August 29, 2019

Memorandum

To: Members, Committee on Financial Services
From: FSC Majority Staff
Subject: September 4, 2019, "Examining Discrimination and Other Barriers to Consumer Credit, Homeownership, and Financial Inclusion in Texas"

The Subcommittee on Oversight and Investigations will hold a field hearing entitled "Examining Discrimination and Other Barriers to Consumer Credit, Homeownership, and Financial Inclusion in Texas," on Wednesday, September 4, 2019, at 10:00 a.m. CDT in Houston, Texas. The witnesses for the first panel are:

- Belinda Everette, Director, Housing Initiative, NAACP Houston Branch
- Judson Robinson III, CEO and Chair, Houston Area Urban League
- Hua Sun, Professor, University of Iowa
- John Wong, Founding Chair, Asian Real Estate Association of America
- Dedrick Asante-Muhammad, Chief, Race, Wealth, and Community, National Community Reinvestment Coalition

The second panel will include:

- Noel Poyo, Executive Director, National Association of Latino Community Asset Builders
- Gary Lindner, President and CEO, PeopleFund
- Jeff Smith, President and CEO, Unity National Bank
- Raymond Ardoin, President, Board of Directors, Brentwood Baptist Church Federal Credit Union
- Jeungho "JP" Park, President and Chairman, Relationship BancShares, Inc.
- Celina Peña, Chief Advancement Officer, LiftFund
- George Johnson, CEO, George E. Johnson Development

Overview

This hearing will examine access to affordable housing, credit, and banking services in low and moderate-income ("LMI") neighborhoods. The witnesses will address discrimination and other barriers to homeownership, credit access, and financial inclusion. The hearing will examine, among other data, recently published research on the perpetuation of systemic impediments to wealth-building and wealth disparities, homeownership, and economic opportunity.

In addition, the hearing will explore potential solutions that would promote financial inclusion and strengthen financial institutions that serve LMI communities, such as minority depository institutions ("MDIs")[1] and community development financial institutions ("CDFIs").[2] The number of MDIs has declined in the last decade, raising the policy challenge of how best to bolster the health and soundness of these community institutions that provide critical services to consumers and businesses underserved by

---

[1] The FDIC defines MDIs as: (1) depository institutions where 51% or more of the voting stock is owned by minority individuals, and (2) depository institutions that predominantly serve minorities and have a Board of Directors comprised mostly of minority individuals.

[2] CDFIs include mission-driven community development banks, credit unions, and funds that target their financing activity on LMI communities. The FDIC estimates that about half of all CDFI banks are MDIs.

underrepresented and obtained proportionally fewer loans than White applicants, even after accounting for income disparities. Moreover, the report found that home values in neighborhoods in which Black homebuyers are concentrated appreciated at a slower rate than those in predominantly White areas. This disparity was evident even in areas such as Houston and Dallas, where home prices were less volatile during and after the financial crisis, and where, for both racial groups, homes in 2017 were valued above 2006 levels.[11] To counteract these trends in the home mortgage market, the report recommended three policies to strengthen fair housing protections: fully reinstate the Affirmatively Furthering Fair Housing rule, increase funding for the Fair Housing Initiatives Program, and stop systematic discrimination by improving HUD's enforcement of the Fair Housing Act and its disparate impact rule.

However, rather than strengthening fair housing protections, the Administration has taken several steps to undermine them:

- On August 19, 2019, HUD proposed a new rule that would fundamentally alter the disparate impact standard under the Fair Housing Act. Among other things, the changes in the proposed rule would make it easier for landlords and mortgage lenders to use algorithms that may discriminate against tenants on previously-prohibited grounds. The proposed rule would also shift much of the burden of proof to plaintiffs, making it harder for victims of discrimination to obtain relief in court, including the large population of Limited English Proficient (LEP) persons in Texas for whom language discrimination is a common barrier to homeownership.[12]
- The Federal Housing Finance Agency recently proposed to remove the language preference question from the Universal Residential Loan Application form for mortgages backed by Fannie Mae and Freddie Mac. Without these data fields, the Uniform Mortgage Data Program cannot accurately capture borrower profiles and language needs.
- As a result of 2018 guidance to FHA lenders, recipients of the Deferred Action for Childhood Arrivals (DACA) program are ineligible for FHA-insured loans, which for decades have afforded low-income, minority borrowers access to homeownership.[13] As a result, each of the 120,000 DACA recipients in Texas faces this added barrier to homeownership.

Discrimination and Insufficient Data in the Small Business Lending Market

According to the Dallas Federal Reserve Bank's most recent survey of small businesses in Texas, 40% of firms applied for new financing in the past 12 months, which is the same rate as the national average.[14] Most firms (62%) sought financing to expand, pursue new opportunities, or replace capital assets. Notably, minority-owned firms are more than twice as likely to receive <u>none</u> of the funding for which they apply (33% of minority-owned firms received none of the financing sought, as compared to 15% of non-minority-owned firms). More than half (55%) of non-minority-owned firms received <u>all</u> of the financing they requested, while only 36% of minority-owned firms received the full amount sought.

Research suggests that these disparities may be attributable to discrimination. In a "mystery-shopping" study by the National Community Reinvestment Coalition (NCRC) utilizing testers, banks were found to

---

[11] *Id*. at 24.

[12] Harris County has the most language minority groups of all Texas counties, and LEP individuals are most commonly from Latinx or AAPI communities. Over one third of Texans speak a language besides English at home and a total of 164 languages are spoken across the state, according to a U.S. Census analysis.

[13] HUD, *Annual Report to Congress on the Fin. Status of the FHA Mutual Mortgage Ins. Fund*, 2018.

[14] Fed. Reserve Bank of Dallas, *Small Business Survey: Report on Employer Firms in Texas* (Dec. 2018).

be twice as likely to offer white entrepreneurs help with their small business loan applications as compared to Black entrepreneurs. The study also found that bankers were three times more likely to invite follow-up appointments with White borrowers than better-qualified Black borrowers. Overall, Black entrepreneurs seeking loans for their businesses were subjected to far more scrutiny than their equal or less creditworthy white counterparts.[15]

Quantification of disparate treatment of Black entrepreneurs by lenders is hampered by the lack of data on small business lending. By law, the Consumer Financial Protection Bureau (CFPB) must require institutions that lend to women-owned, minority-owned, or small businesses to report their lending activities with the same transparency as the home mortgage lending sector. In May 2017, pursuant to § 1071 of the Dodd-Frank Act, the CFPB issued a whitepaper and request for information (RFI) regarding the small business loan market, particularly in relation to financing available to minority-owned and women-owned small businesses. CFPB has failed to issue a proposed or final rule on this statutorily required reporting requirement. According to the Administration's Unified Agenda of Regulatory and Deregulatory Actions, the CFPB will not take action on a rule until January 2020 at the earliest.[16]

The Promise of Community Reinvestment.

The Community Reinvestment Act ("CRA"), 12 U.S.C. § 2901, *et seq*., governs how banks meet credit needs in the areas they serve, including LMI neighborhoods. Banks accrue CRA credits or points for qualifying activities within designated assessment areas and thereby earn performance ratings.[17] In assessing a bank's CRA compliance, federal regulators "may consider as a factor capital investment, loan participation, and other ventures undertaken … in cooperation with minority- and women-owned financial institutions … provided that these activities help meet the credit needs of local communities in which such institutions … are chartered."[18]

By regulation, CRA consideration is also available for community development activities done in cooperation with MDIs, community development lending to MDIs, investments in MDIs, community development services provided to MDIs, and even the donation or sale on favorable terms of bank branch locations to MDIs.[19] Despite these incentives, the number of MDI banks – which were hit disproportionately hard by the financial crisis[20] – has plummeted in the last decade. There are now just 148, down from 215 in 2008.[21] MDI credit unions have seen even more significant declines, with more

---

[15] NCRC, *et al*., *Shaping Small Business Lending Policy Through Matched-Paired Mystery Shopping* (Sept. 2017).

[16] The CFPB's inaction is the subject of a lawsuit seeking a declaratory judgment requiring the CFPB to issue the regulations. In June 2019, an amended complaint was filed adding as plaintiffs the National Association of Latino Community Asset Builders and two women small business owners.

[17] For details on CRA examinations and scoring of eligible activities by regulators, refer to this Committee's Memorandum dated April 4, 2019 in connection with the April 9, 2019 hearing entitled *The Community Reinvestment Act: Assessing the Law's Impact on Discrimination and Redlining*, at 2-3.

[18] 12 U.S.C. § 2903(b).

[19] *See* 12 CFR §§ 25.21(f), 25.23(d), 195.21(f), and 195.12(g) and OCC Fact Sheet: *Partnerships with Minority and Women-Owned Financial Institutions, Low-Income Credit Unions*, Aug. 2016.

[20] Federal Reserve Bank of Chicago, *Capital-raising among minority-owned banks before and after the financial crisis*, 2018.

