Consumer-Lending Discrimination in the FinTech Era*

Robert Bartlett† Adair Morse‡ Richard Stanton§ Nancy Wallace¶

March 22, 2018

Abstract

Racial discrimination in lending can materialize in loan officer decisions or in algorithmic scoring, especially with big-data use by FinTech lenders. To investigate these discrimination channels, we estimate a treatment-based Oaxaca-Blinder decomposition based on the unique mortgage-default-risk setting of the GSEs. Among approved loans, ethnic-minority borrowers pay higher rates of interest, with both traditional and FinTech lenders charging non-white borrowers 0.08% higher interest for purchase mortgages, the present value of which represents nearly 1% of the loan balance. In aggregate, ethnic-minority borrowers pay almost $0.5B per year in extra interest. These findings are consistent with FinTech and non-FinTech lenders extracting rents in weaker competitive environments such as financial deserts. We additionally estimate that lenders reject African-American and Hispanic applicants 5% more often, leaving money on the table. Discrimination in rejection rates is especially pronounced among low-credit-score applicants, but less pronounced for FinTech lenders. These latter results suggest that GSE underwriting plays a crucial role in minimizing algorithmic discrimination, at least with respect to credit rationing.

Key words: Discrimination; FinTech; mortgages; credit scoring; algorithmic underwriting; big-data lending; platform loans; disparate impact

JEL codes: G21, G28, G23, J14, K22, K23, R30

*Financial support from the Fisher Center for Real Estate and Urban Economics is gratefully acknowledged. For helpful comments and suggestions, we thank Sanjiv Das, Raghu Rau, and seminar participants at U.C. Berkeley, U. Michigan, USC, the 2017 Boston JOIM Conference, and the 2017 Annual Research Conference of the Center for International Securities and Derivatives Markets.

†Berkeley Law, U.C. Berkeley, rbartlett@berkeley.edu.

‡Haas School of Business, U.C. Berkeley and NBER, morse@haas.berkeley.edu.

§Haas School of Business, U.C. Berkeley, stanton@haas.berkeley.edu.

¶Haas School of Business, U.C. Berkeley, wallace@haas.berkeley.edu.

1 Introduction

Credit scoring has long been the first step in lenders' processes for approving and pricing household loans. Recently, technology-enabled "FinTech" loan companies have sought to drive significant cost reductions in the underwriting process by improving scoring precision with big-data algorithms. Beyond costs, algorithms reduce the influence of humans, who can produce discriminatory loan decisions with respect to race, gender, and ethnicity. However, algorithms have a potential dark side; by using big data to proxy for legitimate repayment ability variables (e.g., hidden wealth), they may introduce illegitimate statistical discrimination. This potential dark side has not gone unnoticed. 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. Whether FinTech solves an age-old problem of human prejudice or introduces algorithmic discrimination remains an open question, a question that will only grow in importance as more data become readily available.

In this paper, we estimate the level of ethnic discrimination among conventional and FinTech lenders. We focus our examination on the mortgage market, which represents 71% of the $12.73 trillion household loan market as well as the largest market for FinTech loans.[1] The analysis uses a data set that includes never-before-linked information at the loan-level on income, ethnicity, debt-to-income ratios, loan-to-value 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.

To identify the presence of ethnic discrimination, we first need a landscape of what is legal and what is not. Discrimination comes in three types. The first, direct illegal discrimination (termed *disparate treatment* in law), happens when lenders' loan-offer decisions differ explicitly by ethnicity, either in their algorithmic processes (e.g., ethnicity is a scoring variable) or in personnel biases (including both cognitively aware and unaware biases). Redlining, the use of geography to proxy for ethnicity, is also deemed directly illegal under *disparate treatment*.

The second and third types of discrimination emerge from statistical discrimination techniques. Imagine a lender attempting to evaluate a loan applicant's repayment risk. Variables which structurally affect repayment in expectation are life cycle model variables — income, debt, wealth, expenses, and derivatives from these variables. These are legitimate (struc-

[1] See `https://www.nytimes.com/2017/05/17/business/dealbook/household-debt-united-states.html`.

tural) variables for lender algorithms. The lender cannot, however, observe all of these variables; for example, family wealth is difficult to observe. Thus the lender relies on other observable characteristics to proxy for hidden family wealth. Clearly, if the lender uses ethnicity as a proxy, the lender is directly violating the law under disparate treatment. More likely, however, the lender may instead proxy with other characteristics, such as education, which may be correlated with ethnicity. To the extent that the correlation of education with ethnicity is only an artifact of the correlation of education with hidden wealth, this is legitimate statistical discrimination.

Additionally, under existing law, a lender would have the opportunity to justify its disparate impact by demonstrating the strength of its correlation with credit risk (e.g., its correlation to structural life-cycle variables). However, if part of the correlation of education with ethnicity is orthogonal to the structural life-cycle variables, then the use of education implies some illegitimate statistical discrimination. As big-data use increases, the problem gets worse. Algorithms could easily be implemented to predict default based on the exact college or high school one attended. Likewise, whereas the number of children may predict expenses, convex dependence on the number of children may predict default that weights heavily on ethnicity, unrelated to direct child expenses, which should be concave.

