Consumer-Lending Discrimination in the FinTech Era*

Robert Bartlett
School of Law
UC Berkeley
rbartlett@berkeley.edu

Adair Morse
Haas School of Business
UC Berkeley & NBER
adair@berkeley.edu

Richard Stanton
Haas School of Business
UC Berkeley
rhstanton@berkeley.edu

Nancy Wallace
Haas School of Business
UC Berkeley
newallace@berkeley.edu

May 2019

Abstract

Discrimination in lending can occur either in face-to-face decisions or in algorithmic scoring. We provide a workable interpretation of the courts' *legitimate-business-necessity* defense of statistical discrimination. We then estimate the extent of racial/ethnic discrimination in the largest consumer-lending market using an identification afforded by the pricing of mortgage credit risk by Fannie Mae and Freddie Mac. We find that lenders charge Latinx/African-American borrowers 7.9 and 3.6 basis points more for purchase and refinance mortgages respectively, costing them $765M in aggregate per year in extra interest. FinTech algorithms also discriminate, but 40% less than face-to-face lenders. These results are consistent with both FinTech and non-FinTech lenders extracting monopoly rents in weaker competitive environments or profiling borrowers on low-shopping behavior. Such strategic pricing is not illegal per se, but under the law, it cannot result in discrimination. The lower levels of price discrimination by algorithms suggests that removing face-to-face interactions can reduce discrimination. Further silver linings emerge in the FinTech era: (1) Discrimination is declining; algorithmic lending may have increased competition or encouraged more shopping with the ease of platform applications. (2) We find that 0.74-1.3 million minority applications were rejected between 2009 and 2015 due to discrimination; however, FinTechs do not discriminate in loan approval.

JEL classification: G21, G28, G23, J14, K22, K23, R30
Keywords: Discrimination; FinTech; mortgages; credit scoring; algorithmic underwriting; big-data lending; platform loans; disparate impact; statistical discrimination; legitimate business necessity

*Financial support from the Fisher Center for Real Estate and Urban Economics is gratefully acknowledged. For helpful comments and suggestions, we thank Manuel Adelino, Sanjiv Das, Andreas Fuster, Andres Liberman, Manju Puri, Raghu Rau, Amit Seru, Ansgar Walther, Justin Wolfers, and seminar participants at U.C. Berkeley, Cornell, the University of Michigan, USC, NYU Stern, U.T. Austin, Texas A&M, Michigan State University, Chicago-Booth, Johns Hopkins University, Consumer Finance Protection Bureau, the George Washington-Federal Reserve Board Seminar Series on Household Finance, the Spring 2018 NBER Corporate Finance meeting, the 2018 NBER Law & Economics Summer Institute, the 2018 Annual Conference on Empirical Legal Studies, the 2017 Boston JOIM Conference, the 2017 Annual Research Conference of the Center for International Securities and Derivatives Markets, the 2018 NY Fed./NYU Stern Conference on Financial Intermediation, the 2018 American Law and Economics Association (ALEA) annual meeting, the 2018 Financial Intermediation Research Society (FIRS) annual meeting, the 2018 EFA meeting, the 2019 AEA meeting, the 2019 NBER conference on "Big Data: Long-Term Implications for Financial Markets and Firms," the 2019 Adam Smith Workshop in Corporate Finance, the 2019 FMA Wine Country Finance Conference, the Federal Reserve Board, and Freddie Mac.

I. Introduction

Algorithmic decision-making can reduce face-to-face discrimination in markets known to be prone to implicit and explicit biases. But it can also lead to inadvertent discrimination (Barocas & Selbst, 2016). The question of whether algorithmic decision-making promotes or inhibits impermissible discrimination is especially relevant in the context of consumer lending, given both the historical challenge of eliminating discrimination in this domain and the importance of consumer lending for the well-being of households. Household debt in the United States as of 2017 was $13 trillion; of that, minority households account for 17.3%, or $2.25 trillion.[1] The largest component is home loans, on which $1.65 trillion is owed by minority households. Every extra basis point of interest charged on a home loan due to discrimination costs minorities $165 million per year in extra payments. Estimating discrimination, however, is notoriously difficult due to omitted-variable challenges. In this paper, we test for the presence of discrimination and estimate its level in mortgages securitized by the Government Sponsored Enterprises (GSEs), Fannie Mae and Freddie Mac. Using GSE loans allows a novel identification strategy without the omitted variable concerns that have long challenged empirical studies of discrimination. The GSEs charge each loan a guarantee fee that depends only on the (observable) credit score and loan-to-value ratio (LTV). In return, lenders are guaranteed against credit risk. Thus, mortgage interest rate differences between loans within a given GSE grid cell of credit score and LTV cannot reflect differential credit risk, but must instead reflect strategic pricing decisions on the part of lenders. Strategic pricing is not illegal, but the law requires that pricing must not induce a disparate impact.[2] Lenders cannot, even inadvertently, charge a higher strategic mark-up to protected groups. Using this novel identification strategy, we find that discrimination is 7.9 basis points (bps) for purchase mortgages and 3.6 bps for refinance mortgages. Averaging across the distribution of these products in the U.S., lending discrimination currently costs African-American and Latinx borrowers $765 million in extra interest per year.

Consumer lending in the United States is changing rapidly, with loan origination becoming almost exclusively algorithmic. A case in point is the Rocket Mortgage of the platform lender Quicken, which is the largest-volume mortgage product in the U.S. as of 2018. Algorithmic loan origination is not, however, just a feature of FinTech companies. We study the 2,098 largest mortgage lenders (inclusive of all the big banks) over the 2012-2018 period, finding that as of 2018, 45% of them offer complete online or app-based

---

[1] Percent-of-debt estimates are from the 2016 Survey of Consumer Finances. Aggregate debt statistics are from the Federal Reserve.

[2] A disparate impact occurs under U.S. anti-discrimination law when a decision maker's practices do not expressly discriminate on a protected characteristic (e.g., race or ethnicity), but nevertheless disadvantage one or more groups having a protected characteristic without a legitimate business justification. Under disparate impact theory, proof of discriminatory intent is not required to establish liability. As noted below, the U.S. Supreme Court has held that the disparate impact theory of liability extends to the Fair Housing Act, which (among other things) prohibits discrimination in home lending.

mortgage contracting. Value-weighted, that percentage is much higher. While these lenders continue to provide conventional, face-to-face loan applications, the trend is clearly toward automated underwriting. Nor is this simply a mortgage story; one only has to look to the emergence of personal-lending platforms, such as LightStream by SunTrust Bank and Marcus by Goldman Sachs, to see the broader transformation in consumer lending.

With algorithmic credit scoring, the nature of discrimination changes from being primarily concerned with human biases – racism and in-group/out-group bias – to being primarily concerned with illegitimate applications of statistical discrimination. Even if agents performing statistical discrimination have no animus against minority groups, they can induce disparate impact by their use of Big Data variables. Whether these changes induce more or less discrimination was previously unknown. We find that FinTechs do indeed remove some face-to-face discrimination in loan pricing. In particular, relative to the pricing discrimination we find across all lenders, FinTechs lenders discriminate nearly 40% less on average across purchase and refinance mortgages. This reduction is encouraging with regard to the potential for algorithmic lending to reduce discriminatory lending; however, it also highlights the persistence of discriminatory loan pricing even in the FinTech era. In addition to the ability of FinTech lenders to diminish face-to-face bias in loan pricing, our findings also point to two silver linings for the role of FinTechs and algorithmic decision-making. First, we find that discrimination in loan pricing is declining for all lenders from 2009 to 2015, alongside the advent of FinTech Lending. While we cannot prove causation, this finding is consistent with borrowers using online platforms to shop around with greater ease and speed, which should diminish the capacity for lenders to extract rents from minority borrowers. Second, in the loan accept/reject decision (as opposed to pricing), FinTech lenders reveal no evidence of discrimination, in contrast to our evidence of discrimination in rejection rates for traditional lenders.

At the core of our paper is the importance – in our identification and in the legal setting – of statistical discrimination in the new era of algorithmic loan decision-making. Regulators and courts face heightened hurdles to identify which Big Data variables can give rise to a successful claim of illegal discrimination under U.S. fair-lending laws (see *Inclusive Communities*, 2015).[3] For economists, the courts' struggle to untangle legitimate from illegitimate statistical discrimination is the same problem as handling omitted variables in estimating discrimination. Statistical discrimination arises as a solution to a signal extraction problem. The signal extraction setting in consumer lending emerges as follows. Economists can write down a macro-fundamental (life-cycle) model of default risk that applies to everyone.[4] The problem

[3] The potential for illegitimate statistical discrimination toward protected classes of borrowers was a key aspect of Congressman Emanuel Cleaver's 2017 investigation into FinTech lending.
[4] Behavioral models may correctly profile individuals on average, but some individuals would be incorrectly profiled, which could be deemed discrimination by disparate impact under the law.

is that some variables in this macro-fundamental model are not observable. The goal of statistical discrimination is to reconstruct this hidden fundamental information using observable proxies.

In the law, lenders can use proxy variables that produce a disparate impact on minority applicants, but only if the lender can show that these variables have a *legitimate business necessity*. According to the courts, *legitimate business necessity* is the act of scoring credit risk. Furthermore, according to the courts, efforts to use proxy variables that produce a disparate impact for other purposes, including lenders' earning of higher profit margins, do not meet this definition. In business terms, any strategic pricing that causes disparate impact, even inadvertently, is discrimination in the eyes of the law.

Lenders have used the *legitimate-business-necessity* defense to argue that any variable that is correlated with default is acceptable. This definition of *legitimate business necessity* is necessary but not sufficient to comply with the court rulings. An example is illustrative. Surely, the high school that a person attended is an empirically relevant proxy for hidden wealth, where wealth is the endowment variable in a macro-fundamental model of default risk. High school, however, may be correlated with race or ethnicity even after orthogonalizing with respect to wealth. If so, using high school would punish, or have disparate impact on, some minority households.

Our economic mapping of these court rulings on disparate impact to legitimacy in statistical discrimination yields three punchlines: (a) Scoring or pricing loans explicitly on credit-risk macro-fundamental variables is legitimate. (b) Scoring or pricing on a Big Data variable that only correlates with race or ethnicity through hidden fundamental variables is legitimate. (c) Scoring or pricing on a Big Data variable that has residual correlation after orthogonalizing with respect to hidden fundamental credit risk variables is illegitimate.

For policymakers, these punchlines suggest that regulators might take an approach of mandating that lenders provide proof of legitimacy of Big Data proxy variables using their proprietary data on otherwise-hidden fundamental variables such as wealth. This arrangement would be akin to putting the burden of value-at-risk modeling on banks, as is done in banking regulation. We discuss this more in the conclusion.

For researchers, these punchlines imply that in the age of algorithmic decision-making, econometricians require a setting in which all *legitimate-business-necessity* variables are observable in order to identify discrimination without concern for omitted-variable bias. We have been able to find just such a setting, covering a large fraction of consumer lending yet free from omitted-variable concerns. We use this setting to document the extent to which discrimination is happening in the largest consumer-loan market and to illustrate how algorithmic pricing of loans may yet result in discrimination.

It is well known that, post-crisis, the GSEs (Fannie Mae and Freddie Mac) purchase, securitize, and guarantee more than 90% of the conventional conforming mortgage market in the U.S. It is less recognized that, post-crisis, the GSE actions fully determine the price of credit risk via their role as guarantors. In particular, the GSEs produce a predetermined grid that prices credit risk across loan-to-value and credit-score buckets. The pricing grid need not be the optimal model for predicting default among all application variables,[5] but is nevertheless the price lenders must pay the GSE to absorb risk for the MBS market. Thus, any deviation from this grid pricing reflects lenders' competitive agenda in capturing volume or profit per mortgage. Because these non-credit-risk objectives are unrelated to creditworthiness, they fail to qualify as *legitimate business necessities* as determined by the courts. Thus, within the grid, any additional correlation of loan pricing with race or ethnicity is discrimination.

Our analysis uses a data set that includes never-before-linked information at the loan level on income, race, ethnicity, loan-to-value ratios, debt-to-income ratios, all contract terms (such as coupon, loan amount, installment-payment structure, amortization, maturity, loan purpose, and mortgage-origination month), and indicators for whether the lender-of-record primarily used algorithmic scoring.

Our main results are the following twin findings concerning the price of mortgages. First, accepted Latinx and African-American borrowers pay 7.9 and 3.6 basis points more in interest for home purchase and refinance mortgages respectively because of discrimination. These magnitudes represent 11.5% of lenders' average profit per loan.[6] Second, FinTech algorithms discriminate 40% less than face-to-face lenders; Latinx and African-American pay 5.3 basis points more in interest for purchase mortgages and 2.0 basis points for refinance mortgages originated on FinTech platforms.

How discrimination happens is an important question. We leave a full exploration of this topic to a separate research project, but we can fix ideas here. Lenders may be able to extract monopoly rents from minority borrowers because such borrowers might be prone to less shopping on average. The fact that the magnitude of discrimination in refinance loans is lower than in purchase mortgages is consistent with an interpretation that monopoly price extraction of rents is easier in purchase-mortgage transactions, where the borrowers have less experience or are acting in a more urgent time frame. Additionally, because lenders may price loans to capture rents in less-competitive areas, prices might be higher in financial-services deserts, which might have higher minority populations. These pricing mechanisms can play out by human or machine intervention. For instance, one can easily imagine both lending algorithms and human loan officers seeking to detect which types of borrowers are less prone to shopping or which types of geographies have less competitive pricing.

[5] The actuarially fair GSE guarantee fee (or G-fee) is also a central policy question in the determination of the future role of the GSEs in the U.S. mortgage markets (see Elenev, Landvoigt, and Van Nieuwerburgh, 2016; Vickery and Wright, 2013). A standard G-fee is assessed on all mortgages as a percentage of the loan balance and is collected monthly (see Fuster, Goodman, Lucca, Madar, Molloy, and Willen, 2013).

[6] According to the Mortgage Bankers' Association, the average mortgage profit is 50 basis points (see https://www.mba.org/x73719).

We consider the robustness of our estimates to lingering concerns. Although courts have explicitly held that credit risk is the only *legitimate business necessity*, we believe the spirit of these decisions may include room for lenders to differentiate loan pricing based on the fixed cost of providing a loan by lender or by geography. We thus additionally include county and lender fixed effects, as well as county crossed with lender fixed effects, with results remaining robust. We also address the robustness of this result to other concerns of servicing rights and to the use of points.