[21] Today there are just 19 African American MDIs, down from 44 in 2007; 26 Hispanic MDIs, compared to 53 in 2008; 63 Asian American MDIs, down from 99 in 2008-2010; and 18 Native American banks, down from 22 in 2010. FDIC data files on MDIs, updated Aug. 2019.

than one-third of such institutions disappearing since June 2013.[22]

As shown in a recent FDIC study, MDIs originate a greater share of mortgages to borrowers in low- and moderate-income areas and in areas with larger shares of minority populations than non-MDIs.[23] Fewer MDIs could negatively impact the provision of credit to the persons and communities these insititutions serve. To strengthen MDIs, the National Bankers Association has called for enhanced incentives for majority-owned banks to make CRA-qualified investments in MDIs. These include a new MDI investment tax credit, provision of consistent CRA credit throughout the life of an investment, and greater CRA consideration of small business lending and loan amount thresholds.[24]

---

[22] As of December 2018, there were 529 MDI credit unions, which compares to 805 such firms as of June 2013. See NCUA, *Minority Depository Institutions Annual Report to Congress*, 2013 and 2018.

[23] FDIC, *2019 Minority Depository Institutions*, Jun. 2019. Significantly, the study also found that MDI financial performance has improved in the last five years, specifically with respect to loan performance and revenue generation.

[24] Testimony of B. Doyle Mitchell, President and CEO of Industrial Bank, on behalf of the Nat'l Bankers Ass'n, before House Fin. Services Subcomm. on Consumer Protection and Fin. Instit., Apr. 9, 2019.

APPENDIX A

Lending practices to same-sex borrowers

Hua Sun[a,1] and Lei Gao[a,1]

[a]Department of Finance, Ivy College of Business, Iowa State University, Ames, IA 50011

Edited by Stephen L. Ross, University of Connecticut, Storrs, CT, and accepted by Editorial Board Member Susan T. Fiske March 6, 2019 (received for review May 15, 2017)

Using massive US mortgage lending data, we propose a method to infer a borrower's sexual orientation indirectly without a self-identification requirement and demonstrate the method's potential to approximately measure the sexual orientation of the US population at the local level annually over decades. We continue to examine the lending practices to same-sex borrowers and its spillover effects. The persistent results since 1990 reveal that, in contrast with otherwise comparable different-sex loan applicants, the approval rate for same-sex applicants is ∼3–8% lower. Furthermore, conditional on approval, lenders, on average, charge about 0.02–0.2% higher interest to same-sex borrowers, which is equivalent to an annual total of $8.6 million to $86 million in additional interest/fees nationwide. Meanwhile, we find that same-sex borrowers are less risky overall, as they exhibit similar default risk but lower prepayment risk. Finally, we document findings of spillover effects. That is, when the share of a neighborhood's same-sex population increases, both same-sex and different-sex borrowers seem to experience more unfavorable lending outcomes overall. The findings should raise enough concerns to warrant further investigations.

same-sex | LGBT | mortgage | credit rationing | HMDA

In the United States, owning a home has long been associated with achieving the "American Dream." Based on US Census Bureau data, as of 2016, the national homeownership rate was 63.7%. One important vehicle that Americans can use to become homeowners is a mortgage. As of 2016, the total US residential loan balance was $11.45 trillion (1), in contrast to the national debt of $19.98 trillion (2). Due to the heavy reliance of US households on the mortgage market, an economically sound and fair lending mechanism is crucial to sustaining and promoting the welfare of society.

One disturbing concern in the credit market is the likely practice by some lenders of denying loans to selected groups of people for noneconomic reasons. Since the influential seminal work of Black et al. (3), over the last four decades, many researchers have studied lending and other discrimination based on skin color, gender, and race. Relevant discrimination studies can be found regarding the mortgage (4–10), rental housing (11, 12), and auto loan markets (13–16).

Many high-profile cases echo academic studies on discrimination. There have been several lawsuits alleging discrimination by auto loan lenders (and the affiliated auto manufacturers, if any) against minority borrowers by overcharging them on interest. Most of these bias suits ended with multimillion dollars settlements. For example, a case in 2016 against Toyota involved up to $21.9 million on such settlements (17).

Potentially discriminatory practices also are documented in the labor market. In 2009, Walmart agreed to an $11.7 million settlement in response to an accusation of denying jobs to female applicants from 1998 through February 2005 (18). Costco also settled an $8 million bias suit in 2013 for the claim of discriminating against women in promotions to management jobs (19).

Despite the ample archival evidence of discrimination regarding race and gender in consumer and labor markets, fewer cases of potentially unfair treatment based on sexual orientation have been reported. According to *Washington Blade*, a popular lesbian, gay, bisexual, and transgender (LGBT) news outlet, it was not until 2016 that the US Equal Employment Opportunity Commission (EEOC) reached its first settlement in history in an antigay bias case (20). The case is against Pallet Companies for firing a lesbian employee after she complained about being harassed by her supervisor due to her sexual orientation. Walmart also settled a bias suit with a $7.5 million in 2016, after being accused for denied spousal health insurance benefits to same-sex employees between 2011 and 2013 (21).

Motivated by the huge stake that US households have in the mortgage market and the lack of systematic investigation of sexual orientation-based lending practices, we examine potentially different lending treatments toward same-sex mortgage applicants. Although the federal Fair Housing Act (FHA) of 1968 and the Equal Credit Opportunity Act (ECOA) of 1974 prohibit lending discrimination based on a borrower's race, gender, marital status, religion, and so forth, they do not specifically include sexual orientation as a prohibited basis. To the best of our knowledge, no previous academic study examines whether mortgage lenders systematically provide less access to credit (e.g., higher denial rates) or unfavorable terms (e.g., charging higher interest) for LGBT borrowers. We suspect that this is mainly a data availability issue. Unlike gender, race, and related categories, a loan applicant's sexual orientation is not required to be disclosed and, hence, is impossible to be measured directly. In the wake of human rights equality for the LGBT community in recent decades, the study of mortgage lending discrimination for LGBT borrowers is timely.

Using the publicly available Home Mortgage Disclosure Act (HMDA) data, we propose a method to identify potentially homosexual borrowers (households) indirectly without a self-identification requirement. As shown in *Identification of Potentially Homosexual Loan Applicants*, when we compare our measured state-level

Significance

We propose a method to infer people's sexual orientation indirectly through gender disclosure of the borrower and coborrower in a mortgage. Furthermore, we examine lending practices toward same-sex borrowers and its spillover effects. We attempt to extend the research on race/gender discrimination by systematically investigating the potentially different lending treatment toward same-sex borrowers. The data reveal that, compared with otherwise similar different-sex applicants, same-sex applicants are 73.12% more likely to be denied, and they tend to be charged up to 0.2% higher fees/interest. Furthermore, neighborhoods' higher same-sex population density adversely affects both same-sex and different-sex borrowers' lending experiences. Our method might approximately measure the US homosexual population distribution up to the census tract level annually over decades.

Author contributions: L.G. and H.S. contributed equally to this work.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. S.L.R. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: L.G. (lgao@iastate.edu) or H.S. (hsun@iastate.edu).

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1903592116/-/DCSupplemental.

percentage of same-sex borrowers with the 2015 Gallup survey on the adult LGBT population, the correlation is remarkably high, ranging from 0.6 to over 0.8. Furthermore, Washington, DC, data show that, up to the census tract level, our measure matches well with the LGBT population percentage survey released by the US Census Bureau. As the massive HMDA data are compiled and released annually to the public, we can almost reliably measure the demographics of sexual orientation in the United States at up to the census tract level every year, in contrast to a 10-y release of similar information drawn from the census data. To the best of our knowledge, no other dataset allows researchers to measure and track homosexual representation annually at the local level over decades.

Despite the lack of literature on disparate lending outcomes by sexual orientation, there is some research in other markets/fields that may shed light on the possible existence of such lending inequality in the mortgage market. Using experimental data of faked resumes with listed sexual orientation, research shows that lesbians are less likely to be hired (22). A questionnaire-based study (23) finds that the majority of the respondents had experienced anti-LGBT discrimination. Furthermore, LGBT youth experience more depression and a higher tendency to self-harm or to commit suicide, and perceived discrimination is a significant contributor to the observed pattern (24). It is also documented that physicians with different political–ideological orientations tend to provide different care to LGBT patients (25).

Our study begins with an investigation of whether mortgage lenders are more likely to deny same-sex applicants than different-sex applicants with similar backgrounds. We first apply settings that are similar to those in previously studied racial- and gender-based discrimination (5, 6) to our research question on potentially different lending practices toward same-sex applicants. This way, our results are as robust as those from previous literature, and the potential discrimination level is contextualized.

To facilitate cross-comparison, we start our underwriting analysis by using the classical Boston Fed data, which includes an extensive list of property, neighborhood, borrower, and lender characteristics for a random sample of borrowers in Boston in 1990. This dataset has been widely used by racial discrimination literature (5, 6). The result reveals that same-sex applicants are about 73.12% more likely to be denied a mortgage application than are different-sex applicants with similar characteristics, which leads to an ∼8% lower gross approval rate. In contrast, prior research (5) using the same data and a similar methodology finds that, compared with similar White applicants, minority applicants are about 40% more likely to be denied a mortgage.

The Boston Fed data, despite its detailed coverage of underwriting information used by lenders, only represents a sample of loan applicants in Boston back in 1990. To gain some insight into the potentially unequal lending practices against same-sex borrowers to a bigger scope and more recent years, we expand the study with a national sample of data from HMDA from 1990 to 2015. We report qualitatively similar results on loan approval when we use these data. In particular, we find that same-sex applicants have ∼3% lower approval rate nationwide than that of different-sex applicants in the past three decades, after controlling for the observable loan and borrower characteristics. Furthermore, the pattern of lower loan approval rate to same-sex loan applicants is persistent over time.

