. *Review of Economics of the Household*, *16*, 229–254.

Raley, R. K., Sweeney, M. M., & Wondra, D. (2015). The growing racial and ethnic divide in U.S. marriage patterns. *The Future of Children/Center for the Future of Children, the David and Lucile Packard Foundation*, *25*, 89.

Ricks, J. S. (2021). Mortgage subsidies, homeownership, and marriage: Effects of the va loan program. *Regional Science and Urban Economics*, *87*, 103650.