Turning to our findings with regard to loan rejection rates, we first note that any discrimination in loan rejection rates—as opposed to discrimination in loan pricing—would appear to be inconsistent with lenders' profit maximization. In our setting of the GSE guarantee, if lenders were to discriminate in the accept/reject decision, it would imply that money is left on the table. Logic suggests that such unprofitable discrimination must reflect a human bias by loan officers. This is what we find.

Face-to-face lenders reject Latinx and African-American applications approximately 6% more often than they reject similarly situated non-minority applicants for both purchase and refinance loans. In aggregate, our findings suggest that from 2009 to 2015, lenders rejected 0.74 to 1.3 million Latinx and African-American applications that would have been accepted except for discrimination. FinTech lenders, on the other hand, do not discriminate at all in the decision to reject or accept a minority loan application in our sample. This is consistent with algorithms acting in a profit-maximizing manner. Because our findings with respect to rejections must rely on proxies for certain variables utilized by the GSEs in approving loans, we note that these results are preliminary. But they nevertheless point toward the possibility that fully automated underwriting may reduce the incidence of discrimination in loan rejections.

Our paper contributes to a small but growing literature on discrimination in lending. A large literature in labor contributes to the topic of wage discrimination, but even there, our commentary on how courts and regulators can consider Big Data use may be informative. The lending discrimination literature has lagged the wage literature primarily because of the lack of data on ethnicity or race combined with an identification strategy that handles omitted variables in scoring.

Early studies looking at the raw HMDA data found that minority loan applicants were rejected much more often than white applicants even with higher incomes, however, these papers but did not control for variables not collected by HMDA, such as credit history. In a widely cited paper, Munnell, Browne, McEneaney, and Tootel (1996) combined HMDA data on loan applications in Boston in 1990 with additional borrower data collected via survey by the Federal Reserve Bank of Boston, and found that after controlling for borrower characteristics, especially credit history and loan-to-value ratio, white applicants with the same property and personal characteristics as minorities would have experienced a rejection rate of 20% compared with the minority rate of 28%.

Much of the more recent literature focuses on the pre-crisis period, usually looking at subprime lending. Ghent, Hernandez-Murillo, and Owyang (2014) examine subprime loans originated in 2005, and find that for 30-year, adjustable-rate mortgages, African-American and Latinx borrowers face interest rates 12 and 29 basis points, respectively, higher than other borrowers. Bayer, Ferreira, and Ross (2018) find that after conditioning on credit characteristics, African American and Hispanic borrowers were 103% and 78% more likely, respectively, than other borrowers to be in a high-cost mortgage between 2004 and 2007. Similar results were obtained by Reid, Bocian, Li and Quercia (2017).

Cheng, Lin and Liu (2015) use data from the Survey of Consumer Finances to compare mortgage interest rates for minority and non-minority borrowers. They find that black borrowers on average pay about 29 basis points more than comparable white borrowers, with the difference larger for young borrowers with low education, subprime borrowers, and women.

Focusing on the *quality* of consumer credit services, Begley and Purnanandam (2018) study the incidence of consumer complaints about financial institutions to the CFPB. They find that the level of complaints is significantly higher in markets with lower income and educational attainment, and especially in areas with a higher share of minorities, even after controlling for income and education.

In one of the few experimental papers in this area, Hanson, Hawley, Martin, and Liu (2016) show that when potential borrowers (differing only in their name) ask for information about mortgages, loan officers are more likely to respond, and give more information, to white borrowers.

Finally, Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2018) show that the use of machine-learning techniques to evaluate credit quality may result in differential impact on loan provision to minority versus non-minority borrowers. This paper conveys important knowledge in how algorithms are utilized in mortgage markets.

There are also related results from other consumer debt markets. For example, Dobbie, Liberman, Paravisini and Pathania (2018) look at data from a high-cost lender in the UK and find significant bias against immigrant and older loan applicants when measured using long-run profits. However, they find no bias when using the (short-run) measure actually used to evaluate loan examiners, suggesting that the bias is due primarily to a misalignment of firm and examiner incentives.

The rest of the paper is organized as follows. In Section II we discuss our data. We present our methodology for the measurement of mortgage discrimination in Section III, and provide statistics showing the role of the GSE pricing grid in practice. Our empirical results are reported in Section IV. Section V concludes and discusses regulatory implications of our findings.

II. Data & Statistics

A key obstacle for prior studies of mortgage discrimination has been a reliance on Home Mortgage Disclosure Act (HMDA) data. The HMDA compliance surveys cover 90% of mortgage originations in the U.S. (see Engel and McCoy, 2011),[7] and are the only data source with loan-level information on applicant race and ethnicity for both successful and unsuccessful loan applications. What HMDA lacks is information on the contracting structure of the loan (exact date, interest rate, maturity, loan-to-value ratio), on the type of loan (fixed, ARM), on the property characteristics (e.g., address), and on the applicant's credit data used by the GSEs and other lenders (credit score, debt-to-income ratio, etc.).

A challenge with mortgage loan data in the U.S. has been the lack of a unique loan-identification number and thus the lack of a direct way to link the HMDA data and other datasets containing these missing data. We ameliorate this deficiency with a multi-year project of linking loan-level data across the following data providers:

- HMDA data include information on applicant income, race, ethnicity, loan amount, and lender name, as well as the census tract of the property.
- ATTOM data provide transaction and assessor information including lien-holder name, loan performance data (i.e., prepayment and default), borrower and lender names and exact property location, but very little information on mortgage contract terms other than the loan amount, the origination date, the purpose of the loan, and whether it is a fixed or floating contract.
- McDash data provide loan-level data compiled by Black Night Financial Services and include detailed mortgage terms (including interest rates, loan amount, loan-to-value ratio, and zip code of the mortgaged property) and month-by-month mortgage performance information.
- Equifax data provide information on other consumer financing balances that are held by borrowers in addition to their mortgages and the borrower credit score.
- Freddie Mac Single Family Loan-Level Dataset and Fannie Mae Single Family Loan Performance Data: These data were used to construct estimates of the median and the 25^{th} and 75^{th} quartiles for the census tract level "back end" debt-to-income ratio distributions.[8]

Using a machine-learning protocol, we exploit overlapping variables within HMDA, ATTOM, and the McDash/Equifax datasets to construct a merged data set of accepted loans with performance information, contract terms, the mortgage lender, and borrower information. We describe our machine-learning merging algorithm in the Appendix. A key component of the merging was the McDash-ATTOM link, which we

[7] HMDA reporting is not required for institutions with assets (of the entity and its parent corporation) that are below $10 million on the preceding December 31 (see http://www.ffiec.gov/hmda/pdf/2010guide.pdf).
[8] The back end debt-to-income ratio is the sum of a borrower's debt expenses (calculated as all credit report payments plus the payment implied by the current mortgage payment plus taxes and insurance) divided by the borrower's gross monthly income.

accomplished by matching performance strings; i.e., matching loans on the flow of events reported in the property registers reported by ATTOM and in the loan-level performance data reported by McDash. The Equifax-McDash merge was done by Equifax. Our use of the merged dataset is in compliance with IRB standards, and our residual data are anonymized.

The HMDA data include information on both ethnicity and race. For our purposes, we definite a minority applicant to be one with either Latinx ethnicity or African-American race. However, HMDA has missing values on race and ethnicity (Buchak et al., 2017). We therefore augment the HMDA race/ethnicity indicator variable with the additional race/ethnicity data obtained from processing the borrower name field from ATTOM data, using a race and ethnic-name-categorization algorithm from Kerr and Lincoln (2010) and Kerr (2008). In the robustness section, we check for consistency of our results excluding these fixes.

We focus on two types of lender origination decisions – the decision to accept or reject a loan application and the loan pricing conditional on acceptance.

Table 1, Panel (a) reports the summary statistics for the pricing estimations. To standardize our loan pricing analysis, we filter the data to focus on 30-year, fixed-rate, single-family residential loans, securitized by the GSEs over the period 2009 through 2015.[9] We additionally eliminate from our sample any loans made within a census tract covered by the Community Reinvestment Act of 1977 (CRA), given the potential bias these census tracts would introduce into our empirical analysis.[10] The final pricing analysis sample consists of 3,577,010 loans. The dependent variable, the interest rate on the mortgage, has both a mean and a median of 4.50%. The mean loan amount is $234,000, reflecting an LTV of 0.744, from an applicant that has an 11% probability of being a Latinx or African-American borrower. This borrower has $107,200 in income and a high credit score of 755.8.

For our accept/reject analysis, the only loan-level source of data is HMDA, which records application status in the "Action Taken" field. An Action Taken equal to one is a reject ("Application denied by financial institution").

Three weaknesses exist in the accept/reject data for purposes of estimating discrimination in our accept/reject analysis. First, public HMDA data do not indicate whether loan applications are to be securitized through the GSEs, if approved. We mitigate this problem by limiting the accept/reject sample to HMDA mortgages qualifying as being *conventional* (not being backed by the Federal Housing Administration, the Veterans Administration, the Farm Service Agency or the Rural Housing Service) and *conforming* (the loan size falls below the annual conforming loan limit set by the Federal Housing Finance Agency). We use these two criteria as a proxy for the mortgage being a GSE loan. Second, we are unable

[9] We also do a rate analysis on 15 and 20 year mortgages. Statistics and estimation results are included as Appendix Table 1.
[10] Under the CRA Act, financial institutions are required to provide a certain level of lending to CRA districts to counter the lack of financial services in lower-income districts.

to further filter to 30-year loan applications since HMDA does not report the applied-for loan maturity. Therefore our data pool together 15-, 20-, and 30-year applied-for mortgages for the accept/reject analysis. Third, while we know the precise (observable) variables that the GSEs use to determine loan acceptance/rejection in their automated underwriting systems, we do not have loan-level data on some of these application variables. We therefore augment the rejection data with measures, described in the Appendix, of credit scores, debt outstanding, debt-to-income ratios, and loan-to-values using medians that are computed at the census-tract level. A census tract is on average 1,600 households (4,000 inhabitants), designed by the Census Bureau to reflect relatively uniform economic situations. Because our data on rejected applications have these potential weaknesses, we exert more caution in interpreting the accept/reject results (as opposed to the pricing results).

Table 1 reports the summary statistics for the accept (Panel (b)) and reject (Panel (c)) applications. Slightly over half of the mortgages are acceptances. The final sample for the accept/reject analysis consists of 6,648,413 accepted loans and 6,535,664 rejected loans. Latinx and African-Americans account for 11.9% of accepted and 18.6% of rejected applications. As expected, accepted applicants have stronger credit-risk profiles than those rejected. Accepted applicants exhibit a higher mean income of $108,300 (versus $97,400 for rejected loans) and a higher mean applied-for loan amount of $213,900 (versus $187,300 for rejected). The summary statistics for the census-tract proxies reflect a similar pattern. Accepted applicants exhibit a higher average credit score of 750.8 (versus 744.2 for rejected loans) and a slightly lower average loan-to-value ratio (LTV) of 0.791 (versus 0.812 for rejected loans). Accepted and rejected applications appear similar with respect to the mean total debt outstanding and mean debt-to-income (DTI) in the census-tract proxies.

Table 1 also reports summary information concerning the types of lending institutions that received the loan applications in our sample. Using the list of firms identified as FinTech in Buchak et al. (2017), we find that FinTech lenders originated approximately 4.3% of accepted loans (Panels (a) and (b)) and were responsible for 5.5% of all loan rejections in our sample. Table 1 also highlights the dominance of the largest originators in the mortgage lending industry. The top 25 originators (by origination volume in their respective loan-origination year) both accepted and rejected over 50% of all loans processed.[11]

In all of our analysis, we divide the market between purchase and refinance loans. Purchase loans represent 41.8% of the loans in the pricing analysis, probably because we focus on 30-year maturities rather than 15- or 20-year maturities, which are preferred for refinances. Consistent with this conjecture, for the accept/reject estimations, purchase loans represent 30.8% of accepted loans and 17.3% of rejected loans.

[11] We create a variable of the top 25 mortgage originators per year by matching HMDA lender names with mortgage origination statistics obtained from Inside Mortgage Finance.

III. Method

III.a. The GSE Lending Process

GSE involvement in the mortgage process begins with the submission of an application into the GSE underwriting system. The lender feeds application observables (the credit score, income, liquid reserves, debt-to-income ratio, loan-to-value ratio, property value, etc.), into the GSE 'black box,' an automated underwriter system (Desktop Underwriter for Fannie Mae; Loan Prospector for Freddie Mac). The GSE black box produces the accept/reject decision based on a specified set of observables contained in the application. If the GSE accepts the loan and the lender issues it, the lender sells the mortgage to the GSE. In return, the GSE compensates the lender with a cash transfer.[12] The GSE then packages the loan with a pool of similar mortgages into a mortgage-backed security (MBS), issues a default-risk guarantee on this product, and sells it to the MBS market.

In light of the sale of the mortgage to a GSE, the lender is not exposed to any prepayment or default risk. The only risk that the lender faces is put-back risk. Put-backs can occur when the documentation on income (tax returns, pay stubs, etc.), credit score, loan purpose (residential vs. non-occupancy) or property value (the appraisal) is falsified or missing. After the 2008 financial crisis, because of put-backs and large fines for misrepresentation, lenders ceased no-documentation GSE loans and adjusted their policies to lessen the potential for falsified documentation.[13] The magnitudes of put-backs on post-2008 originations have become a trickle compared to the early 2000 issuances. Figure 1, taken from Goodman, Parrott and Zhu (2015), plots put-back rates over the time horizon of the loan, highlighting the different put-back rates across loan-vintage years for loans issued between 2000 and 2010. The figure supports our assumption that put-back risk is immaterial post 2008.