Although the findings using HMDA data are consistent with lending discrimination, we acknowledge that there might be other potential missing variables in the tests and HMDA has a serious limitation because it has very limited coverage of borrower and loan characteristics, which are crucial for a credible underwriting study. Even after controlling for all essential underwriting variables, lower loan approval rates alone do not establish a complete and convincing case for lending discrimination. As Gary Becker (26), while commenting on minority lending discrimination, points out in his 1993 Nobel Prize speech, "If banks discriminate against minority applicants, they should earn greater profits on the loans made to them than on those to Whites." Inspired by this observation, and to address the concern on the limited information covered in HMDA, we further merge the HMDA data with Fannie Mae Single-Family Loan Performance data, which provide more information on borrower characteristics, mortgage terms, and performance. Our results on loan cost and performance analyses show that, after controlling for loan and borrower characteristics, lenders charge 0.02–0.2% higher interest to same-sex borrowers. Based on our inferred loan balance of $43 billion from same-sex borrowers in the United States, this is equivalent to about $8.6 million to $86 million more interest/fees paid by same-sex borrowers nationwide every year. Furthermore, the results seem to reveal that borrowers' same-sex status is associated with lower prepayment risk but not associated with higher default risk.

Another finding from this study is a spillover effect. We use a county's (or tract's) percentage of the same-sex population each year as a proxy for the attractiveness of the neighborhood to LGBT people. We find that when the share of a neighborhood's same-sex population increases, both same-sex and different-sex borrowers residing in the same neighborhood seem to experience more unfavorable lending outcomes overall. The findings should raise enough concerns to warrant further investigations on this topic by researchers, relevant government agencies, and other interested groups.

Although this study focuses on potential disparate lending practices against same-sex borrowers, our proposed method of measuring the same-sex population enables demographers and other social scientists to explore a wide range of LGBT-related research topics that are otherwise difficult to conduct due to the previous measurement barrier. We proceed as follows: In *Data* and *Identification of Potentially Homosexual Loan Applicants*, we describe the data and our research design; in *Results*, we summarize our results; and, in *Conclusion*, we conclude.

Data

There have been three influential data sources regarding the homosexual population in literature since the 1990s. Two of the datasets come from surveys from the National Opinion Research Center that rely on self-identification and voluntary participation: the 1992 National Health and Social Life Survey (NHSLS) and the ongoing General Social Survey (GSS). The sample sizes are relatively small, and survey data might not be consistent across the years. For example, there are only 3,432 observations in the NHSLS sample (27, 28), and the ongoing GSS covers only up to 450 gay/lesbian/bisexual individuals from 1988 to 1996 (29).

A systematic source for nationwide same-sex population information is the decennial census. These data are released every 10 y. There are also miscoding (30) and data inconsistency issues in the data because the US Census Bureau changes the survey forms over time. Some additional data sources collected for health surveys and other specialized purposes might be able to provide insight from different angles, depending on the variables that they cover. Black et al. (31) provide a more comprehensive review of the datasets. Overall, although these datasets can provide some information about same-sex couples, none of the available individual-level data can be used to match individual mortgage lending and serve in our research setting.

In this paper, we propose a method to infer homosexuality indirectly by using mortgage data. In this section, we first introduce the relevant datasets but defer our discussion of the inference procedure until *Identification of Potentially Homosexual Loan Applicants*. HMDA, a federal law enacted in 1975, requires certain financial institutions to provide mortgage data to the public. HMDA data cover information such as the applicant's race, gender, income, loan purpose, loan approval, and related categories. The details of the HMDA data can be found at the website of the Federal Deposit Insurance Corporation (32). We collect HMDA national data from 1990 to 2015. Because our HMDA data cover the entire nation, resulting in a very large sample size, we draw a 20% random sample from the full data.

One significant limitation of the original HMDA data are their lack of coverage of some important information on loan and applicants, which is crucial to lenders' approval decisions and any credible investigation of lending discrimination. In response, in 1990, the Federal Reserve Bank of Boston surveyed a sample of lenders in the Boston Metropolitan Statistical Area, collecting detailed information on the borrowers' financial and credit strength, employment, and other demographic and property characteristics. The data were later used to generate the famous Boston Fed Study (5), which is the classic research on minority lending discrimination. The Boston Fed data have been made public by the authors and have since been used by numerous researchers.

To investigate loan cost and performance, we merge HMDA's full sample with Fannie Mae Loan Performance data. This public-use Fannie Mae data include all 30-y, fully amortizing, full documentation, single-family, and conventional fixed-rate mortgages that Fannie Mae has purchased since 2000. In addition to loan performance records, the Fannie Mae data provide crucial information that is not reported in HMDA, such as borrowers' credit scores, contractual interest rates, loan-to-value ratios, debt-to-income ratios, and mortgage insurance premiums, if any.

In *SI Appendix*, we provide additional details on data sources, features, and the procedure for data merging and cleansing. A list of the key variables is reported in *SI Appendix*, Table S1, and their summary statistics are reported in *SI Appendix*, Table S2.

Identification of Potentially Homosexual Loan Applicants

Because sexual orientation is not a required disclosure item, it is impossible to identify homosexual loan applicants directly from the available mortgage data. A fair concern, therefore, is how to attribute any findings to homosexual discrimination if we cannot identify homosexual applicants. The definition of "discrimination" from the *Cambridge Dictionaries Online* is, "treatment or consideration of, or making a distinction in favor of or against, a person or thing based on the group, class, or category to which that person or thing is perceived to belong to rather than on individual merit." Hence, it is the external perception, instead of the authenticity of the underlying perception, that matters in discrimination identification. In this regard, research shows that it is very likely that perceived discrimination is a contributor to emotional distress among LGBT 9th- to 12th-grade students (24). Fortunately, measuring such external perception is possible.

To comply with FHA and ECOA, disclosing the gender information of the applicant and the coapplicant, if applicable, is mandatory. As gender discrimination is illegal, we anticipate that, compared with some other surveys, the reporting accuracy of gender information will be high in mortgage data. Therefore, based on HMDA data, whenever we observe a joint presence of a loan applicant and coapplicant with the same gender, we assign value 1 on the dummy variable, *Same-Sex*.

As identifying same-sex borrowers requires gender information for both the main applicant and coapplicant, we drop the HMDA observations that include the main applicant only. Processing this way helps us to avoid the attenuation bias in the estimated population percentage of same-sex borrowers, as observations without a coapplicant can include both same-sex and different-sex applicants, making potential sexual orientation undistinguishable to researchers.

To verify that the indirectly identifying homosexual loan applicants strategy is mostly valid, we calculate the percentage of same-sex pairs out of all identifiable applicant-pairs from the full HMDA dataset in 2015 and for each state and Washington, DC, and we compare this percentage with the Gallup estimate of the state-level adult LGBT population percentage of the same year (33). We augment the Gallup survey with a 2016 Williams Institute study (34) that estimates the state-level transgender population percentage. We then subtract the estimated percentage of transgender people from the LGBT population to infer the state-level LGB population percentage. Our prior expectation is that, because the *Same-Sex* measure directly ties to the gay and lesbian population, the correlation should be higher when we use the Gallup–Williams LGB estimate instead of the Gallup LGBT one. The correlation results are reported in Table 1.

Several important findings emerge. First, using the full sample of HMDA data in 2015, the correlation between our measure and the Gallup LGBT estimate is 0.8063 when Washington, DC, is included in addition to the 50 states in the United States. In contrast, it drops to 0.6186 when we remove Washington, DC. The higher correlation when Washington, DC, is included suggests that cities tend to have a higher representation of same-sex couples; hence, the accuracy of the same-sex population percentage estimate tends to improve.

Second, the same-sex population percentage is best estimated when we use the refinance subsample of the HMDA data. In this case, the correlation is further increased to 0.8469 and 0.6587, respectively, depending on whether Washington, DC, is included. The correlation becomes lower when we use home purchase loans. We believe that this finding makes sense, as Gallup surveys are designed to reflect the static LGBT representation of current residents (the same as borrowers who refinance their existing loans), not the dynamic pattern of people who plan to move into a neighborhood (the same as borrowers who purchase a home). In contrast, using home improvement mortgages is not very informative for same-sex population estimation, suggesting that, compared with those of different-sex couples, same-sex borrowers' decisions about home improvement can be much more idiosyncratic, probably due to their unique household structure.

Finally, it is surprising but interesting to see that HMDA-based same-sex percentages have higher correlations with the Gallup LGBT percentages than with the presumably better-suited Gallup-Williams LGB percentages. This finding suggests that our HMDA-based measure manages to capture the transgender population as well. For example, it is certainly possible that a male and a transgender female couple is identified as a same-sex couple in HMDA because the transgender female applicant still reports her biological sex of male when applying for a loan.

To help readers to visualize the matching quality of same-sex population, in Fig. 1, we present a scatter plot of HMDA-implied same-sex percentage (based on the refinance subsample) and Gallup LGBT percentage in the 50 states of the United States.

Table 1. Correlation between HMDA-implied and survey-based LGBT percentages (state level in 2015)

Location and purpose	Gallup LGBT	Gallup–Williams LGB
US 50 states and Washington, DC		
Refinance	0.8469	0.7281
Purchase	0.7297	0.6854
Improvement	0.6248	0.4488
All included	0.8063	0.7066
US 50 states only		
Refinance	0.6587	0.6168
Purchase	0.6132	0.5772
Improvement	0.2551	0.1525
All included	0.6186	0.5699

Note: This table presents the correlation of HMDA-implied LGBT percentages and two survey-based LGBT percentages, Gallup LGBT estimate and Gallup–Williams LGB estimate.

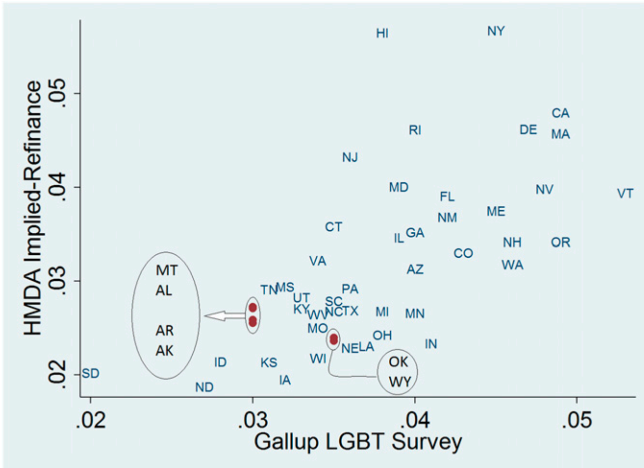

Fig. 1. Percentage of 2015 state-level same-sex population from HMDA and Gallup (correlation: 0.66). Gallup source: ref. 33.

Given the high correlation between these two aggregate estimates, we are confident that our measure covers a significant portion of homosexual loan applicants.

As the HMDA dataset allows us to map each observation up to the census tract level, we further choose Washington, DC, and calculate the percentage of same-sex applicants in each census tract, using post-2010 data. We then compare our measure with (*i*) an LGBT population survey from 2010 US Census, and (*ii*) other anecdotal evidence (35) related to LGBT communities in Washington, DC, such as Logan Circle and DuPont Circle, which are considered to have large LGBT populations in the area. We include the survey results of US Census data in Fig. 2, *Left*, and we plot the heat map of our estimated same-sex percentage in Fig. 2, *Right*. We can see that the pattern of the varying same-sex population density from our measure is consistent with both the results of the US Census and anecdotal evidence that Logan Circle and DuPont Circle are the commonly perceived as neighborhoods with concentrated homosexual population.

The summary statistics from *SI Appendix*, Table S2 reveal that, compared with different-sex couples, same-sex couples are, on average, younger and with lower credit scores but higher incomes. They are more likely to be the first-time home buyers and tend to borrow at a higher loan-to-income ratio, probably due to smaller down payments.

Although it seems that our proposed *Same-Sex* measure closely tracks the underlying homosexual population, we acknowledge that our measure has its flaws. For example, it is certainly possible that some of the identified pairs in our sample are merely family members. As a result, in *Results*, *Extended Analysis*, we conduct a series of robustness checks to address the concern of potential measurement errors.

Results

Loan Approval Analysis, Using Boston Fed Data. Before we conduct the loan approval analysis, we calculate the distribution of different-sex and same-sex applicants based on different loan types and programs. The results, calculated based on a 20% random sample of mortgages with a coapplicant, are reported in *SI Appendix*, Table S3. *SI Appendix*, Table S3 shows that applications filed by same-sex borrowers account for 4.03% of the total received. Conditional on the approved loans, however, this percentage drops to 3.75%. Given the total residential loan balance of \$11.45 trillion as of 2016, we infer that same-sex applicants borrow approximately \$43 billion. The approval rate for different-sex applicants is 82.74%, in contrast to only 76.82% for same-sex applicants. The lower approval rate for same-sex applicants is persistent among various loan purposes and programs, although the gaps are smaller for government-sponsored programs.

The lower raw approval rate is not enough to establish a case of potentially unfair treatment of same-sex applicants, as we do not control for other confounding factors that affect a lender's approval decision. As a result, based on Eq. **1** in *Materials and Methods*, we run a series of logit models with available controls from Boston Fed data. As discussed earlier, a significant advantage of Boston Fed data over HMDA is its detailed coverage of borrower and loan characteristics, including borrowers' credit and employment histories, educational backgrounds, loan-to-value ratios, mortgage expense-to-income ratios, etc. A consumer's mortgage lending approval and interest rate are affected by the applicant's neighborhood. For example, the cohort of the most recent purchase or refinancing is influential in predicting defaults (36). Before the subprime mortgage crisis, lenders targeted geographic areas with historically low rates of home ownership, where loan sellers identified many underserved consumers. To test for the potential neighborhood heterogeneity effect, and due to the narrow geographic coverage, we calculate the percentage of *Same-Sex* applicants at the census tract level (*LG_TractPct*), using the full sample of HMDA data in 1990, and include it in our analysis here.

We report our findings in Table 2. The model specification progresses as follows. As seen in column 1, we begin with a baseline model that includes only those controls that are also

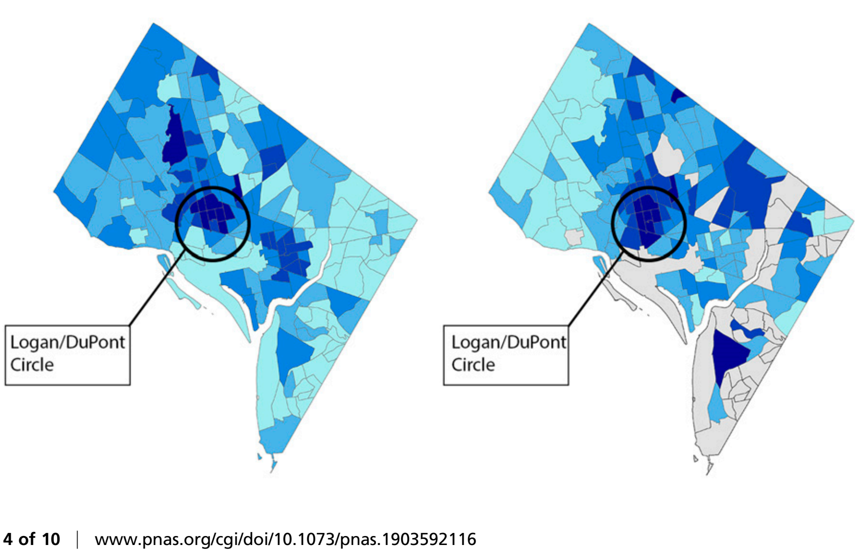

Fig. 2. *Left* shows the choropleth heat map of the population density of same-sex couples in Washington, DC, using 2010 census data. The light-to-dark color represents from 0.00 to 79.09 of same-sex couples per 1,000 households. *Right* shows the choropleth heat map of the percentage of same-sex borrowers, measured using post-2010 HMDA full data. The light-to-dark color represents from 1.89 to 29.53% of the inferred same-sex mortgage applicants. Gray color represents no data available in the area.

Table 2. Boston Fed data results

Variables	Logit (1)	Logit (2)	Logit (3)	Logit (4)	Logit (5)
Same-Sex	−0.4585	−1.0274***	−1.1878***	−1.2913**	−0.3855
	(0.3330)	(0.3885)	(0.4127)	(0.5478)	(0.7507)
[Average marginal effect]	[−0.0542]	[−0.0851]	[−0.0864]	[−0.0846]	[−0.0252]
LG_TractPct	−0.0181*	−0.0273*	−0.0262	−0.0330**	−0.0250
	(0.0093)	(0.0160)	(0.0175)	(0.0161)	(0.0172)
[Average marginal effect]	[−0.0021]	[−0.0023]	[−0.0019]	[−0.0022]	[−0.0016]
Same-Sex × *LG_TractPct*					−0.0642**
					(0.0256)
[Average marginal effect]					[−0.0042]
County fixed effects	Y	Y	Y	Y	Y
Lender fixed effects	N	N	Y	Y	Y
Key underwriting variable and lender portfolio interactions	N	N	N	Y	Y
N	2,316	2,316	2,316	2,316	2,316

Note: See *SI Appendix* for the full display of this table on other controls. SEs in parentheses are robust and clustered at the lender level. *$P < 0.1$, **$P < 0.05$, and ***$P < 0.01$.

available from HMDA and without lender fixed effects. We then include Boston Fed-only controls for additional loan and borrower characteristics in column 2. As seen in column 3, we further add lender fixed effects. In column 4, in addition to lender fixed effects, we allow lenders to put different underwriting weights on various loan, borrower, and lender characteristics. Ross and Yinger (6) emphasize the importance of controlling for these types of variations. Following their suggestion, we first identify a list of key underwriting variables (i.e., house expense-to-income ratio, total debt expense-to-income ratio, loan-to-value ratio, bankruptcy history, and borrowers' consumer and mortgage credit history). Also, we identify a group of key lender portfolio variables (i.e., the percentage of conventional loans sold to a secondary market, average loan size, average applicant's income, and average loan-to-income ratio). We then construct pairwise interaction terms for these variables and add them to our model in column 4. Finally, we add the interaction term between *Same-Sex* and *LG_TractPct* in column 5. Throughout all specifications, we also control for county fixed effects.