Within this GSE process, the lender has decisions to make, which constitute occasions when discrimination can happen. The first is pricing. The mortgage interest rate that a household sees consists of three parts (see Fuster et al., 2013). First, all mortgages face the same market price of capital, determined by the base mortgage rate, which reflects the primary market interest rate for loans to be securitized by the GSEs, in essence, the credit-risk-free rate. Second, when a GSE buys a mortgage from a lender, the GSE takes a guarantee fee (or g-fee) to cover projected borrower default and operational costs. Starting in March 2008 and adjusted a handful of times since then, this g-fee varies in an 8×9 matrix of LTVs and credit scores

[12] If the originator is a large-volume lender, the lender will transfer loans to the GSE in bulk and, instead of receiving cash for the mortgages, the originator receives back an MBS with a pool of similar-characteristic mortgages produced by that lender. (Sometimes these MBS products have mortgages originated by other lenders to fill out the MBS, but one should think of this pool as primarily being the lender's own issuance.) These MBS products are equally guaranteed by the GSE, but because the lender also retains servicing rights, the lender may be exposed to the extra servicing costs (e.g., additional phone calls and outreach) that happen when loans become delinquent. For this reason, we show all of our results with and without the large-volume lenders.

[13] The GSEs put back $4.2 billion of pre-crisis loans in 2010 alone (American Banker, July 14, 2016).

to reflect varying credit risk across the GSE grid. Figure 2 depicts a typical GSE grid of Fannie Mae, also called the Loan Level Price Adjustments (LLPAs) (see FHFA 2000; 2010; 2011; 2012; 2013 as well as Fuster and Willen, 2010). In practice these one-time fees are commonly converted into monthly "flow" payments, which are added into the interest rate as rate pass-throughs to borrowers. In quoting rates to customers, originators utilize rate sheets that expressly incorporate both the base mortgage rate (generally reflected as the "par rate") as well as LLPA adjustments (colloquially referred to as "hits" to the par rate).

The third component of pricing comes from lenders' discretion in quoting rates that deviate above the par rate plus g-fee "hit." Deviations may reflect simple lender fixed-effects differences in overhead costs, but they also might reflect strategic volume positioning or monopoly rent-taking. Of particular interest is the possibility that lenders use monopoly-like pricing based on the competitive environment of a location (e.g., in areas of collusion or in financial-desert environments) or as a rent-extraction strategy against borrowers who shop around less. These examples (which may be human or machine-coded) may induce a disparate impact on minorities, i.e., inadvertent discrimination.

III.b. Identification of Rate Discrimination

Our identification of discrimination relies on the legal setting established by U.S. fair lending law.[14] A lender accused of discrimination under U.S. fair lending law can assert a defense based on the principle of *legitimate business necessity*. One might imagine that many activities fall under business necessity related to a lender's goal of profit maximizing, but the courts have consistently limited the *legitimate business necessity* defense to a lender's use of variables and practices to ascertain creditworthiness.[15] Thus, the use of variables or practices that induce higher profit-taking (above creditworthiness) from, for example, charging higher rates to applicants in financial deserts or applicants with low shopping characteristics cannot be justified as legitimate business necessity, even though the use of these variables may be profit maximizing. Using strategic pricing variables is not illegal, but the use of these variables cannot fall disproportionately on minorities or other protected categories.

---

[14] We define U.S. fair lending law as including the Fair Housing Act and the Equal Credit Opportunity Act (ECOA), together with all implementing regulations and judicial interpretations relating to them.

[15] See A.B. & S. Auto Service, Inc. v. South Shore Bank of Chicago, 962 F. Supp. 1056 (N.D. Ill. 1997) ("[In a disparate impact claim under the ECOA], once the plaintiff has made the prima facie case, the defendant-lender must demonstrate that any policy, procedure, or practice has a manifest relationship to the creditworthiness of the applicant…"). See also Lewis v. ACB Business Services, Inc., 135 F.3d 389, 406 (6th Cir. 1998) ("The [ECOA] was only intended to prohibit credit determinations based on 'characteristics unrelated to creditworthiness.'"); Miller v. Countrywide Bank, NA, 571 F.Supp.2d 251, 258 (D. Mass 2008) (rejecting argument that discrimination in loan terms among African American and white borrowers was justified as the result of competitive "market forces," noting that prior courts had rejected the "market forces" argument insofar that it would allow the pricing of consumer loans to be "based on subjective criteria beyond creditworthiness.")

The legal environment provides guidance for identification of discrimination, but also requires that an econometrician be able to observe all variables determining creditworthiness. However, as described above, the GSEs' role in guaranteeing loans provides a setting (in the largest consumer loan market in the United States) in which we can fully see the price of credit risk by observing a borrower's LTV and credit score.

In Figure 3, we depict the importance of the GSE grid for purposes of pricing loans. In Panel (a), we plot histograms of raw mortgage interest rates sorted by minority status. The histograms reveal a wide distribution of rates for both minority and non-minority loans, as one might expect given the length of our sample period and the large number of loans in the sample. However, when we level interest rates within the grid by subtracting out the month-year-grid cell mean, Panel (b) shows a dramatic reduction in the distribution of excess interest rates for both groups of borrowers, highlighting the central role of the LLPA grid in determining interest rates for GSE mortgages.

We translate the Panel (b) figure into an empirical model for application *i* occurring in the month-year *t*:

$$interest\ rate_{it} = \alpha\ LatinxAfricanAmerican_i + \mu_{GSEgrid} + \mu_{month\_year} + \varepsilon_{it}, \tag{1}$$

where the mortgage interest rate is regressed on the indicator for the applicant being a Latinx- or African-American, 72 GSE grid fixed effects $\mu_{GSEgrid}$, and month-year fixed effects μ_{month_year}. Under the GSE identification, any loading of pricing on race/ethnicity is discrimination. For robustness, we also implement models using the richness of our data environment. For instance, one concern would be that the interpretation of U.S. fair lending law might evolve such that a court would permit a lender to advance as a *legitimate-business-necessity* defense the ability to recoup overhead costs that differ by lender or that differ for originating loans in different geographic locations. To address this possibility, we implement a model with lender fixed effects, μ_{lender} crossed with county fixed effects, μ_{county}:

$$\begin{aligned} interest\ rate_{it} = {} & \alpha\ LatinxAfricanAmerican_i + \mu_{GSEgrid} + \mu_{month_{year}} \\ & + \mu_{lender} * \mu_{county} + \varepsilon_{it}. \end{aligned} \tag{2}$$

This specification forces the appealing interpretation on the discrimination coefficient to be the differences in average interest rate charged to a minority applicant as compared to that offered to an non-minority applicant by the same lender in that same county within the GSE credit model. Although appealing econometrically, this rigor throws out some of the variation in which we are interested. In the results section we also tackle robustness concerns that may arise from the use of points and exposure to servicing costs.

III.c. Identification of Accept/Reject Discrimination

The second decision that a lender makes is the accept/reject decision, which occurs before the pricing decision. Even though an application might receive a creditworthiness approval in the GSE underwriter system, the lender may still reject an application. If the lender uses the GSE market for securitizing mortgages, no credit risk would remain post-transacting; hence, money would be left on the table.

Why would a lender choose to reject a GSE-accepted applicant? (i) The lender might feel that a particular borrower reflects additional put-back risk. As we have argued, such put-back risk is so small, especially in the latter half of our sample, that even if this put-back risk were residually correlated with race or ethnicity (which is not established), it would not be able to explain any material differences that we find in rejection rates. Thus, this argument would amount to a biased belief affecting loan decisions. (ii) The lender might be directly racist or have other in-group biases. (iii) The lender might prefer to cater to non-minority borrowers in branch banking. None of these explanations falls under *legitimate business necessity*.[16]

Our linear probability model of rejection discrimination for application *i* in year *t* is:

$$rejection_{it} = \beta\ LatinxAfricanAmerican_i + f(HMDA\ income_i, HMDA\ loan\ amount_i) \quad (3)$$
$$+\ g(LTV_c, credit\ score_c, debt\ outstanding_c, DTI_c) + \mu_{year} + \varepsilon_{it}.$$

Rejection is an indicator for an application being rejected by the lender. The $f(\cdot)$ function is a non-parametric function of the original HMDA data for income and loan amount. Since we do not know the exact scoring function of lenders on these variables, we use 21 piecewise-linear splines of income and 47 of loan amount, rather than a linear form. We control for year fixed effects, μ_{year}, rather than month fixed effects because HMDA does not provide precise dates for the rejections. As mentioned above, we do not have loan-level data on all underwriting variables entering the GSE black box underwriter system. Thus, we proxy using the equivalent variable at the census tract (1,600 households) of the property, *c*. To capture the distribution within the census tract, we include the 25th, 50th, and 75th percentile of *LTV*, *credit score*, *debt outstanding*, and *DTI* in the census tract, denoted by the function $g(\cdot)$.

Two concerns come to mind that may affect the robustness of our estimation. First is the concern that the lender may keep some loans on its balance sheet after using the GSE accept/reject underwriting. Some large lenders may have their own credit risk model with *legitimate business necessity* variables unseen to the GSE underwriter (and thus to us as econometricians). If these lenders use fundamental models to cherry-pick loans to keep on their balance sheet, our empirical model may load credit risk on our estimation of the Latinx-/African-American variable (assuming this latter variable is correlated with any supplemental, unseen credit risk variables). We address this concern by eliminating large lenders in our

[16] Note that a lender that intentionally treated applicants differently based on a protected characteristic would be liable under the *disparate treatment* theory of discrimination, for which there is no formal legitimate business necessity defense. In effect, the disparate treatment theory of liability assumes that intentional discrimination can never constitute a legitimate business necessity.

robustness tests, since small lenders are unlikely to hold many mortgages on their balance sheet. Removal of large lenders also addresses the concern that servicing-cost risks may affect accept/reject decisions, since only the large lenders obtain servicing contracts for the GSE pools of loans.

Second is the concern that a loan officer might deter a potential minority borrower from applying or coach non-minority applicants to increase their likelihood of acceptance. If we had a perfect set of both accepted and rejected applicants' data, we could recreate the GSEs' black-box algorithms and eliminate these possibilities from biasing our results. Without perfect data, this concern creates the possibility for bias in our estimation of discrimination in rejection rates. Importantly, however, this concern is not valid for our FinTech discrimination estimations given the absence of any loan officer-to-application interaction.

Additionally, even within our non-FinTech results, these two scenarios do not create the same risk of bias, and may very well bias us away from a finding of discrimination. Imagine, for instance, that all lenders coach every non-minority applicant to delay submitting a loan application until the non-minority applicants can increase their credit scores. In such a setting, actual non-minority credit scores will be stronger than the census-tract proxy that we observe. As a result, our specification will estimate discrimination through the fact that non-minority applicants are approved at a greater rate than minority applicants from the same census tract due to the stronger credit scores of the (coached) non-minority applicants. Conversely, imagine all lenders overtly refuse to consider minority applicants except for those applicants having the strongest credit scores. In this setting, among applications that are submitted, actual minority credit scores will likewise be greater than the census-tract proxy that we observe. As a result, our specification will underestimate discrimination due to the fact that minority applicants will be disproportionately accepted within a census tract given the heightened application standard.

Although we cannot fully address this latter possibility, we again can implement a model of lender crossed with county fixed effects because of the richness of our data. To the extent this latter scenario arises from branch-wide behavior by loan officers, these granular fixed effects models will permit us to estimate discrimination in rejection rates that goes beyond this form of overt bias against minority applicants.

IV. Results

IV.a. Interest Rate Discrimination Estimates - Main

Table 2, Panel (a) presents within-GSE grid estimates of interest rate discrimination for 30-year fixed-rate mortgages that are approved and originated. (Appendix Table 1, Panel (b) reports an identical table for 15- and 20-year maturity mortgages.) Because lenders' pricing strategies vary by mortgage type, we present estimates for purchase mortgages (columns (1) and (2)) separately from refinance mortgages (columns (3) and (4)). Columns (1) and (3) present simply the overall mean differences in prices paid by Latinx-/African-Americans versus everyone else. Columns (2) and (4) are our full credit-risk model of equation (1), containing the 72 GSE grid fixed effects to capture the pricing in the grid plus the month-year fixed effects to capture market-price fluctuations.[17]

The overall mean difference in the purchase-mortgage interest rate between Latinx/African-American and non-minority borrowers is 0.090%, or 9.0 basis points. Of this amount, column (2) shows that 1.1 basis points are explained by the credit-risk model, leaving 7.9 basis points of discrimination. For refinance mortgages (columns (3) and (4)), we identify an economically-smaller price discrimination of 3.6 basis points. The interpretation of this main credit-risk-model result is that conditional on being given a loan, African-American and Latinx borrowers pay an average of 7.9 basis points more than other similar borrowers for their purchase mortgages and 3.6 basis points more for refinance mortgages. Given the Mortgage Bankers' Association mean profit of 50 basis points, these discrimination premiums would represent 16% and 7% respective increases in profits for lenders, or 11.5% on average.

Also of interest in column (2) is the ability of the credit-risk model to explain 73% of the variation across nearly 1.5 million purchase mortgages and 69% of the variation in 2.1 million refinance mortgages. The unexplained variation (one minus the R^2 = 27%-31%) may reflect strategic pricing either on borrowers' location (perhaps due to collusion or to opportunistic pricing in financial deserts) or on borrowers' behavioral characteristics (perhaps reflecting profiling using variables or soft information that correlate with a lack of shopping). The disparity between purchase and refinance mortgage discrimination suggests that borrower sophistication and hurriedness matter. Refinancing borrowers are, by definition, experienced and may be in less of a hurry to re-contract than the average purchase-mortgage borrower, who may be time constrained to bid on a house on the market.

To put these magnitudes in more context, Panel (b) shows a back-of-the-envelope calculation of extra interest paid due to discrimination in loan pricing. The total U.S. mortgage market float is $9.5 trillion.[18] Assuming the existing float of mortgages consists of 75% refinance loans and 25% purchase loans and that Latinx- and African-Americans borrowers make up 17.3% of the total float (from the Survey of Consumer Finances (SCF)) and that the average mortgage has a term of 30 years with a 3.5% coupon, we find that discrimination in mortgage interest rate costs Latinx- and African-Americans $765 million extra in interest annually.