Progressing as outlined above allows us to investigate the impacts from the potentially confounding factors on our key *Same-Sex*–related coefficients. In particular, it enables us to understand better the possible implications of some key missing variables from HMDA, the data that are the basis for the expanded underwriting analysis later.

With HMDA controls only, the coefficient for *Same-Sex* in column 1 is negative (−0.4585, with an average marginal effect of −0.0542) but insignificant (with a SE of 0.3330). Interestingly, when we include additional controls from Boston Fed in column 2, the negative coefficient on *Same-Sex* becomes significant at 1%, and the associated average marginal effect increases to −0.0851, suggesting an even stronger adverse treatment of same-sex borrowers. The effect remains stable and significant when we add lender fixed effects, as seen in column 3, and adopt the more flexible underwriting model, as seen in column 4. In the meantime, we find some evidence that neighborhoods with a stronger presence of same-sex populations tend to have lower approval rate, on average. The coefficient on *LG_TractPct* is negative throughout all models, although it is insignificant (at 15% level) in columns 3 and 5. The average marginal effect is about −0.002 across models, suggesting that, when the same-sex population percentage in a census tract increases by 1%, the heterosexual applicant's loan approval rate declines by 0.2%.

Notably, when we further add the interaction term between *Same-Sex* and *LG_TractPct* in column 5, the main coefficient on *Same-Sex* becomes negative (an average marginal effect of −0.0252) but insignificant, and the interaction term is negative and significant at 5%, suggesting that same-sex couples who live in denser same-sex neighborhoods are more likely to be rejected. Coupled with the negative coefficient on *LG_TractPct*, column 5 suggests that when the same-sex population percentage in a census tract increases by 1%, the approval rate for same-sex applicants further declines by about 0.58% than the base group.

Although the lower approval rate of same-sex applicants is persistent, as our model in Table 2 progresses, we note the instability of the key coefficients of *Same-Sex*–related variables when additional key underwriting variables missing from HMDA are added (variables available from Boston Fed data). This is a clear indication that these coefficients tend to be biased with HMDA controls only. The specific direction of the bias, however, is inconclusive.

Finally, following Boston Fed study (5), we run a full logit regression (i.e., column 5 of Table 2, but excluding *Same-Sex*–related control variables) on non–same-sex observations. Then, for the same-sex applicants, we plug their attributes into the model, pretending that they are different-sex, and compare the average predicted denial probability with the actual denial rate observed in the data. Compared with otherwise similar different-sex applicants, same-sex applicants are 73.12% more likely to be rejected. The results are reported in Table 3.

Loan Approval Analysis, Using HMDA Data. To expand our investigation to a broader scope and more recent years, we run a series of logit and linear probability models with available controls from HMDA and census data, keeping in mind that HMDA data lack extensive coverage on loan and borrower characteristics and hence may bias the estimated coefficients on the variables of interest. Although a 20% random sample is chosen due

Table 3. Plug-in of *Same-Sex* attributes to non–*Same-Sex* regression: Boston Fed data

Characteristics and experience	Rates
Actual denial rate for *Same-Sex* borrowers in sample (n = 69)	20.29%
Predicted denial rate for *Same-Sex* borrowers with their characteristics but non–*Same-Sex* experience	11.72%
Odds ratio	1.7312

Note: This table presents the predicted denial rate and actual denial rate for *Same-Sex* borrowers and odds ratio.

to the extremely large data size, the small percentage of same-sex couples in the population means that we potentially are excluding a great deal of information on same-sex borrowers. Thanks to the property of unbiased estimation from the exogenous stratification, in the following HMDA-based loan approval analysis, we use the full sample of identifiable same-sex observations and pool them with our 20% random sample on different-sex couples, unless stated otherwise.

We report the results in Table 4. Column 1 contains the baseline logistic regression result. The coefficient for same-sex applicants is −0.1972 (1% level of significance). The corresponding average marginal effect indicates that the chances are 2.72% lower for same-sex applicants to be approved for a mortgage loan after controlling for an entire set of basic mortgage loan control variables, such as property type, natural log of applicants' income, loan purpose, and so on.

In column 3, we include county-level percentages of same-sex applicants measured each year (*LG_CountyPct*) and their interaction with *Same-Sex* dummy as control variables, and the results become more negative for the main effects. We also report the average marginal effect in column 4. The *Same-Sex* coefficient is −0.3708 at the 1% significance level. The margin coefficient of −0.0511 suggests that, relative to an otherwise similar different-sex applicant, the approval rate for the base group same-sex applicant is 5.11% lower, on average.

The coefficient of *LG_CountyPct* is −0.0650 at the 1% significance level. Its margin of −0.0090 indicates that, when the share of a county's same-sex population increases by 1%, the loan approval rate to different-sex applicants who reside in that county drops by 0.9%. The interaction term between *Same-Sex* dummy and *LG_CountyPct* is 0.0372 at the 1% significance level (with a corresponding margin of 0.0051). This finding seems to contradict our previous finding using Boston Fed data, which has a positive and significant coefficient on the interaction term. The opposite signs of the interaction term from the two datasets suggest that factors such as the differences in sample size, time frame, and available controls may affect the estimates. That said, a closer look reveals that there is no qualitative inconsistency here. The reason is that, as a neighborhood's same-sex population density increases, the gross incremental effect to the same-sex applicant shall be the sum of the marginal effects from *LG_CountyPct* (−0.9%) and its interaction with *Same-Sex* dummy (0.51%). This combined effect turns out to be −0.39%, which is close to the estimated gross effect of −0.58% (although measured at census tract level) from Boston Fed data. Our findings from both datasets seem to suggest a two-sided spillover effect when neighborhood LGBT population density changes. That is, in more LGBT-populated neighborhoods, both same-sex and different-sex borrowers residing in the same neighborhood experience lower approval rates. Given its robustness with extensive controls of neighborhood/time fixed effects, as to be seen in the subsequent analysis, this finding does not seem to be driven by the underlying unobservable characteristics of the neighborhoods.

Although the results in columns 1 and 3 are consistent with a pattern of lending discrimination, several concerns warrant more in-depth analysis. First, it is possible that much of the "observed" discrimination is driven by the heterogeneous search patterns when borrowers choose their lenders. Thus, it is worthwhile to look at how major coefficients change when we further control for lender fixed effects. Also see *SI Appendix*, section 3.2, for additional discussion of lender level analysis that also helps to address the issue of heterogeneous searching patterns from loan applicants. Second, the use of county fixed effects may disguise much geographic heterogeneity regarding where homosexual couples reside within a county. Although in columns 1 and 3 we control for census tract demographics in addition to county fixed effects, a more convincing case can be made if we directly control for census tract fixed effects. The only caveat of doing so is that the boundary definition of a census tract can change over time. As HMDA is not always using the most up-to-date tract definition, this can make the specification of a time-unvarying tract fixed effects problematic.

To address the above concerns, and due to the greatly expanded number of control variables, we use a linear probability model in the remaining analyses, as also presented in Table 4. In column 5, we begin with a baseline model that includes only the observed HMDA controls, census tract demographics, and year fixed effects. We then separately add the lender fixed effects in column 6, the census tract fixed effects in column 7, and both lender and tract fixed effects in column 8. Finally, to address the concern of the changes in tract definition over time, we add lender and tract-by-year fixed effects in column 9.

A valuable feature of the linear probability model is that it allows for the simple probabilistic interpretation of coefficients, and, as can be seen, the findings are qualitatively unchanged from column 3. We once again find significant evidence of a lower approval rate for same-sex applicants and the two-sided

Table 4. HMDA-based national loan approval (1990–2015)

Variables	Logit (1)	Average marginal effect (1)	Logit (2)	Average marginal effect (2)	Linear probability (3)	Linear probability (4)	Linear probability (5)	Linear probability (6)	Linear probability (7)
Same-Sex	−0.1972***	−0.0272	−0.3708***	−0.0511	−0.0590***	−0.0362***	−0.0489***	−0.0360***	−0.0350***
	(0.0248)		(0.0277)		(0.0049)	(0.0026)	(0.0042)	(0.0024)	(0.0022)
LG_CountyPct			−0.0650***	−0.0090	−0.0075***	−0.0023***	−0.0093***	−0.0039***	N/A
			(0.0095)		(0.0012)	(0.0004)	(0.0019)	(0.0008)	
Same-Sex × *LG_CountyPct*			0.0372***	0.0051	0.0062***	0.0017***	0.0042***	0.0016***	0.0016***
			(0.0049)		(0.0008)	(0.0004)	(0.0008)	(0.0004)	(0.0003)
Tract demographics	Y		Y		Y	Y	Y	Y	N/A
Year fixed effects	Y		Y		Y	Y	Y	Y	N/A
County fixed effects	Y		Y		N	N	N	N	N
Lender fixed effects	N		N		N	Y	N	Y	Y
Tract fixed effects	N		N		N	N	Y	Y	N/A
Tract × year fixed effects	N		N		N	N	N	N	Y
N	33,664,547		33,664,547		33,664,547	33,664,547	33,664,547	33,664,547	33,664,547
Adjusted R^2	N/A		N/A		0.0665	0.1914	0.0825	0.1973	0.2073

Note: See *SI Appendix* for the full display of this table on other controls. These results are based on a 100% *Same-Sex* sample plus a 20% heterosexual sample. SEs in parentheses are robust and clustered at the lender level. ***$P < 0.01$.

spillover effect, as discussed above. It is also clear, however, when comparing columns 5 and 6 that the potential discrimination effect drops significantly after adding lender fixed effects. The main coefficient on *Same-Sex* drops from −0.0590 to −0.0362, and the coefficient of the interaction term between *Same-Sex* and *LG_CountyPct* drops from 0.0062 to 0.0017. This suggests that the borrowers' self-selection on lenders can drive a large portion of across-lender effects.