IV.b. Interest Rate Discrimination Estimates – FinTech Lenders

The twin goals of our paper are to estimate the extent of consumer lending discrimination ($765 million in mortgages alone) and to ask whether FinTech originators perform any better in avoiding discrimination. Although face-to-face lenders provide loan officers with personal contact with applicants

[17] Our estimates are almost identical if we instead use GSE grid fixed effects interacted with the month-year dummies.
[18] Source: Federal Reserve.

that can induce racism and in-group bias in decision-making, platforms may have equal opportunity to cause inadvertent discrimination. Algorithmic pricing of loans applies estimation techniques over large sets of data to enable profit-maximizing pricing strategies. An algorithm would naturally discover that higher prices could be quoted to profiles of borrowers or geographies associated with low-shopping tendencies. As described earlier, if such pricing induces higher markups for minorities, the lender must have a *legitimate business necessity* defense for this form of algorithmic profiling. However, as noted, courts have consistently limited the *legitimate business necessity* defense to a lender's use of variables and practices to ascertain creditworthiness. In the case of mortgage lending in the GSE system, no residual creditworthiness assessment is needed within the GSE grid to price credit risk; therefore, pricing strategies that cause higher markup for minorities using this strategy would constitute impermissible discrimination. (We note below that face-to-face lenders may also seek to charge higher rates to borrowers having a lower propensity to shop around by preparing different rate sheets by branch or geography.)

Table 3 shows the results for FinTech lenders for purchase mortgages (column (1)) and refinance mortgages (column (2)). Both columns use our full credit-risk model containing the 72 GSE grid fixed effects plus month-year fixed effects. We find, reported in column (1), that FinTech purchase-mortgage discrimination is 5.31 bps, approximately 2/3rds the magnitude of the estimate for the full sample of issuers in Table 2. Column (2) reports that discrimination by FinTech lenders in the pricing of refinance mortgages is 1.97 bps, 55% of the magnitude for the sample of all issuers.

We conclude two things from Table 3. FinTechs do indeed remove some face-to-face biases. In particular, FinTechs discriminate 40% less on average across the two mortgage products. This non-trivial reduction is encouraging with regard to the potential for algorithmic lending to reduce discriminatory lending. Yet the result also clearly tells a flip-side. Both FinTechs and face-to-face lenders may discriminate in mortgages issuance through pricing strategies. We are just scratching the surface in the role of pricing strategy discrimination in the algorithmic era of data use. In short, algorithmic lending may reduce discrimination relative to face-to-face lenders, but algorithmic lending is not alone sufficient to eliminate discrimination in loan pricing.

We end this section on a positive note. We find two additional silver linings for discrimination in the era of FinTech. The first can be seen in Figure 4, which shows discrimination in loan pricing by the year of loan issuance. Discrimination has declined between 2009 and 2015. Although we cannot prove causality, this result could be due to competition from the platforms and/or the ease of shopping around made possible by online applications. The second additional silver lining for the role of FinTechs comes in rejection decisions, to which we return in Section IV.d.

IV.c. Interest Rate Discrimination Estimates- Robustness

The identifying assumption in equation (1) – and therefore, the identifying assumption behind the estimates in Tables 2 and 3 – is that any legally permissible differences in loan pricing must be a function of borrowers' differing placement within the GSE grid, as well as the month of loan origination as an adjustment variable for the credit risk-free rate. As discussed previously, this assumption should be satisfied by (a) the fact that all credit-related information relevant to GSE loan pricing is captured by the GSE grid and (b) the fact that courts have limited the *legitimate-business-necessity* defense to pricing differences that arise from differences in borrower credit risk. Nonetheless, we explore several potential robustness concerns.

The first relates to the possibility that courts may be amenable to the legal argument that variation in loan pricing depends not only on borrower credit risk but also on a lender's costs, which may differ by lender and by geography. While the courts have not ruled that *legitimate business necessity* includes locational or lender fixed costs, we can imagine such an argument might arise. Table 4 therefore introduces lender and county fixed effects to the main credit-risk model specification of Table 2, which allows appealing within-county, within-lender, and within-county-lender comparisons as robustness. Across all specifications in Table 4, we find discrimination in loan pricing. Even within the same lender originating loans in the same county (column (4)), Latinx- and African-Americans pay 5.2 basis points more for purchase mortgages and 2.0 basis points more for refinance mortgages.

While econometrically interesting, we caution that these specifications may underestimate the incidence of illegal discrimination. In particular, these specifications could fail to capture variation in loan pricing that is indicative of illegal discrimination. Consider, for instance, a lender that establishes a branch office in a county with primarily minority residents for the express purpose of issuing premium-priced loans to the branch's largely minority customers. If the lender uses a branch-wide rate sheet with exorbitant rates (a practice that has been documented in several enforcement actions),[19] the lender crossed with county fixed effect would absorb all variation in loan pricing by this institution. Yet such behavior would almost surely be viewed as conventional red-lining, which is illegal under the disparate treatment theory of discrimination (Gano, 2017). Similar scenarios can be envisioned for the county fixed effect (e.g., all lenders engage in this practice within various counties) and the lender fixed effect (e.g., small lenders establish offices in minority-majority counties and use higher rate sheets for the purpose of placing minority borrowers in higher-priced loans, rather than for cost reasons).

[19] For example, in United States v. Sage Bank, Sage Bank agreed to pay a fine of $1,175,000 and take other remedial action after the United States Department of Justice alleged that Sage bank had assigned a higher "target price" to loan officers who disproportionately served African-American and Hispanic borrowers, resulting in higher priced mortgages for these customers. See United States v. Sage Bank, Sage Bank, Consent Order, November 11, 2015, available at https://www.justice.gov/opa/file/796371/download.

In addition to the foregoing concern about legal robustness, three additional concerns might confound the interpretation of our estimates: the use of points, servicer cost risk, and the designation of race or ethnicity. Table 5 presents a series of tests for these concerns. All estimations are in the form of the full credit-risk specification of Table 2, columns (2) and (4).

The term "paying points" refers to the borrower's act of paying a lump sum to a lender to reduce the interest rate. An interpretation consistent with our results is that minorities could be more fully utilizing their cash for down-payment, leaving no cash to pay points to reduce the rates. This would imply different interest rates for minority borrowers without also implying discrimination. We test for this story in purchase (Panel (a)) and refinance (Panel (b)) mortgages in column (1) of Table 5. In particular, we limit the sample to borrowers precisely at the 0.80 LTV threshold and whose non-mortgage debt is greater than 30% of reported income. This set is likely to contain borrowers who are scraping up funds to just make the down-payment required to meet an LTV of 0.80, which is the LTV threshold at which borrowers are exempt from purchasing mortgage loan insurance. Going over this threshold also results in a higher loan interest rate. A histogram of LTVs depicted as Figure 5 illustrates how important the LTV threshold is. Table 5, column (1) shows that discrimination is even higher (8.73 bps as opposed to 7.88 bps) for these borrowers, inconsistent with our results being driven by (positive) points. (Results are virtually identical if we examine borrowers at the 0.80 LTV threshold without regard to their non-mortgage debt).

Another points story consistent with our results is that Latinx and African-American borrowers may be paying *negative* points (incurring a higher interest rate) to get a rebate in cash to pay closing costs. Although we cannot see the pricing in our dataset, conversations with mortgage brokers suggest that the interest rate costs of taking these yield rate spreads are high. Thus, if a borrower could pay a slightly lower down-payment, retaining some cash for closing costs, without inducing any interest rate increase, this choice would be optimal for most borrowers. The borrowers who are likely to be able to slightly decrease their down-payments while not increasing their rates are those not facing the upper threshold of the LTV grid (i.e., 0.80, 0.75, etc., see Figure 2 for the grid definitions). Thus, in column (2) of the panels of Table 5, we consider the robustness of our results to negative points by estimating our full credit risk model for the sample of LTV borrowers that have LTVs between (0.70, 0.74) and (0.75, 0.795). These borrowers are not at the edge of the GSE grid bucket in LTV and should therefore have little incentive to pay negative points. The results in these columns are very similar to those in Table 2, with larger magnitude estimates for purchase mortgages.

The third robustness concern taken up in Table 5 involves the role of residual risk for lenders who hold the GSE-guaranteed MBS or those who service these loans. GSE loans are special in that many lenders do not retain servicing rights in the GSE process nor do they hold the asset on their balance sheets once the loan is put through the GSE system. This is not always true for the very large bank lenders, who provide a

GSE with a pool of mortgages and who are often repaid in kind (the GSE pays the lender with MBS rather than in cash). These large lenders may hold the MBS on their balance sheets and may be servicers of their own mortgages. In such a situation, the GSE still guarantees the loan. However, servicing costs surely increase with delinquent or defaulting loans. Thus, if the GSE grid does not perfectly model the underlying credit risk (e.g., because it does not incorporate other fundamental variables such as wealth), a large lender might rationally implement a better pricing model using fundamental variables to estimate hidden servicing cost risk. Likewise, a large lender who plans to hold some balance sheet MBS prepayment risk that is not guaranteed by the GSE may implement a better-than-the-grid model to estimate prepayment risk. Since both of these actions would imply adjustment in prices for credit risk, they would be deemed *legitimate business necessity* by the courts.[20] In column (3) of Table 5, Panel (a) and (b), we estimate our full credit-risk specification, but limiting the sample to the non-top-25-volume lenders for every year. The results are not materially different from those in Table 2.

Finally, we note that there is the potential for inaccuracies in our estimation in discrimination due to errors in identifying borrower race or ethnicity. These are determined in our analysis by combining self-reported data from HMDA and, for mortgages in HMDA that lack an indicator of borrower race/ethnicity, borrower's likely race/ethnicity based on a race and ethnicity name-categorization algorithm from Kerr and Lincoln (2010) and Kerr (2008). Given the possibility that this algorithm might misclassify a borrower's race or ethnicity, the final column of Table 5 estimates the full credit risk model of Table 2 using only observations where a borrower's race or ethnicity is provided by HMDA. For both purchase mortgages (Panel (a)) and refinance mortgages (Panel (b)), our estimates for discrimination in pricing are slightly higher than the estimates obtained in Table 2, confirming that our results are not driven by potential errors in identifying a borrower's race or ethnicity using the name-categorization algorithm.

IV.d. Rejection Discrimination Estimates – Main

Table 6 reports the main estimation of discrimination in accept/reject decisions across all lenders for purchase (Panel (a)) and refinance (Panel (b)) applications. For each panel, we report three columns. Column (1) presents raw mean differences in rejection rates by minority treatment. Column (2) includes the GSE underwriting variables that HMDA provides at the loan level – income, loan amount, and year. Because we do not know the functional form of the underwriter, we include 21 piecewise-linear splines of income and 47 of loan amount, amounting to 136 variables in all. Column (3) adds census-tract-level proxies for the other GSE underwriting variables; namely, an applicant's LTV, credit score, debt-to-income,

[20] Note that minorities on average prepay less in aggregate statistics, which means our main estimate is probably conservative on this point of large lenders pricing prepayment risk differentially for minorities.

and total debt. We add three percentiles (the 25th, 50th, and 75th) of these variables to capture the distributional spread in the census tract.

The raw mean difference, reported in the column (1) specifications, indicate that minority applicants are more likely to be rejected by 14.1 percentage points in purchase applications and 12.2 percentage points in refinance applications compared to everyone else. Our estimation covers 3.2 million purchase applications and 10.0 million refinance applications. Of these 13.2 million applications, Table 1 reports that the overall likelihood of rejection is 49.6%. Decomposing these numbers, our minorities groups face rejection rates of greater than 60.6%, and everyone else, 47.6%.

Columns (2) and (3) present the main rejection results for all lenders, controlling for the underwriting variables used in the GSE algorithm. Focusing on column (3), lenders reject minority borrowers 9.6 percentage points more often for purchase loans and 7.3 percentage points for refinance loans. Across the sample of 13.2 million applications, applied across the purchase and refinance distribution from Table 1, our result would imply 1.03 million mortgages were rejected during 2009-2015 due to discrimination.

IV.e. Rejection Discrimination Estimates – FinTech

Table 7 reports the same rejection rate estimation specifications as Table 6, but only for the sample of FinTech lenders. As before, the overall raw mean rejection rates are higher for Latinx- and African-American applicants among FinTech lenders (5.3 percentage points higher in purchase applications and 5.4 percentage points higher in refinance applications). However, once we implement the models with the underwriting variables, we find that FinTech lenders do not discriminate in mortgage accept/reject decision-making. The evidence points to a story of FinTech lenders implementing little-to-no discrimination in rejection rates, consistent with the idea that algorithms are programmed to not leave money on the table.

IV.f. Rejection Discrimination Estimates –Robustness

Table 8 presents robustness tests addressing the concern that rejection discrimination in Table 6 stems from unobserved creditworthiness for lenders who keep balance sheet or servicing rights risk or from unobserved interactions between lenders and potential applicants ahead of formal application submission. As in the price analysis, we begin our robustness tests by utilizing the richness of our data in lender-crossed-with-county fixed-effects. Column (1) of Table 8 (Panel (a) for purchase mortgages and Panel (b) for refinance mortgages) reproduces the main accept/reject result from Table 6, Column (3), including all the underwriting variables. Columns (2), (3), and (4) add in, respectively, lender fixed effects, county fixed effects, and lender-crossed-with-county fixed effects. Interpreting column (4), we find that a given lender rejects Latinx and African-American applicants with a 7% higher probability for purchase applications and a 6.3% higher probability for refinance mortgages compared with the same lender's rejection decision for all other applicants in the same county. These econometrically appealing within-lender-county results are somewhat smaller than those of Column (1), but remain statistically and economically robust in magnitude.

Our other robustness concern involves unobserved creditworthiness, applicable only for lenders who might be holding loans on their balance sheet or for lenders who obtain servicing contracts for the mortgages once issued. As noted previously, large volume lenders may retain servicing rights in arrangements with the GSEs post-securitization (Aldrich et al. 2001). Likewise, it is unlikely that small lenders would be in a position to hold mortgage risk on their balance sheets. Thus, we repeat the analysis for Small-Volume Lenders, where we drop the top 25 originators calculated by year. In Column (5), we implement the most econometrically stringent model, with lender-crossed-with-geography fixed effects. We find results very similar to the results in column (4), which includes the full sample of lenders, despite the sample size being less than half as large.

To be conservative, we update our economic calculations to the most conservative interpretation from our robustness tests in Table (8). In sum, our results imply that between 0.74 million and 1.3 million mortgages were rejected during 2009-2015 due to discrimination, with minority applicants in our sample facing a 6% higher rejection rate due to discrimination.

Although we must maintain our caveat that we cannot claim perfect identification in our accept/reject analysis because we do not have access to the exact GSE underwriting model and because we must proxy for certain borrower characteristics, with this granularity of the fixed-effect model, the evidence seems compelling that discrimination exists in accept/reject decisions, but its incidence is diminished among FinTech lenders.