Recall that same-sex couples tend to be younger and have lower credit scores. Imagine that, for some reason, many of them choose to borrow from some particular lenders. Although HMDA data do not contain information on an applicant's age and credit score, by adding lender fixed effects, we can at least partially control for the impacts driven by these unobserved borrower characteristics if there is significant self-selection by loan applicants. This observation motivates us to look further at the sorting patterns among borrowers, as well as to investigate the discrimination effect at the individual-lender level. We present these further analyses in *SI Appendix*, section 3.2.

As previously discussed, if same-sex couples exhibit extreme sorting to neighborhoods within a county, then some effects can be driven by the unobserved neighborhood characteristics rather than by discrimination. The results in column 7 provide supporting evidence of this. For example, compared with the results in column 5, when adding census tract fixed effects, the coefficients on both *Same-Sex* and its interaction with *LG_CountyPct* decline in magnitude, although to a lesser extent than when we add lender fixed effects. Finally, when we add tract-by-year fixed effects in column 9, the little change in the key coefficients suggests that our major findings on potential discrimination effect are unlikely affected by the time-varying census tract definitions.

Another concern about our finding of the lower approval rate for same-sex applicants is that the result might be driven by observations of early years or a few extreme years. We run the regressions year by year, using the full sample of national HMDA data, and report the estimates and confidence interval in *SI Appendix*, Fig. S1. The pattern of lower approval rate to same-sex applicants is persistent over time, except in 1996 and 2000, in which the negative coefficient is not significantly different from zero at 5%. Surprisingly, the extent of discrimination seems unmitigated in the last decade, when the Democratic Party was in power.

Our HMDA-based analysis has its limitations. One serious concern is the unbalanced representation of same-sex borrowers in the sample. For example, as can be seen in *SI Appendix*, Table S2 and the discussion in *Identification of Potentially Homosexual Loan Applicants*, same-sex couples tend to be younger, with lower credit scores and higher incomes. They are more likely to be first-time home buyers and tend to borrow with higher loan-to-income ratios. Unfortunately, many of these characteristics are missing from the HMDA data. A good example is credit score. As a proxy for borrower's risk, we expect that, ceteris paribus, lenders will reject more applications from borrowers with lower credit scores, which has nothing to do with discrimination.

Suppose we find the coefficient of *Same-Sex* to be negatively significant when using HMDA data. As credit score is unavailable from HMDA (and, hence, is not included in the regression), we do not know whether the observed lower approval rate for same-sex borrowers is due to lending discrimination or the fact that we fail to include credit score as a control variable. The same problem also applies to the interaction of *Same-Sex* with *LG_CountyPct*, a key term that is used to measure the spillover effect. In *SI Appendix*, section 3.1, we report on a simple test on the conditional correlation (conditional on *Same-Sex* and *LG_CountyPct*) between *Same-Sex* × *LG_CountyPct* and some key variables observed in HMDA. The test result rejects the null hypothesis of zero conditional correlation at 1%, suggesting another balance failure.

To investigate the property of unbalanced representation in *Same-Sex*–related coefficients, we conduct a series of linear probability regressions by gradually adding more HMDA controls to observe their impacts on the key coefficients. We begin with an extremely basic approval model that includes only *Same-Sex*, *LG_TractPct*, and census tract and year fixed effects as controls. We then gradually add loan attributes, borrower attributes, census tract demographics, lender fixed effects, and, finally, lender tract fixed effects to the model. We report the results in *SI Appendix*, Table S4. Compared with what is seen in the baseline model, the addition of loan characteristics leads to a more negative *Same-Sex* coefficient. The size of this estimated "discrimination" effect shrinks when borrower and census demographic controls are added, and it further declines (but remains economically and statistically significant) when we add lender fixed effects, as seen in the last column of *SI Appendix*, Table S4. The major takeaways are twofold. On the one hand, there is clear evidence that the unbalanced representation of *Same-Sex* sample may affect the estimation accuracy of the size of the discrimination effect, as shown by the varying magnitudes of the *Same-Sex*–related coefficients. On the other hand, the qualitative finding of unfavorable treatment to same-sex borrowers and the spillover effect turns out to be rather persistent and robust to the omitted variables, which is consistent with a pattern of lending discrimination.

Loan Cost Analysis: Annual Percentage Rate Based. Measuring mortgage costs to borrowers can be tricky. One obvious choice is to use the contractual rate. Nevertheless, we are aware that the contractual rate is often not an accurate measure of the effective borrowing cost. Unlike the annual percentage rate (APR), the contractual rate ignores the impact of loan fees and other closing costs charged by the lender. As a simple example, consider a 30-y fixed-rate mortgage for $200,000. Suppose a lender offers a 4.5% contractual rate with 2% upfront loan fees. The terms observed from the merged HMDA–Fannie Mae data will have a loan amount of $200,000 and a contractual rate of 4.5%, and the monthly mortgage payment will be $1,013.37. Nevertheless, due to the loan fees, the lender will charge an extra $4,000 upon closing. Assuming that the borrower chooses to amortize the fees over the loan term, the actual monthly payment will be $1,033.64, which is based on the principal value of $204,000. In other words, this borrower receives only $200,000 from the lender, but his or her mortgage payment will be calculated as if he or she had borrowed $204,000. The effective borrowing cost (hence the APR) is 4.67%, which is 0.17% higher than its contractual rate. This example demonstrates why the contractual rate often underestimates the true financing cost. As a result, we focus our analysis here on APRs.

Under HMDA, since 2004, a rate spread for a loan must be reported if it is above a certain threshold. Between January 2004 and September 2010, the rate spread was defined as the difference between the APR on a loan and the prevalent rate for Treasury securities of comparable maturity. HMDA mandates disclosure of such a rate spread if it is at least 3% for a loan secured by a first lien. Due to this high threshold, only 3.89% of the loans in our usable sample have reported rate spreads. In October 2010, HMDA changed its definition of rate spread to the difference between a loan's APR and a survey-based estimate of prevalent APR (instead of Treasury rate) for comparable loans. Given this new definition, the disclosure is required if a rate spread is above 1.5%. The new disclosure threshold is even higher than before, as only 1.01% of the loans in our sample are above this threshold and, hence, have reported spreads.

One data restriction is that our sample is censored below the threshold that triggers the rate spread disclosure. We, therefore, adopt a Tobit model to address the data censoring issue. Due to the 2010 change to the reporting threshold and the definition of rate spread, we split our sample into periods before and after 2010, and report the Tobit regression results in Table 5. Here, we drop the loans originated in the last quarter of 2009 as it is the

Table 5. Tobit regression on rate spread

Variables	Model 1 (pre-2010)	Model 2 (pre-2010)	Model 3 (pre-2010)	Model 4 (post-2010)	Model 5 (post-2010)	Model 6 (post-2010)
Same-Sex	0.1718***	0.1722***	0.1902***	0.0452***	0.0455***	0.1032***
	(0.0113)	(0.0113)	(0.0155)	(0.0096)	(0.0096)	(0.0142)
LG_CountyPct		−0.0027	−0.0024		−0.0018	−0.0008
		(0.0043)	(0.0043)		(0.0034)	(0.0034)
Same-Sex × *LG_CountyPct*			−0.0036*			−0.0128***
			(0.0026)			(0.0028)
Census tract demographic controls	Y	Y	Y	Y	Y	Y
Loan month fixed effects	Y	Y	Y	Y	Y	Y
Lender × County fixed effects	Y	Y	Y	Y	Y	Y
N	176,502	176,502	176,502	237,131	237,131	237,131

Note: See *SI Appendix* for the full display of this table on other controls. This table reports Tobit regression results. SEs in parentheses are robust and clustered at the lender level. **P* < 0.1 and ****P* < 0.01.

transitional period of the changing reporting rules. Models 1 and 4 correspond to the baseline model for pre- and post-2010 periods separately, and models 3 and 6 are full models after controlling for the interaction term between *Same-Sex* and county-level *Same-Sex* percentage variable and additional fixed effects for the corresponding periods. Following prior literature on loan cost analysis (8), we include the combined loan to value ratio as a series of dummy variables for below 0.6, 0.6–0.8, 0.8–0.85, 0.85–0.9, 0.9–0.95, 0.95–1, and above throughout all models. We add dummy variables for debt-to-income ratio around the threshold of 0.36 with bins as small as 0.03. We also add dummy variables for borrower and coborrower credit score separately for below 600, above 820, and in 20-point bins otherwise.

The results show that same-sex borrowers are charged higher rate spreads. For example, the coefficient of *Same-Sex* borrowers for the full model before 2010 (model 3) is 0.1902 at a 1% significance level, which implies that the base group of *Same-Sex* borrowers on average pay 0.1902% more on their mortgages compared with otherwise similar borrowers The result is both statistically and economically significant. The same coefficient for post-2010 (model 6) is smaller at 0.1032, still highly significant.