V. Conclusion

Using a unique data set of mortgage loans that includes never-before-linked information at the loan level on income, race, ethnicity, loan-to-value, and other contract terms, we exploit the unique structure of the GSE pricing grid to identify discrimination in mortgage loan pricing. Overall, we find that conditional on obtaining a loan, Latinx and African-American borrowers pay interest rates that are 7.9 bps higher for purchase mortgages and 3.6 bps higher for refinance mortgages. In addition, Latinx and African-American borrowers face higher hurdles in being accepted for a mortgage. Our evidence suggests that at least 6% of Latinx and African-American applications are rejected, but would have been accepted had the applicant not been in these minority groups. This amounts to a rejection of 0.74 to 1.3 million creditworthy minority applications.

Focusing on the effect of FinTech, we find that FinTech lenders discriminate approximately one-third less than lenders overall in terms of pricing. This finding is consistent with FinTech lenders removing discrimination arising from face-to-face interactions between originators and borrowers. Yet, it is also consistent with FinTech lenders using pricing strategies and data analytics that nevertheless produce discriminatory pricing. These results underscore the fact that even if algorithmic lending can reduce discrimination relative to face-to-face lenders, it is insufficient to eliminate discrimination in loan pricing.

We supplement these findings regarding discrimination among FinTech lenders with two additional silver linings associated with the emergence of FinTech lending. First, our evidence suggests that over our short time period, discrimination in loan pricing is declining, perhaps due to the ease of applying and shopping around afforded by the growth of FinTech platforms. Second, we also find that algorithmic lenders do not discriminate in accept/reject decisions. Thus, in addition to any efficiency gains of algorithmic innovations in credit scoring, our results suggest that these innovations may also serve to make the mortgage lending markets more accessible to African-American and Latinx applicants.

In this paper, we have also mapped the court's definition of a *legitimate business necessity* defense in lending discrimination cases to the signal extraction problem at the heart of statistical discrimination. In particular, discrmination in loan decisions must be rooted an applicant's creditworthiness and not the ability of a lender to extract rents. This insight points to a workable standard to deploy as lenders turn increasingly to Big Data tools to price and allocate credit. Operationally, in fair lending examinations, we propose a simple test as to whether a particular Big Data variable legitimately proxies for a borrower's creditworthiness and not a protected characteristic: lenders must demonstrate (a) that the Big Data variable (e.g., high school) is correlated with historical data relating to a fundamental lifecycle variable (e.g., income growth),[21] and (b) that this Big Data variable does not predict a protected characteristic after orthogonalizing it to the fundamental lifecycle variable. The use of a Big Data variable that passes such a test could thus be empirically validated as serving a *legitimate business necessity* given that any disparate impact associated with its use would arise solely through its correlation with the fundamental lifecycle variable of interest.

Finally, our results also speak to ongoing debates concerning the future structure of the GSEs. The GSE underwriting process that informs our identification strategy establishes clear rules for assessing borrower creditworthiness. Accordingly, it is possible that the GSE process itself may be serving to attenuate the incidence of discrimination, given that private lenders' benefit of greater use of variables is eliminated since the GSEs take on the credit risk of the mortgages. To date, this less-well-understood role of the GSEs has not been considered in GSE reform proposals, nor is it obvious how such a role could be supported within a fully privatized, conventional conforming secondary mortgage market. Likewise,

[21] Even if fundamental variables (e.g., income growth) are unobservable for a particular applicant, lending institutions can nevertheless observe how these variables are distributed in the population using historical data. For instance, a bank may have historical fundamental variable data from existing clients, such as historical income growth, but the bank will not have these data for new applicants.

outside of mortgage lending, it is possible that our estimates of discrimination are conservative, since these markets lack formal underwriting and pricing standards.

Appendix: Algorithm for Merging Mortgage Data Sets

Since there are no unique mortgage loan identifiers in the U.S., we develop an algorithm using machine-learning techniques to match loans found in two independent datasets: the McDash dataset, which contains loan-level data compiled by Black Night Financial Services, and the ATTOM dataset, which provides detailed property transaction and ownership information in addition to a time-series history of all recorded mortgage-lien events such as new mortgage originations, prepayments, REO, foreclosure, short sales, and arms-length sales and loan payoffs.

Our merging process applies a modified k-nearest neighbor classifier (see Hastie, Tibshirani, and Friedman, 2009; James, Witten, Hastie, and Tibshirani, 2015) using *sklearn.neighbors* in Python to fit radial kernels using *BallTree*. The k-nearest-neighbors classifier implements learning based on the 25 nearest neighbors in the corresponding zip code in the McDash data. We represent each loan in each data set with a thirteen element vector that includes: 1) the original loan balance; 2) the lien position, 3) the origination date of the loan, 4) the ending date of the loan, 5) the foreclosure date of the loan (maybe null), the prepayment date of the loan (maybe null), 6) the appraised market value of the property, 7) the loan purpose (refinance or purchase), 8) loan distress dates (may be null), 9) loan REO date (may be null), 10) loan liquidation date (may be null), 11) short sale indicator variable (may be null), 12) interest rate type (fixed or variable loans), 13) property transaction value if there is a sale. Each of these elements is assigned a category subscore between 0 and 1. Each subscore is then squared to achieve a greater penalty for matches on key elements such as the loan amount. The category subscore is then scaled by a factor which represents the category's importance to the match quality relative to other elements used in the match. Each category factor is an integer between 0 and 100, and the sum of the category factors is 100. Our scoring algorithm (*get.score* in Python's *sklearn*) takes into account the 13 different elements of each matched pair of loans to calculate a score. The score roughly corresponds to the estimated error for each match, measured in hundredths of a percent. Thus, a match score of 1689 corresponds to a 16.89% chance of an incorrect match, or an 83.11% confidence in the match. We use only matches with scores of 2000 or less. For fixed-rate GSE loans originated between 2009 and 2015, we obtain a 90% merge rate.

Our prior machine learning strategy is less applicable for the merge of the HMDA data to McDash data, because HMDA has a greatly reduced set of loan characteristics at origination and has no loan-level performance strings. For this merge, we instead standardize the lenders' names between ATTOM and

HMDA and then merge these data sets using lender names, loan amount, lien type, and the loan-purpose fields. For the final merge, we unite the ATTOM-to-McDash data to the ATTOM-to-HMDA data using the crosswalk developed with the k-nearest neighbor algorithm and we obtained a final data set of 6.8 million loans that are single-family fixed-rate GSE loans originated between 2008 and 2015 (3.5 million of those loans have maturities of 360 months).

Ethnicity matching using ATTOM data

Because there are missing ethnicity data in HMDA, we augment the HMDA ethnicity variable. We first apply the race and ethnicity name-categorization algorithm from Kerr and Lincoln (2010) and Kerr (2008) to assign an ethnicity to all the lien-holder names that are found in ATTOM. We then create an analysis subset of the ATTOM data that includes only the ATTOM crosswalk identification number and the ethnicity matches with no lien-holder name information. This data subset is then merged with the ATTOM-McDash-Equifax-HMDA merged data. We report our pricing results with both the HMDA only race/ethnicity indicators and with the enhanced HMDA race/ethnicity indicators using our ethnicity matches for the originated loans. The accept/reject estimations use only the HMDA race/ethnicity indicators.

The Equifax-enhanced subsample of originations

To obtain a final data that includes the full spectrum of underwriting characteristics that would have been available to the lender, we again merge the HMDA/ATTOM/McDash data set of fixed rate GSE loans that were originated between 2008 and 2015 to the McDash loans that are merged to Equifax data. The Equifax-enhanced originated loan sample includes other consumer credit positions of the borrowers such as: the total sum of retail, consumer finance and bank card balances; total student loan debt, total auto loan debt (sum or auto finance and auto bank debt); age of the borrower, and Vantage 3.0 score.

References

Abowd, J., F. Kramarz, and D. Margolis, 1999, High wage workers and high wage firms, Econometrica 67, 251–333.

Aigner, D. J., and G. Cain, 1977, Statistical theories of discrimination in the labor market, Industry and Labor Relations Review 30, 175–187.

Aldrich, S. P. B., W. R. Greenberg, and B. S. Payner, 2001, A capital markets view of mortgage servicing rights, Journal of Fixed Income 11, 37–53.

Arnold, D., W. Dobbie, and C. S. Yang, 2018, Racial bias in bail decisions, Quarterly Journal of Economics (forthcoming).

Arrow, K., 1973, Higher education as a filter, Journal of Public Economics 2, 193–216. Ayres, I., 2002, Outcome tests of racial disparities in police practices, Justice Research and Policy 4, 131–142.

Barocas, S., and A. Selbst, 2016, Big Data's Disparate Impact, California Law Review 104, 671-732.

Bayer, P., F. Ferreira, and S. L. Ross, 2016, The vulnerability of minority homeowners in the housing boom and bust, American Economic Review: Economic Policy 8, 1–27.

Bayer, P., F. Ferreira, and S. L. Ross, 2018, What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders, Review of Financial Studies 31, 175– 205.

Becker, G. S., 1957, The Economics of Discrimination (University of Chicago Press, Chicago).

Begley, T. A., and A. K. Purnanandam, 2017, Color and credit: Race, regulation, and the quality of financial services, Working paper, University of Michigan.

Black, H., R. L. Schweitzer, and L. Mandell, 1978, Wage discrimination: Reduced form and structural estimates, American Economic Review, Papers and Proceedings 68, 186–191.

Black, H. A., and R. Schweitzer, 1977, A canonical analysis of mortgage lending terms: Testing for lending discrimination at a commercial bank, Urban Studies 22, 13–19.

Blinder, A. S., 1973, Wage discrimination: Reduced form and structural estimates, Journal of Human Resources 8, 436–455.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru, 2017, Fintech, regulatory arbitrage, and the rise of shadow banks, Working Paper 23288, NBER.

Card, D., A. R. Cardoso, and P. Kline, 2016, Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women, Quarterly Journal of Economics 131, 633–686.

Elenev, V., T. Landvoigt, and S. van Nieuwerburgh, 2016, Phasing out the GSEs, Journal of Monetary Economics 81, 11–132.

Engel, K. C., and P. A. McCoy, 2011, The Subprime Virus: Reckless Credit, Regulatory Failure, and Next Steps (Oxford University Press, New York).

Federal Housing Finance Administration, 2009, Fannie Mae and Freddie Mac single-family guarantee fees in 2007 and 2008.

Federal Housing Finance Administration, 2010, Fannie Mae and Freddie Mac single-family guarantee fees in 2008 and 2009.

Federal Housing Finance Administration, 2011, Fannie Mae and Freddie Mac single-family guarantee fees in 2009 and 2010.

Federal Housing Finance Administration, 2012, Fannie Mae and Freddie Mac single-family guarantee fees in 2010 and 2011.

Federal Housing Finance Administration, 2013, Fannie Mae and Freddie Mac single-family guarantee fees in 2012.

Fortin, N., T. Lemieux, and S. Firpo, 2011, Decomposition methods in economics, in O. Ashenfelter, and D. Card, eds., Handbook of Labor Economics, volume 4A, chapter 1 (Elsevier, Amsterdam).

Fulks, M., 2017, Living with the robust causality requirement for disparate impact under the Fair Housing Act, Baker Donelson.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther, 2018, Predictably unequal? the effects of machine learning on credit markets, Working paper, Federal Reserve Bank of New York.

Fuster, A., L. Goodman, D. Lucca, L. Madar, L. Molloy, and P. Willen, 2013, The rising gap between primary and secondary mortgage rates, Federal Reserve Bank of New York Economic Policy Review 19.

Fuster, A., and P. Willen, 2010, $1.25 trillion is still real money: Some facts about the effects of the Federal Reserve's mortgage market investments, Public Policy Discussion Paper 10-4, Federal Reserve Bank of Boston.

Gano, A., 2017, Disparate impact and mortgage lending: A beginner's guide, University of Colorado Law Review 88, 1109–1166.

Ghent, A. C., R. Hernandez-Murillo, and M. T. Owyang, 2014, Differences in subprime loan pricing across races and neighborhoods, Regional Science and Urban Economics 48, 199–215.

Glassman, A. M., and S. Verna, 2016, Disparate impact one year after Inclusive Communities, Journal of Affordable Housing & Community Development Law 25, 12–24.

Gronau, Reuben, 1998, A useful interpretation of R-squared in binary choice models, Princeton University Working Paper.

Gruenstein Bocian, D., K. Ernst, and W. Li, 2008, Race, ethnicity and subprime home loan pricing, Journal of Urban Economics 60, 110–124.

Hanson, A., Z. Hawley, H. Martin, and B. Liu, 2016, Discrimination in mortgage lending: Evidence from a correspondence experiment, Journal of Urban Economics 92, 48–65.

Hastie, T., R. Tibshirani, and J. Friedman, 2009, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second edition (Springer, New York).

Haughwout, A., C. Mayer, and J. Tracy, 2009, Subprime mortgage pricing: The impact of race, ethnicity, and gender on the cost of borrowing, Staff Report 368, Federal Reserve Bank of New York.

Heckman, J., 1998, Detecting discrimination, Journal of Economic Perspectives 12, 101–116.

Heckman, J., and P. Siegelman, 1993, The Urban Institute audit studies: Their methods and findings, in M. E. Fix, and R. J. Struyk, eds., Clear and Convincing Evidence: Measurement of Discrimination in America, 187–258 (Urban Institute Press, Washington, DC).

James, G., D. Witten, T. Hastie, and R. Tibshirani, 2015, An Introduction to Statistical Learning with Applications in R (Springer, New York).

Kaye, D., 1982, Statistical evidence of discrimination, Journal of the American Statistical Association 77, 773–783.

Kerr, W. R., 2008, Ethnic scientific communities and international technology diffusion, Review of Economics and Statistics 90, 518–537.

Kerr, W. R., and W. F. Lincoln, 2010, The supply side of innovation: H-1B visa reforms and U.S. ethnic invention, Journal of Labor Economics 28, 473–508.

Maddala, G., and R. P. Trost, 1982, On measuring discrimination in loan markets, Housing Finance Review 1, 693–709.

Munnell, A. H., L. Browne, J. McEneaney, and G. Tootel, 1996, Mortgage lending in Boston: Interpreting HMDA data, American Economic Review 86, 25–54.

Phelps, E. S., 1972, Money, public expenditure and labor supply, Journal of Economic Theory 5, 69–78.