As far as the spillover effects, the coefficient for *LG_CountyPct* is negative across all models, although insignificant. For same-sex borrowers who live in these counties, the combined effect is also negative. Overall, we find some weak evidence of a two-sided spillover effect on rate spread as well.

Loan Cost Analysis: Contractual-Rate Based. Although the Tobit model is suitable to handle censored data, the extremely high censoring rate in our sample and the strong normality assumption made by the Tobit model can still be of concern. In response, in this section, we report on our examination of the original contractual rate, which is available for all observations, thanks to Fannie Mae data.

Understanding that the contractual rate usually underestimates the effective financing cost, we report the ordinary least-squares (OLS) regression results when using contractual rate as our cost measure, as seen in Table 6. The model specifications are otherwise identical to those in Table 5. As the changing disclosure rule on rate spread is no longer relevant, we first report the full sample results in columns 1–3. Then, to facilitate comparisons with Table 5, we break our sample into the same pre- and post-2010 periods in columns 4 and 5. The findings from the contractual rate are qualitatively consistent with the previous APR-based rate spread analysis. For example, in column 1, the coefficient for *Same-Sex* is 0.0183 (SE, 0.0030), which suggests that *Same-Sex* borrowers, on average, pay 1.83 more basis points on the contractual rate compared with otherwise-similar different-sex borrowers. The pattern is persistent across different subperiods, suggesting that overpricing to same-sex borrowers is a common business practice. Finally, the larger coefficient for *Same-Sex* after 2010 in column 5 echoes the trend of decreasing approval rate for same-sex applicants since 2004, as shown in *SI Appendix*, Fig. S1.

Loan Performance Analysis: Mortgage Default. Our findings so far suggest that same-sex borrowers are more likely to be rejected when they apply for loans, and, conditional on being approved, they are charged higher financing costs by lenders, primarily through upfront fees. Although we have controlled for an extensive list of characteristics of both borrowers and their loans, it may still be plausible that there could be some unobserved characteristics that cause same-sex borrowers to be riskier and, therefore, more likely to default. In this case, the higher APRs charged to same-sex borrowers could simply reflect the premium of their higher default risk. To investigate whether the higher financing costs to same-sex borrowers can be justified by the default risk premium and, hence, a reflection of financial necessity, we look at the loan performance, using our merged HMDA–Fannie Mae sample.

We estimate logit models for the mortgage defaults. The dependent variable is a dummy variable that indicates when a mortgage becomes delinquent for at least 60 d within 5 y of its origination date. As the most recent performance updates in our data were from the end of 2015, we restrict our sample to loans that originated before 2010 to generate a 5-y observation window. The results are shown in *SI Appendix*, Table S5. Model 1 is our baseline model. In model 2, we add *LG_CountyPct*. In model 3, we add the interaction between *Same-Sex* and *LG_CountyPct*. Although there is no reason to associate subsequent loan performance with the original lender after loan characteristics are controlled for, we still include the annually measured lender's county-level market share measured and lender fixed effects in model 4. None of the coefficients of *Same-Sex* borrowers across all four models is significant. Hence, we feel comfortable saying that same-sex status exhibits no greater risk of default.

Loan Performance Analysis: Mortgage Prepayment. Another key risk factor to lenders is regarding prepayment. For example, when the market rate drops, borrowers tend to prepay the existing higher-rate loan through refinancing. These prepayments can be risky to lenders, as the lenders need to reinvest a large amount of unexpected cash inflows (due to prepayment) at the lower prevalent market rate. Even worse, in many cases, it is illegal for lenders to charge a prepayment penalty as a remedy because of the consumer protection laws. In this section, we report on our investigation of whether same-sex borrowers exhibit higher prepayment risk, which can potentially justify the higher interest rates and lower approval rates for their loans.

Table 6. Linear regression on contractual rate

Variables	OLS 1 (full sample)	OLS 2 (full sample)	OLS 3 (full sample)	OLS 4 (pre-2010)	OLS 5 (post-2010)
Same-Sex	0.0183***	0.0181***	0.0288***	0.0185**	0.0290***
	(0.0030)	(0.0030)	(0.0050)	(0.0084)	(0.0058)
LG_CountyPct		0.0012	0.0014*	0.0027**	0.0019**
		(0.0008)	(0.0008)	(0.0013)	(0.0008)
Same-Sex × *LG_CountyPct*			−0.0021**	0.0004	−0.0029***
			(0.0008)	(0.0013)	(0.0010)
Census tract demographic controls	Y	Y	Y	Y	Y
Loan month fixed effects	Y	Y	Y	Y	Y
Lender × County fixed effects	Y	Y	Y	Y	Y
N	420,175	420,175	420,175	176,502	237,131
Adjusted R^2	0.9097	0.9097	0.9097	0.7511	0.7511

Note: See *SI Appendix* for the full display of this table on other controls. This table presents linear regression results for contractual rates. SEs in parentheses are robust and clustered at the lender level. *$P < 0.1$, **$P < 0.05$, and ***$P < 0.01$.

Similar to what was discussed in *Loan Performance Analysis: Mortgage Default*, we create a dummy variable that equals 1 when a mortgage is prepaid within 5 y of its origination date and focus on the loans that originated between 2004 and 2010. Our model specifications are otherwise identical to the ones that we use in the loan default analysis. We report the results in *SI Appendix*, Table S6. The coefficient on *Same-Sex* is negatively significant at 1% in columns 1 and 2. It is worth noting that, in columns 3 and 4, the appropriate test for *Same-Sex* performance is an *F* test for the coefficients of *Same-Sex* and its interaction with *LG_CountyPct* (which are both negative) for joint significance. The corresponding *F* statistics of 19.05 ($P < 0.0001$) and 19.58 ($P < 0.0001$), respectively, present strong evidence that same-sex borrowers are less likely to prepay compared with their peers.

Extended Analysis. As seen in *SI Appendix*, we conduct a series of robustness checks. In particular, we examine the following:

i) To support the findings presented in *Loan Approval Analysis, Using HMDA Data*, we run a balanced representation test on *Same-Sex* × *LG_CountyPct* in HMDA.

ii) In the Boston Fed approval analysis, we use appropriate subsamples of observations to address the concern that same-sex borrowers are younger, on average; more likely to live in multifamily units than single-family units; and more likely to have a cosigner.

iii) In the HMDA approval analysis, for identified same-sex applicants, we drop the observations if the races of the main applicant and coapplicant are the same to rule out potential father/son (or brother/sister) types of relative pairs.

iv) In the cost analysis, we use a linear probability model to test whether same-sex borrowers are more likely to have a disclosed high APR spread. This treatment frees us from the data-censoring problem.

v) In the performance analysis, we use the Cox proportional hazard model to test whether, conditional on the subsample of loans in default, same-sex borrowers default sooner, on average, which is costlier to lenders.

In short, our key findings remain qualitatively unchanged across these robustness checks. In *SI Appendix*, we also provide a more in-depth discussion of the types of lending discrimination and discriminatory behaviors at the individual-lender level.

Conclusion

We propose a method to infer a borrower's sexual orientation indirectly without self-identification and examine potentially unequal lending practice to same-sex borrowers and its spillover effects. The results indicate that same-sex mortgage applicants in the United States are more likely to be denied than are different-sex applicants with similar characteristics. We then check the cost of the approved loans, and our results show that lenders tend to charge same-sex borrowers higher financing costs, primarily through upfront fees. Furthermore, our investigation on mortgage performance reveals that same-sex applicants are less likely to prepay mortgages and are no more likely to default than their peers, indicating that they are less risky to lenders. Given the absence of evidence that suggests that same-sex status is a reliable signal for loan underperformance, potential disparate lending practices against sexual orientation might exist in the mortgage market.

Our analysis also provides evidence of spillover effects. That is, holding other factors constant, when a neighborhood's same-sex population density increases, both same-sex and different-sex applicants seem to experience either lower approval rate or higher financing cost, or both.

Our research might be subject to omitted variable bias, data errors, and endogeneity problems in several of the explanatory variables, especially in our HDMA-based loan approval analysis. We acknowledge the potential overstatement of the discriminations and constraints of the limited data items available to test, and we further check the loan performance after we check the approval rate of the mortgage applications. We also try to address these concerns with both different econometrics specifications and robustness tests for all of the helpful additional data we can find. For example, the original HMDA data suggest that differential treatment by sexual orientation is occurring in the mortgage market, but important confounding variables related to both loan approval and sexual orientation are missing from the HMDA data. To cross-validate the reliability of our findings, we employ Boston Fed data that cover in detail the lenders' information, including essentially all of the information the lenders use in their decision-making process, and find that same-sex status still plays a significant role in mortgage-lending decisions. The bottom line is that our results could still be refuted with potential evidence that the homosexual–heterosexual differences in loan approval and loan performance can be entirely explained by nondiscriminatory differences in business-justifiable underwriting standards across lenders.

Although it is still premature to conclude the existence of lending discrimination against sexual orientation, we believe the documented findings in this study should raise enough concern to warrant further investigations on this topic by researchers, relevant government agencies, and other interested groups.

In addition to its original contribution to the literature on discrimination studies, this paper sheds light on a broad group. For demographers, economists, and other social scientists, we propose a reliable and low-cost method to approximately measure the sexual orientation of US households at the local level annually over decades. This method is in contrast to alternative ways, such as the 10-y release of similar information from the US Census survey. For policymakers, lawmakers, and researchers, as the current federal fair lending acts do not explicitly list sexual orientation as a prohibitory class, our findings have direct implications for the urgency of protecting the LGBT community regarding fair credit accessibility. Furthermore, for credit monitoring agencies, such as the Consumer Financial Protection Bureau, and given its data advantage, we provide a possible approach to measure and investigate sexual orientation-based disparate lending practices.

Materials and Methods

There is a large body of literature on lending discrimination, although the research focuses primarily on racial/gender-based discrimination. To facilitate comparison with other studies and for consistency, we adopt a similar methodology to that used in prior classic studies (5, 6).

We first investigate lenders' loan approval decisions, using both Boston Fed and HMDA data. Our baseline model is a logistic regression (and linear probability model) with the following specifications:

$$\begin{aligned} Approve = {} & \beta_0 + \beta_1 \mathit{Same\text{-}Sex} + \beta_2 \mathit{Borrower\ Characteristics} \\ & + \beta_3 \mathit{Loan\ Characteristics} + \beta_4 \mathit{Census\ Tract\ Demographic\ Controls} \\ & + \mathit{Fixed\ Effects} + \mathit{error\ term}. \end{aligned} \quad [1]$$

To examine the potential spillover effect, we further expand our analysis based on Eq. **1** by adding *LG_CountyPct*, which measures the percentage of same-sex applicants annually out of the total paired applicants. We use it as a measure of a neighborhood's prevalent LGBT population density. We also add the $\mathit{Same\text{-}Sex} \times \mathit{LG\_CountyPct}$ interaction term as well as various lender/location fixed effects to account for varied lenders' underwriting models in a systematic way. Using similar model specifications, we also examine the potential differential outcomes for same-sex borrowers on loan cost and performance.

The key challenge that scholars face as they try to understand lending discrimination is that different lenders may use different underwriting standards (6). As a result, it is possible that, at the individual-lender level, there is a common underwriting standard for everybody and, therefore, no "discrimination." Because applicants might be sorted disproportionately to different lenders, however, the sorting might lead to the case that certain types of applicants are more (or less) likely to have their applications approved. To address this concern, we follow the methodology proposed in the lending discrimination literature (6) and examine several variations of lenders' underwriting models. In *SI Appendix*, we provide a more in-depth discussion of and justification for various model specifications that evolved from our baseline models and correspond to different datasets.

ACKNOWLEDGMENTS. We thank James Brown, Ioannis Floros, Dermot Hayes, Youngme Seo, three anonymous reviewers, and the editor for their insightful comments and suggestions. We also thank participants at Central University of Finance and Economics, Iowa State University, Shandong University, Shandong University of Finance and Economics, the 2017 China International Conference in Finance, the 2017 Financial Management Association Annual Conference, and the 2018 American Economic Association Annual Conference for their helpful comments. The authors are responsible for all errors and omissions.

1. Board of Governors of the Federal Reserve System (2019) Mortgage debt outstanding. Available at https://www.federalreserve.gov/data/mortoutstand/current.htm. Accessed March 27, 2019.
2. TreasuryDirect (2016) The debt to the penny and who holds it. Available at https://treasurydirect.gov/govt/reports/pd/pd_debttothepenny.htm. Accessed March 27, 2019.
3. Black H, Schweitzer RL, Mandell L (1978) Discrimination in mortgage lending. *Am Econ Rev* 68:186–191.
4. Ladd HF (1982) Equal credit opportunity: Women and mortgage credit. *Am Econ Rev* 72:166–170.
5. Munnell AH, Tootell GMB, Browne LE, McEneaney J (1996) Mortgage lending in Boston: Interpreting HMDA data. *Am Econ Rev* 86:25–53.
6. Ross SL, Yinger J (2002) *The Color of Credit* (MIT, Cambridge, MA).
7. Cheng P, Lin ZG, Liu YC (2014) Racial discrepancy in mortgage interest rates. *J Real Estate Finance Econ* 51:1–20.
8. Bayer P, Ferreira F, Ross SL (2018) What drives racial and ethnic differences in high cost mortgages? The role of high risk lenders. *Rev Finance Stud* 31:175–205.
9. Ross SL, Turner MA, Godfrey E, Smith RR (2008) Mortgage lending in Chicago and Los Angeles: A paired testing study of the pre-application process. *J Urban Econ* 63: 902–919.
10. Hanson A, Hawley Z, Martin H, Liu B (2016) Discrimination in mortgage lending: Evidence from a correspondence experiment. *J Urban Econ* 92:48–65.
11. Ahmed AM, Hammarstedt M (2008) Discrimination in the rental housing market: A field experiment on the internet. *J Urban Econ* 64:362–372.
12. Ahmed AM, Andersson L, Hammarstedt M (2008) Are lesbians discriminated against in the rental housing market? Evidence from a correspondence testing experiment. *J Hous Econ* 17:234–238.
13. Ayres I (1991) Fair driving: Gender and race discrimination in retail car negotiations. *Harv Law Rev* 104:817–872.
14. Ayres I, Siegelman P (1995) Race and gender discrimination in bargaining for a new car. *Am Econ Rev* 85:304–321.
15. Goldberg PK (1996) Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey. *J Polit Econ* 104:622–654.
16. Morton FS, Zettelmeyer F, Silva-Risso J (2003) Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities? *Quant Mark Econ* 1:65–92.
17. Korent JR (2016) Toyota will compensate black and Asian borrowers to settle bias investigation. *The Los Angeles Times*. Available at https://www.latimes.com/business/la-fi-toyota-settlement-20160202-story.html. Accessed March 10, 2019.
18. US Equal Employment Opportunity Commission (2010) Walmart to pay more than $11.7 million to settle EEOC sex discrimination suit. Available at https://www.eeoc.gov/eeoc/newsroom/release/3-1-10.cfm. Accessed March 27, 2019.
19. Egelko B (2013) Costco to pay $8 million to settle gender-bias suit. *The SFGATE*. Available at https://www.sfgate.com/nation/article/Costco-to-pay-8-million-to-settle-gender-bias-5083273.php. Accessed March 10, 2019.
20. Johnson C (2016) EEOC reaches first-ever settlement in anti-gay bias case. *The Washington Blade*. Available at https://www.washingtonblade.com/2016/06/29/eeoc-reaches-first-ever-settlement-anti-gay-discrimination-case/. Accessed March 10, 2019.
21. Gupta S (2016) Walmart agrees to $7.5 million settlement in discrimination suit. *The CNN Business*. Available at https://money.cnn.com/2016/12/02/news/companies/walmart-discrimination-settlement/. Accessed March 10, 2019.
22. Weichselbaumer D (2003) Sexual orientation discrimination in hiring. *Labour Econ* 10: 629–642.
23. Schneider MS, Dimito A (2010) Factors influencing the career and academic choices of lesbian, gay, bisexual, and transgender people. *J Homosex* 57:1355–1369.
24. Almeida J, Johnson RM, Corliss HL, Molnar BE, Azrael D (2009) Emotional distress among LGBT youth: The influence of perceived discrimination based on sexual orientation. *J Youth Adolesc* 38:1001–1014.
25. Hersh ED, Goldenberg MN (2016) Democratic and Republican physicians provide different care on politicized health issues. *Proc Natl Acad Sci USA* 113:11811–11816.
26. Becker G (1993) Nobel lecture: The economic way of looking at behavior. *J Polit Econ* 101:385–409.
27. Michael RT, Gagnon JH, Laumann EO, Kolata G (1994) *Sex in America: A Definitive Survey* (Little, Brown, Boston).
28. Laumann EO, Gagnon JH, Michael RT, Michaels S (1994) *The Social Organization of Sexuality: Sexual Practices in the United States* (Univ Chicago Press, Chicago).
29. Black DA, Gates G, Sanders S, Taylor L (2000) Demographics of the gay and lesbian population in the United States: Evidence from available systematic data sources. *Demography* 37:139–154.
30. Black DA, Gates G, Sanders S, Taylor L (2007) *The Measurement of Same-Sex Unmarried Partner Couples in the 2000 US Census* (California Center for Population Research, Los Angeles).
31. Black DA, Sanders SG, Taylor LJ (2007) The economics of lesbian and gay families. *J Econ Perspect* 21:53–70.
32. Federal Deposit Insurance Corporation (2012) FDIC Law, Regulations, Related Acts. 6500–Consumer Protection. Available at https://www.fdic.gov/regulations/laws/rules/6500-3030.html#6500hmda1975. Accessed March 27, 2019.
33. Gates GJ (2017) Vermont leads states in LGBT identification. *Social and Policy Issues* (Gallup, Washington, DC). Available at https://news.gallup.com/poll/203513/vermont-leads-states-lgbt-identification.aspx. Accessed March 27, 2019.
34. Flores AR, Herman JL, Gates GJ, Brown NT (2016) How many adults identify as transgender in the United States? (The Williams Institute, Los Angeles).
35. Borderstan (2012) Where is DC's gay neighborhood in 2012? The winner is. . .. Available at https://www.borderstan.com/2012/05/30/the-winner-is-dcs-gay-neighborhood/. Accessed March 27, 2019.
36. Bayer P, Ferreira F, Ross SL (2016) The vulnerability of minority homeowners in the housing boom and bust. *Am Econ J Econ Policy* 8:1–30.