Rachlis, M., and A. Yezer, 1993, Serious flaws in statistical tests for discrimination in mort- gage markets, Journal of Housing Research 4, 315–336.

Reid, C.K., D. Bocian, W. Li, and R. G. Quercia, 2017, Revisiting the subprime crisis: The dual mortgage market and mortgage defaults by race and ethnicity, Journal of Urban Affairs, 39, 469-487.

Ross, S., 2002, Paired testing and the 2000 Housing Discrimination Survey, in W. Foster, A. F. Mitchell, and S. E. Fienberg, eds., Measuring Housing Discrimination in a National Study: Report of a Workshop (National Academies Press, Washington DC).

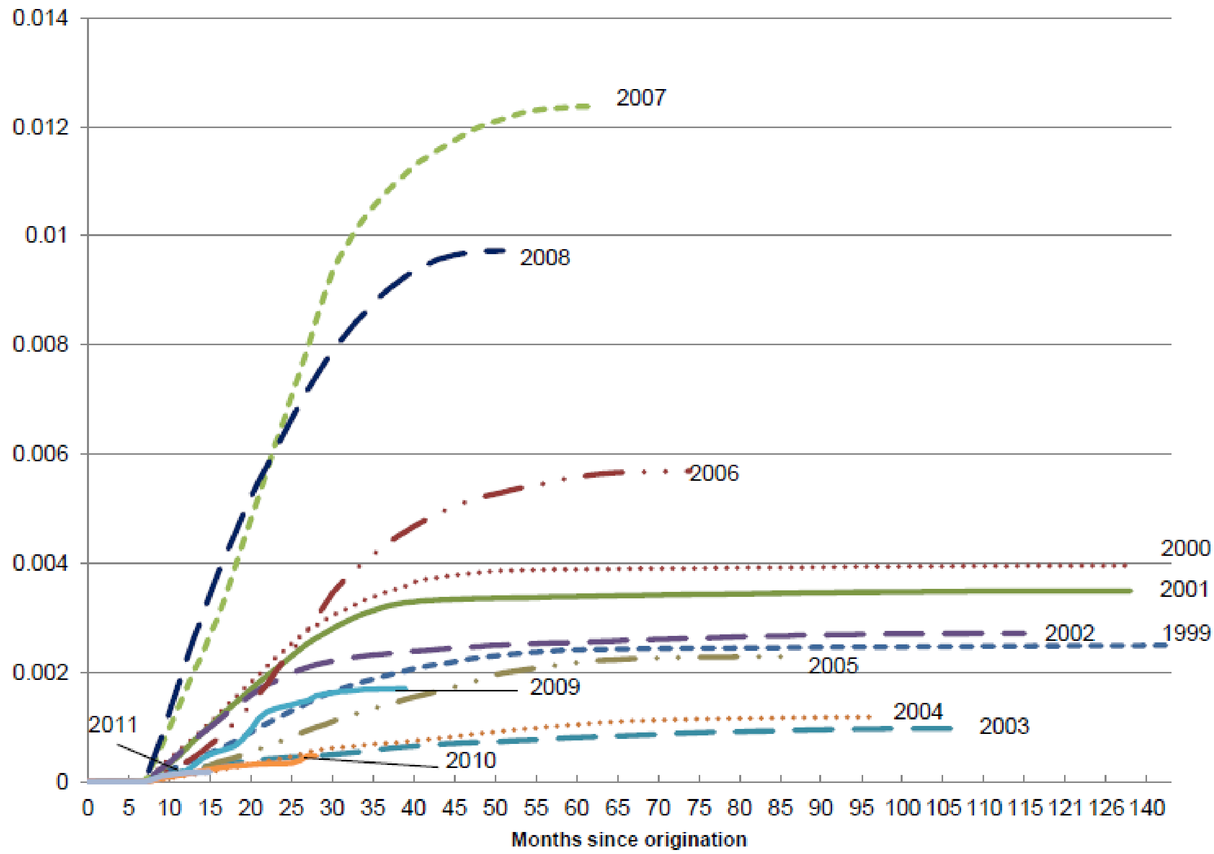
Ross, S. L., M. A. Turner, E. Godfrey, and R. R. Smith, 2008, Mortgage lending in Chicago and Los Angeles: A paired testing study of the pre-application process, Journal of Urban Economics 63, 902–919.

Sandler, A. L., and J. Biran, 1995, The improper use of statistics in mortgage lending discrimination actions, in A. Yezer, ed., Fair Lending Analysis (American Bankers Association, Washington, DC).

Shafer, R., and H. Ladd, 1981, Discrimination in Mortgage Lending (MIT Press, Cambridge, MA).

Smith, R., and M. Delair, 1999, New evidence from lender testing: Discrimination at the pre- application stage, in M. Turner, and F. Skidmore, eds., Mortgage Lending Discrimination: A Review of Existing Evidence (Urban Institute Press, Washington, DC).

Vickery, J. I., and J. Wright, 2013, TBA trading and liquidity in the agency MBS market, Economic Policy Review 19.



Sources: Fannie Mae and Freddie Mac credit database and Urban Institute calculations.

Figure 1: Put-Backs for Issuances 2000 - 2010

> *Source: Goodman, Laurie S., and Jun Zhu, 2013. "Reps and Warrants: Lessons from the GSEs Experience". Urban Institute: Housing Policy Center White Paper.*

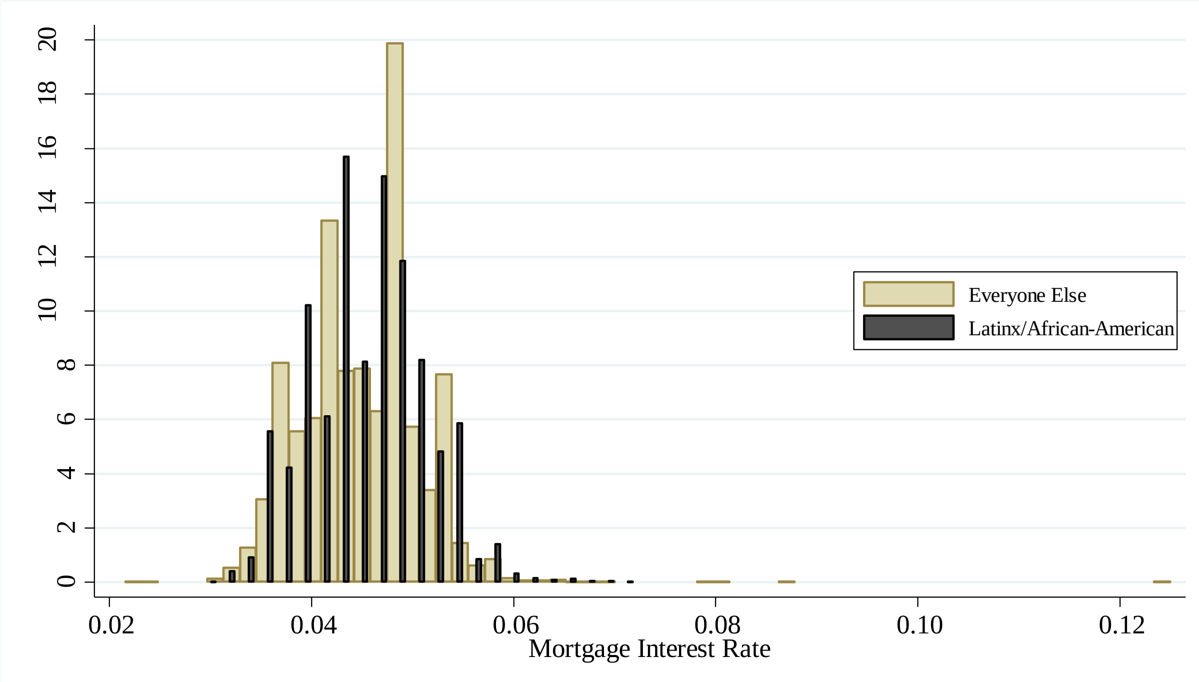
Presented is a copy of Figure 2 from the aforementioned Urban Institute White Paper (permission granted in copyright.). The Figure shows the dollars put back on Freddie Mac loans by issuance vintage. The take-away of the figure for our purposes are twofold. First is the small value of put-backs after 2008. Second is the low volume of put-backs in the first 30 days of loan life, when lenders remain exposed to credit-risk for originated loans.

Table 2: All Eligible Mortgages (Excluding MCM): LLPA by Credit Score/LTV									
PRODUCT FEATURE	LLPAs by LTV Range								
	≤ 60.00%	60.01 – 70.00%	70.01 – 75.00%	75.01 – 80.00%	80.01 – 85.00%	85.01 – 90.00%	90.01 – 95.00%	95.01 – 97.00%	SFC
Representative Credit Score	Applicable for all mortgages with greater than 15 year terms For whole loans purchased on or before March 31, 2011, or loans delivered into MBS pools with issue dates of March 1, 2011 or earlier								
≥ 740	-0.250%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	N/A
720 – 739	-0.250%	0.000%	0.000%	0.250%	0.000%	0.000%	0.000%	0.000%	N/A
700 – 719	-0.250%	0.500%	0.500%	0.750%	0.500%	0.500%	0.500%	0.500%	N/A
680 – 699	0.000%	0.500%	1.000%	1.500%	1.000%	0.750%	0.750%	0.500%	N/A
660 – 679	0.000%	1.000%	2.000%	2.500%	2.250%	1.750%	1.750%	1.250%	N/A
640 – 659	0.500%	1.250%	2.500%	3.000%	2.750%	2.250%	2.250%	1.750%	N/A
620 – 639	0.500%	1.500%	3.000%	3.000%	3.000%	2.750%	2.750%	2.500%	N/A
< 620 (1)	0.500%	1.500%	3.000%	3.000%	3.000%	3.000%	3.000%	3.000%	N/A

Figure 2: An Example of the GSE Grid

Presented is the LLPA (Loan-Level Price Adjustment) Grid of Fannie Mae for 2011. The figure is from the Fannie Mae Selling Guide, dated 12/23/2010. (MCMs, now retired, refers to "My Community Mortgages", a program of subsidized loans for low-income target areas.) The LLPA Grid has a parallel grid at Freddie Mac called the Credit Fees in Price chart. These grids provide the additional g-fee (guarantee fee) that lenders must pay the GSE for guaranteeing the mortgage, varying by LTV and credit score. In practice, these lump-sum fees are translated to flows concepts to be added to the interest rate passed on to borrowers to pay for credit risk.

Panel A: Raw Interest Rates by Race/Ethnicity



Panel B: Excess Interest Rates over Month-Year GSE Grid Rate by Race/Ethnicity

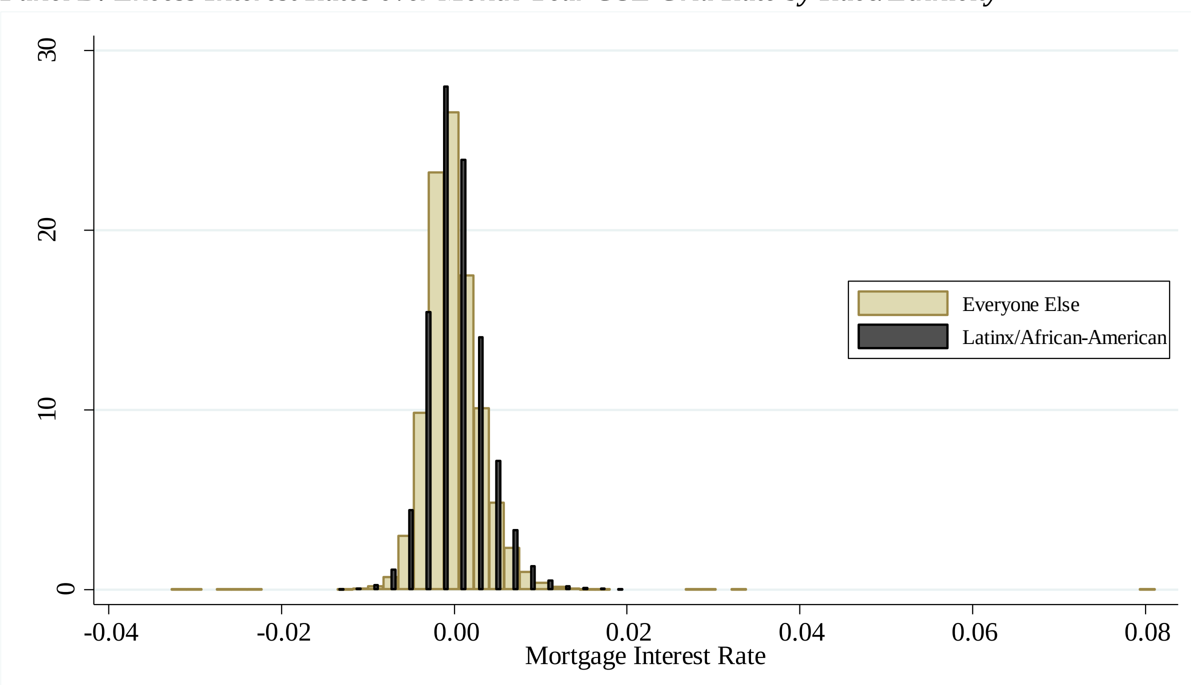


Figure 3: Interest Rate Histograms by Race/Ethnicity: The Role of the GSE Grid

Presented are two histograms of interest rates originated on 30-year fixed-rate mortgages, 2009-2015, that are processed in the GSE system. Panel A shows the raw data histogram of interest rates. Panel B de-means the histogram to the GSE grid for the month and year. The histograms are plotted for Latinx and African-Americans and for everyone else.

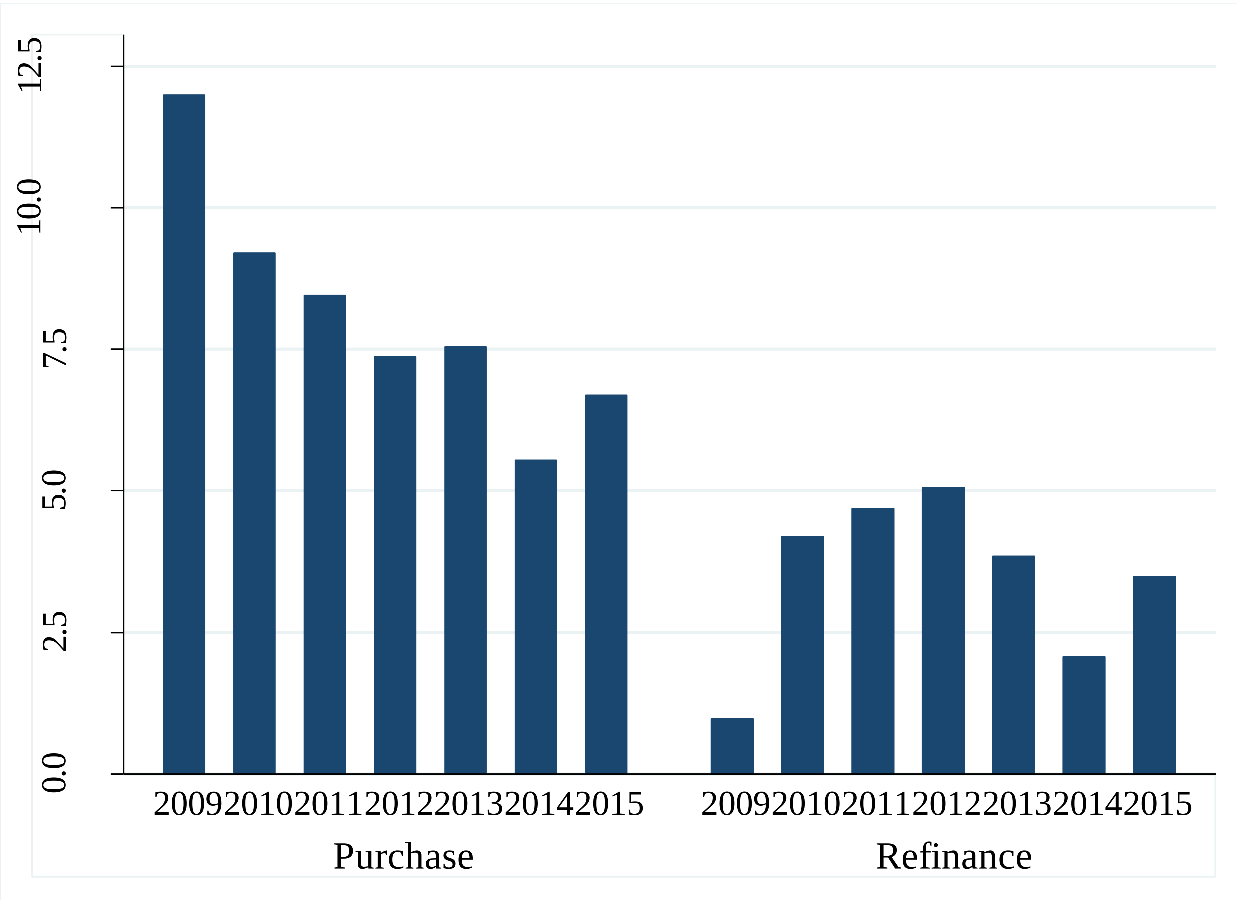
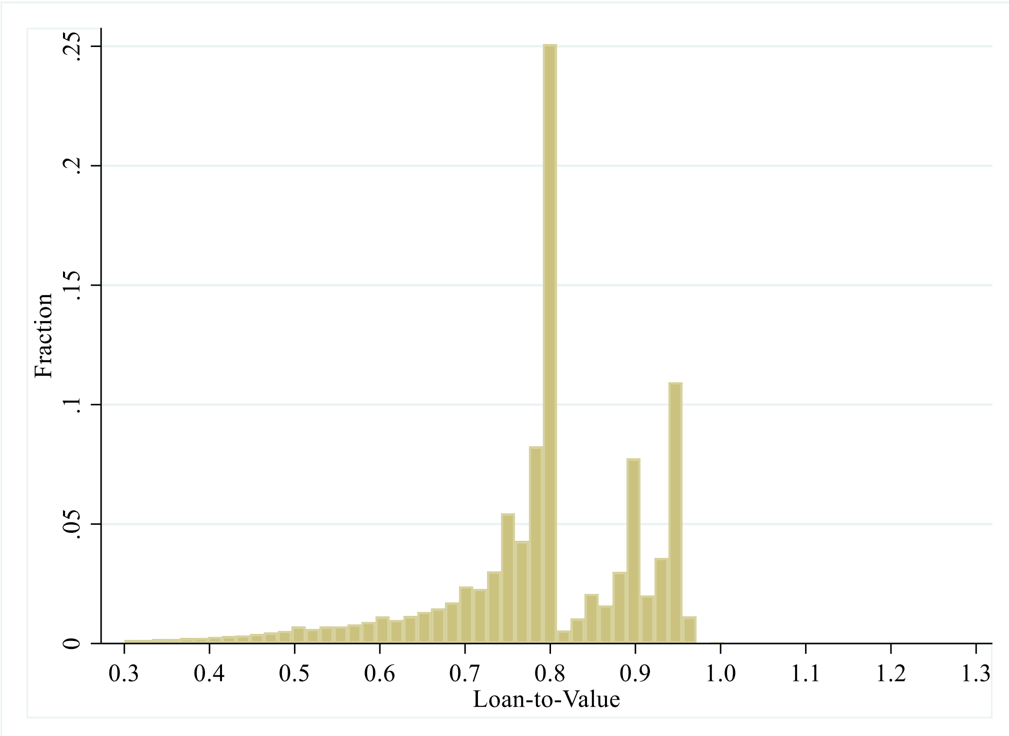


Figure 4: Interest Rate Discrimination Estimates by Year

Plotted are the purchase and mortgage discrimination estimates (the beta coefficient on Latinx/African-American) by year for the credit risk model of interest rate discrimination (akin to Table 2, columns 2 and 4, but by year). The sample is all loans for 30-year fixed-rate mortgages, securitized through the GSE system, 2009-2015. The estimation regresses interest rates on the GSE grid fixed effects and month-year fixed effects. Estimates are converted to basis points (1 basis point =0.1%) for ease of conveyance.

Panel A: Purchase Mortgage



Panel B: Refinance Mortgage

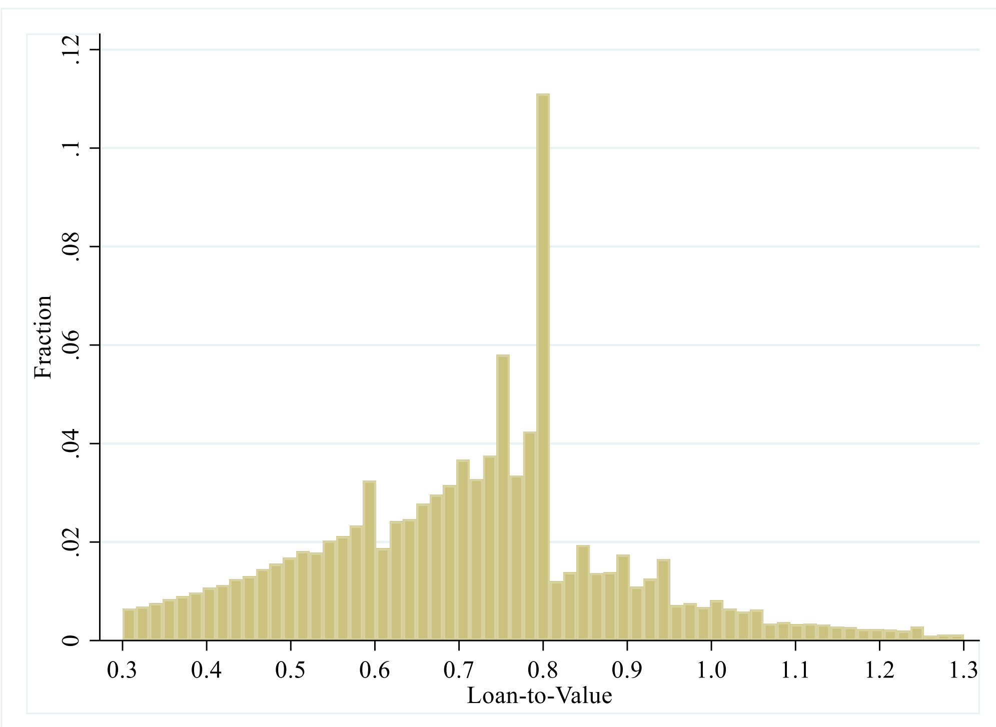


Figure 5: Histogram of GSE Purchase and Refinance Mortgages across Loan-to-Value

Panels A and B present the distribution of purchase and refinance mortgages by loan-to-value ratios. The sample is 2009-2015, 30-year fixed mortgages, securitized through the GSE system.

Table 1: Summary Statistics

Panel A reports summary statistics for the pricing estimations. Data are GSE, 30-year fixed rate mortgage originations obtained from a loan-level merge of HMDA, ATTOM, McDash, and Equifax data. Loan amount, applicant income and Latinx-/African-American are from HMDA. Interest rate, LTV, and credit score are from McDash-Equifax. Top 25 Volume Lender is calculated annually from volume of loans by lender. FinTech is a platform identifier from Buchak et al (2017). Panels B and C report statistics for accept/reject analyses. Because we only have HMDA data for rejections, we proxy for LTV, credit score, debt outstanding, and debt-to-income ratio using census tract medians. For all panels, Top 25 Volume Lender is calculated annually from the volume of loans by lender. FinTech is a platform identifier from Buchak et al (2017).

Panel A: For Pricing Analysis: GSE, 30-Year Fixed Rate Mortgage Acceptances (N = 3,577,010)					
	Mean	St. Deviation	Minimum	Median	Maximum
Interest Rate % (McDash)	4.50%	0.56%	2.00%	4.50%	12.50%
Loan Amount $,000	$234.0	$122.6	$30.0	$210.0	$729.0
Applicant Income $,000	$107.2	$92.0	$19	$89	$9,980
Credit Score (McDash-Equifax)	755.8	43.4	620	766	850
Loan-to-Value(McDash-Equifax)	0.744	0.165	0.300	0.774	1.300
FinTech	0.043	0.203			
Top 25 Lender	0.523	0.499			
Latinx-/African-American	0.110	0.313			
Purchase=1; Refinance=0	0.418	0.493			
Panel B: For Accept/Reject Analysis: Conventional Mortgage Acceptances (N = 6,648,413)					
	Mean	St. Deviation	Minimum	Median	Maximum
Loan Amount $,000	$213.9	$114.7	$30.0	$191.0	$729.0
Applicant Income $,000	$108.3	$103.2	$19	$89	$9,999
Credit Score (census tract)	750.8	24.1	620	756	832
Loan-to-Value (census tract)	0.791	0.101	0.300	0.799	1.283
Debt Outstanding (census tract)	18,180	8,145	0	17,739	529,506
Debt-to-Income% (census tract)	32.7	3.6	1.0	32.5	63.0
FinTech	0.042				
Top 25 Lender	0.522				
Latinx-/African-American	0.119				
Purchase=1; Refinance=0	0.308				
Panel C: For Accept/Reject Analysis: Conventional Mortgage Rejections (N = 6,535,664)					
	Mean	St. Deviation	Minimum	Median	Maximum
Loan Amount $,000	$187.3	$101.2	$30.0	$166.0	$428.0
Applicant Income $,000	$97.4	$129.7	$19	$75	$9,999
Credit Score (census tract)	744.2	26.9	620.0	749.0	830.0
Loan-to-Value (census tract)	0.812	0.099	0.300	0.800	1.283
Debt Outstanding (census tract)	18,322	9,290	0	17,715	513,857
Debt-to-Income% (census tract)	32.9	3.9	1.0	33.0	61.0
FinTech	0.055				
Top 25 Volume Lender	0.515				
Latinx-/African-American	0.186				
Purchase=1; Refinance=0	0.173				

Table 2: Interest Rate Discrimination

Panel A reports discrimination results using the GSE grid for identification. The dependent variable is the interest rate on origintated GSE 30-year fixed-rate mortgages. Estimates for purchase mortgages are in Columns (1) and (2); estimates for refinances are in Columns (3) and (4). Columns (1) and (3) report raw differences in means, as a starting point for understanding the role of the credit risk model. Columns (2) and (4) report discrimination estimates for the full credit risk model. We regress the interest rate on the GSE grid fixed effects and the month-year effects; identifying discrimination as the estimate on a Latinx/African-American indicator variable. Standard errors are clustered at the lender level. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels. Panel B presents economic magnitude calculations, aggregating the findings in Panel A to the mortgage market outstanding as of the end of 2018. Latinx-/African-American percentage representation in the mortgage market float (17.3%) is from the Survey of Consumer Finances. The aggregate housing debt is from the Federal Reserve Bank of New York.

Panel A: Estimates

	Dependent Variable: Mortgage Interest Rate			
	Purchase Mortgages		*Refinance Mortgages*	
	(1)	(2)	(3)	(4)
Latinx-/African-American	0.000903***	0.000788***	0.000298***	0.000356***
	[0.000102]	[3.11e-05]	[7.98e-05]	[2.92e-05]
Observations	1,495,021	1,495,021	2,081,989	2,081,807
R-squared	--	0.729	--	0.694
Month-Year FE	N	Y	N	Y
GSE Grid FE	N	Y	N	Y

Panel B: Economic Magnitude Calculation

A . Market Size of Housing Debt (Federal Reserve of New York) ($M)	$9,536,000
B . African-American/Latinx % of the Float in Mortgage Market (SCF data)	0.173
C . Extra Interest Payments per Basis Point of Discrimination ($M)	$164.97
Discrimination Estimates from Panel (A):	
D . Extra Interest Rate (bps): Purchase Mortgages (estimate from column 2)	7.88
E . Extra Interest Rate (bps): Refinance Mortgages (estimate from column 4)	3.56
F . Share of refinance loans in stock of float	0.75
G . Weighted average extra interest rate (= D*(1 -F) + E*F)	4.64
Annual Aggregate Extra Interest Paid by Latinx-/African-Americans (= C*G) ($M)	$765.47

Table 3: Interest Rate Discimination - FinTech Results

This table replicates our main credit risk specification, but only for the subsample of FinTech lenders. The columns reproduce the specification of column (2) and (4) of Table 2, regressing interest rates on the GSE grid dummy variables, month-year effects, and an indicator for whether the borrower is Latinx- or African-American. The sample is the list of FinTech platforms from Buchak et al (2017). Column (1) reports purchase-mortgage estimates, and column (2) reports refinance mortgages. Standard errors clustered at the lender level are in brackets. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels.

	Dependent Variable: Mortgage Interest Rate	
	Purchase	*Refinance*
Sample:	*FinTech Lender*	*FinTech Lender*
	(1)	(2)
Latinx-/African-American	0.000531***	0.000197**
	[4.22e-05]	[6.20e-05]
Observations	42,318	111,912
R-squared	0.729	0.707
Year FE	Y	Y
GSE Grid FE	Y	Y

Table 4: Interest Rate Discrimination - Robustness to Lender and Geography Fixed Effects

This table mitigates the concern that our Table 2 estimates are picking up differential costs of delivering a mortgage by lender or by geography. Panels A and B report estimates for interest rate discrimination for purchase and refinance mortgages, respectively. Column (1) repeats the OLS estimate of Table 2, regressing interest rates on the GSE grid-dummy variables, month-year fixed effects, and an indicator for whether the borrower is ethnically Latinx or African-American. This is the main credit risk model. Column (2) adds county fixed effects; column (3) adds lender fixed effects; and column (4) adds lender crossed with county fixed effects. Standard errors in brackets are clustered at the lender level. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels.

Panel A: Purchases

	Dependent Variable: Mortgage Interest Rate			
	Varying Fixed Effects			
	(1)	(2)	(3)	(4)
Latinx-/African-American	0.000788***	0.000695***	0.000545***	0.000516***
	[3.11e-05]	[2.41e-05]	[2.67e-05]	[3.26e-05]
Observations	1,495,021	1,495,005	1,493,797	1,468,357
R-squared	0.729	0.733	0.748	0.759
Month-Year FE	Y	Y	Y	Y
GSE Grid FE	Y	Y	Y	Y
County FE	N	Y	Y	Y
Lender FE	N	N	Y	Y
County x Lender FE	N	N	N	Y

Panel B: Refinances

	Dependent Variable: Mortgage Interest Rate			
	Varying Fixed Effects			
	(1)	(2)	(3)	(4)
Latinx-/African-American	0.000356***	0.000364***	0.000223***	0.000202***
	[2.92e-05]	[2.80e-05]	[2.19e-05]	[1.78e-05]
Observations	2,081,807	2,081,798	2,080,699	2,052,246
R-squared	0.694	0.697	0.712	0.721
Month-Year FE	Y	Y	Y	Y
GSE Grid FE	Y	Y	Y	Y
County FE	N	Y	Y	Y
Lender FE	N	N	Y	Y
County x Lender FE	N	N	N	Y

Table 5: Interest Rate Discimination - Robustness

This table addresses robustness concerns for interpreting disrimination in pricing results. All specifications use the formulation of the main credit risk model of Table 2, columns (2) and (4), which include GSE-grid dummies and month-year fixed effects. Panels A and B present results for purchase and refinance mortgages, respectively. Column (1) address the possibility that our loading on the Latinx/African-American variable is due to points paid by non-minority borrowers. We limit the sample to mortgages with LTVs precisely at 0.8 and borrower's whose non-mortgage debt is greater than 30% of income. These are on average likely to be borrowers fully utilizing their cash for down payment, leaving no cash to pay (positive) points. Column (2) addresses the possibility that our results are driven by minority borrowers taking negative points (a rebate) to pay closing costs. Borrowers who have LTVs between (0.70, 0.74) and (0.75, 0.795) are not at the edge of the GSE grid bucket in LTV. These borrowers could pay slightly less in down payment without incurring an interest rate increase, so that taking a negative point rebate for closing costs would be suboptimal. Column (3) restricts the analysis to only non-top-25-volume lenders, addressing robustness of our results to the concern that large lenders retain the MBS and servicing rights after securitizing through the GSEs. Small lenders do not service GSE loans. Column (4) drops borrowers who do not designate their race or ethnicity directly in HMDA. Standard errors clustered at the lender level are in brackets. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels. mean (0.795 to 0.801).

Dependent Var:	Mortgage Interest Rate			
	(1)	(2)	(3)	(4)
Robustness Concern:	*Positive Points*	*Negative Points*	*Residual Risk via Servicing or MBS Holding*	*Ethnicity Designation*
Sub-Sample:	*0.795 < LTV < 0.801 and Other Debt/Income>0.3*	*LTV not near Grid Cell Maximum*	*Small Lenders*	*Only use HMDA-Classified Ethnicity Observations*
Reasoning:	*At budget constraint*	*Borrower will not face higher interest rate for slightly larger loan to cover closing costs*	*Unlikely to service the loans or hold as MBS on balance sheet*	*Eliminate software errors in race/ethnicity classification*
Panel A: Purchases				
Latinx-/African-American	0.000873***	0.000938***	0.000865***	0.000809***
	[5.27e-05]	[4.02e-05]	[3.50e-05]	[3.26e-05]
Observations	241,847	337,115	846,547	1,370,384
R-squared	0.749	0.725	0.729	0.729
Month-Year FE	Y	Y	Y	Y
GSE Grid FE	Y	Y	Y	Y
Panel B: Refinances				
Latinx-/African-American	0.000251***	0.000333***	0.000384***	0.000384***
	[5.96e-05]	[4.28e-05]	[2.21e-05]	[3.19e-05]
Observations	102,627	380,625	859,715	1,857,191
R-squared	0.746	0.705	0.732	0.695
Month-Year FE	Y	Y	Y	Y
GSE Grid FE	Y	Y	Y	Y

Table 6: Application Rejection Discrimination

This table reports discrimination in rejection rates for mortgages in our full sample of HMDA mortgages. The dependent variable is an indicator for an application being rejected by the lender. The sample is the set of all HMDA 30-year, fixed-rate mortgage applications run through GSE underwriting. The credit-risk model controls consist of three sets of variables, with the inclusion noted beneath the estimation. The first column presents raw mean differences. Column (2) presents the full set of HMDA data publicly available, which includes the year, applicant income, and applicant loan amount, including 21 piecewise-linear functions of income and 47 of loan amount. Column (3) includes the full set of variables used in the black box of the GSE underwriter system that determines GSE acceptability of applications, which adds in census tract proxies for LTV, credit score, debt-to-income (DTI), and total debt. We include quartiles (25th, 50th and 75th percentiles) of each variable to capture within-tract dispersion. We construct these census tract variables from McDash (for LTV, total debt, and credit score) and from the GSE data (for DTI). Standard errors clustered at the lender level are in brackets. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels.

Panel A: Purchase Mortgages

	Dependent Variable: Application Rejection		
	(1)	(2)	(3)
Latinx/African-American	0.141***	0.109***	0.0957***
	[0.00510]	[0.00508]	[0.00448]
Observations	3,179,813	3,179,813	3,179,273
R-squared	--	0.055	0.063
Application Income & Application Loan Amount: 68 piecewise-linear splines	N	Y	Y
Year FE	N	Y	Y
LTV, Credit Score, Debt-to-Income, and Total Debt: Quartile Variables by Census Tract	N	N	Y

Panel B: Refinance Mortgages

	Dependent Variable: Application Rejection		
	(1)	(2)	(3)
Latinx/African-American	0.122***	0.0896***	0.0728***
	[0.00867]	[0.00577]	[0.00505]
Observations	10,004,264	10,004,264	10,002,727
R-squared	--	0.049	0.055
Application Income & Application Loan Amount: 68 piecewise-linear splines	N	Y	Y
Year FE	N	Y	Y
LTV, Credit Score, Debt-to-Income, and Total Debt: Quartile Variables by Census Tract	N	N	Y

Table 7: Application Rejection Discrimination - FinTech Lenders

This table reports discrimination in rejection rates among FinTech lenders for mortgages in our full sample of HMDA mortgages. The dependent variable is an indicator for an application being rejected by the lender. The sample is the set of stand-alone FinTech platforms from Buchak et al (2017). The credit risk model controls consists of three potential sets of variables, with the inclusion noted beneath the estimation. The first column presents raw mean differences. Column (2) presents the full set of HMDA data publicly available, which includes the year, applicant income, and applicant loan amount, including 21 piecewise-linear functions of income and 47 of loan amount. Column (3) includes the full set of variables used in the black box of the GSE underwriter system that determines GSE acceptability of applications, adding in census tract proxies for LTV, credit score, debt-to-income (DTI), and total debt. We include quartiles (25th, 50th and 75th percentiles) of each variable to capture within-tract dispersion. We construct these census-tract variables from McDash (for LTV, total debt, and credit score) and from the GSE data (for DTI). Standard errors clustered at the lender lever are in brackets. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels.

Panel A: Purchase Mortgages

	Dependent Variable: Application Rejection		
	Sample: FinTech Lenders		
	(1)	(2)	(3)
Latinx/African-American	0.0527*	0.0411	0.0328
	[0.0281]	[0.0224]	[0.0220]
Observations	70,813	70,813	70,791
R-squared	--	0.043	0.048
Application Income & Application Loan Amount: 68 piecewise-linear splines	N	Y	Y
Year FE	N	Y	Y
LTV, Credit Score, Debt-to-Income, and Total Debt: Quartile Variables by Census Tract	N	N	Y

Panel B: Refinance Mortgages

	Dependent Variable: Application Rejection		
	Sample: FinTech Lenders		
	(1)	(2)	(3)
Latinx/African-American	0.0540**	0.0288*	0.0233
	[0.0213]	[0.0154]	[0.0132]
Observations	337,582	337,582	337,508
R-squared	--	0.052	0.058
Application Income & Application Loan Amount: 68 piecewise-linear splines	N	Y	Y
Year FE	N	Y	Y
LTV, Credit Score, Debt-to-Income, and Total Debt: Quartile Variables by Census Tract	N	N	Y

Table 8: Application Rejection Discrimination - Lender & Location Fixed Effects

This table reports robustness tests for our analysis of discrimination in rejection rates. The dependent variable is an indicator for an application being rejected by the lender. Column (1) repeats the specification from the third column of Table 6. Column (2) adds in lender fixed effects. Column (3) includes county and lender fixed effects. Column (4) includes lender crossed with county fixed effects. Column (5) repeats the column (4) specification but limits the sample to only small lenders. Panel (A) presents purchase mortgage application results, and Panel (B) presents results for refinance applications. Standard errors clustered at the lender level are in brackets. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels.

Panel A: Purchase Mortgages

	Dependent Variable: Application Rejection				
	(1)	(2)	(3)	(4)	(5)
Sample Restriction:					Small Lenders
Latinx/African-American	0.0957***	0.0745***	0.0714***	0.0698***	0.0657***
	[0.00448]	[0.00342]	[0.00319]	[0.00311]	[0.00223]
Observations	3,179,273	3,178,331	3,178,324	3,137,844	1,838,121
R-squared	0.063	0.224	0.227	0.264	0.337
Application Income & Application Loan Amount: 68 piecewise-linear splines	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
LTV, Credit Score, Debt-to-Income, and Total Debt: Quartile Variables by Census Tract	Y	Y	Y	Y	Y
Lender FE	N	Y	Y	N	N
County FE	N	N	Y	N	N
Lender # County FE	N	N	N	Y	Y

Panel B: Refinance Mortgages

	Dependent Variable: Application Rejection				
	(1)	(2)	(3)	(4)	(5)
Sample Restriction:					Small Lenders
Latinx/African-American	0.0728***	0.0617***	0.0628***	0.0633***	0.0527***
	[0.00505]	[0.00484]	[0.00477]	[0.00476]	[0.00285]
Observations	10,002,727	10,001,968	10,001,964	9,954,232	4,415,763
R-squared	0.055	0.164	0.167	0.188	0.327
Application Income & Application Loan Amount: 68 piecewise-linear splines	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
LTV, Credit Score, Debt-to-Income, and Total Debt: Quartile Variables by Census Tract	Y	Y	Y	Y	Y
Lender FE	N	Y	Y	N	N
County FE	N	N	Y	N	N
Lender # County FE	N	N	N	Y	Y

Appendix Table 1: Statistics & Estimation of Interst Rate Discrimination for Shorter Maturities

This table reports statisitics and analysis on GSE-issued conforming mortgages, with maturities of less than 30 years. Panel A reports statistics akin to Table 1, panel A except for the shorter maturity sampling. Data are from a loan-level merge of HMDA, ATTOM, McDash, and Equifax data. Loan amount, applicant income and Latinx-/African-American are from HMDA. Interest rate, LTV, and credit score are from McDash-Equifax. Top 25 Volume Lender is calculated annually from volume of loans by lender. FinTech is a platform identifier from Buchak et al (2017). Panel B reports discrimination results using the GSE grid for identification. The dependent variable is the interest rate. Columns (1) and (3) report raw differences in means, as a starting point for understanding the role of the credit risk model. Columns (2) and (4) report discrimination estimates for the full credit risk model. We regress the rate on the GSE grid fixed effects and the month-year effects, identifying discrimination as the estimate on an Latinx/African-American indicator variable. Standard errors are clustered at the lender level. ***, **, and * indicate significance at the 1%, 5%, and 10% conventional levels.

Panel A: Statistics for Accepted GSE <30 Year Fixed Rate Mortgages (N = 1,390,286)

	Mean	Standard Deviation	Minimum	Median	Maximum
Interest Rate % (McDash)	3.94%	0.64%	2.00%	3.88%	7.63%
Loan Amount $,000	$185.4	$100.1	$30.0	$163.0	$729.0
Applicant Income $,000	$111.4	$101.5	$19	$91	$9,600
Credit Score (McDash-Equifax)	758.6	45.6	620	772	850
Loan-to-Value(McDash-Equifax)	0.669	0.177	0.300	0.685	1.300
FinTech	0.050				
Top 25 Lender	0.591				
Latinx-/African-American	0.113				
Purchase=1; Refinance=0	0.101				

Panel B: Estimates of Interest Rate Discrmination for Shorter Maturity Originations

	Dependent Variable: Mortgage Interest Rate			
	Purchase Mortgages		*Refinance Mortgages*	
	(1)	(2)	(3)	(4)
Latinx-/African-American	0.00134***	0.00117***	0.000292**	0.000555***
	[0.000249]	[0.000162]	[0.000137]	[6.85e-05]
Observations	140,613	140,613	1,249,673	1,249,634
R-squared	--	0.674	--	0.6
Month-Year FE	N	Y	N	Y
GSE Grid FE	N	Y	N	Y