Under the law, this last use of statistical discrimination—that is, where a scoring variable is largely orthogonal to structural life cycle variables—can create liability for lenders under the legal theory of *disparate impact*. The term conveys the idea that a scoring process disparately impacts an ethnic group, unrelated to structural fundamentals. The challenge posed by FinTech's use of big data is that if a variable predicts default at all, a lender may deem the variable to have credit-risk assessment validity. However, one can imagine that the indicator vector of attendance in all possible high schools or colleges could predict default, but would also pick up ethnic sorting unrelated to structural credit risk.

This paper's identification strategy is the Oaxaca-Blinder decomposition method of the labor literature, with the modern implementation in a treatment frame (Fortin, Lemieux, and Firpo, 2011; Kline, 2011). Our outcome variables are the interest rate borrowers face conditional on acceptance and lenders' accept/reject decisions. The method decomposes ethnicity effects that emerge because of differences in the life-cycle covariates (legitimate statistical discrimination) versus differences that are either disparate treatment (explicit bias) or disparate impact (illegitimate statistical discrimination). Treatment is being in the African-American or Hispanic ethnic groups.

Our empirical challenge is asserting that any residual differences among ethnic groups is not due to lenders' use of legitimate life-cycle variables unobservable to us. Our identification rests on framing our tests within the underwriting and pricing standards that

are imposed on the conventional conforming securitized loan market by the Government Sponsored Enterprises (GSEs). It is well known that, post-crisis, the GSEs (Fannie Mae and Freddie Mac) purchase and securitize more than 90% of the conventional conforming mortgage market in the U.S. It is less recognized that, post-crisis, it is the GSEs who dictate the allowable combinations of contract and eligible borrower characteristics. We exploit these GSE structures to estimate discrimination with respect to rates charged for accepted loans, as well as with respect to loan approval rates. With respect to loan pricing, the GSEs produce a predetermined grid pricing that prices credit risk across loan-to-value and credit score buckets. Any deviation from this grid pricing must reflect lenders' competitive agenda in capturing volume or profit per mortgage. Any ethnic differences in pricing emerging from competitive position, and which are unrelated to creditworthiness, are likely to fail to satisfy the courts as justified in a legitimate business necessity.

Likewise, with respect to loan-application rejection rates, the GSE structure allows us to identify what is likely to be illegitimate discrimination, satisfying omitted-variable concerns. Specifically, once a mortgage clears the credit-quality assessment of the GSEs based on specific observables (e.g., debt-to-income ratios, loan-to-value ratio, credit scores), lenders face only documentation risk (no residual credit risk) that a future defaulting loan will be "put back" to the lender. Because the only variables exposed to this risk are observables, our empirical design for accept/reject estimation satisfies the mean-conditional-independence assumptions of being able to disentangle legal from illegal/illegitimate discrimination. Given this identification strategy, we focus our tests on a large dataset of loan applications and approved loans that are identified as being sourced to a GSE between 2007 and 2012.

In loan pricing, we find that accepted borrowers in Hispanic and African-American ethnic groups pay a 0.08% higher interest rate for purchase mortgages and a 0.03% higher interest rate for refis. Comparing two hypothetical 30-year amortizing mortgages at 3.50% vs. 3.58%, this pricing difference over 30 years represents a present value at the time of mortgage origination of approximately 1% of the loan balance. Annually, this finding translates into almost half a billion dollars per year in extra interest payments.[2] In contrast, FinTech lenders charge only a 0.01% higher interest rate to minority borrowers. These results are largely uniform across lender types and within credit scores.

The differential price result suggests that either lenders try to generate more volume among white majority borrowers by offering rate discounts or that lenders feel that minority borrowers shop around less, either because of personal characteristics or because of financial

[2]This is calculated by multiplying the rate differential (assuming 80% refinance loans and 20% purchase loans) by the total US mortgage market size of $9.09 trillion (`https://www.nytimes.com/2017/05/17/business/dealbook/household-debt-united-states.html`), and then by the 13.20% share reported in Table 1 for the share of African-American and Hispanic borrowers.

services deserts, affecting the ability of lenders to impose monopoly pricing. In work in progress, we study the geographical distribution of differential pricing and rejection rates, to address the possibility that pricing may depend on concepts such as monopoly rents, financial deserts, preying-on-race-based-trust models, and strategic positioning in local markets.

With respect to loan acceptances, we find that African-American and Hispanic applicants are 5% more likely to be rejected for a mortgage than other applicants. These higher rejections are most pronounced in low-credit-score buckets. We interpret this result as being consistent with lender biases against ethnic minorities in terms of assessing the reliability of documents. Consistent with this interpretation, we find that traditional lenders discriminate more, while FinTech firms discriminate less. This finding appears consistent with loan officer facial bias; these loan officers need not be cognitively aware of their biases.

The paper is organized as follows. In Section 2, we discuss anti-discrimination regulations in U.S. mortgage lending, mapping the law to economic methods in statistical discrimination. We present our methodology for the measurement of mortgage discrimination in Section 3, and introduce the role of the GSE pricing and scoring grids in practice. In Section 4, we discuss our multifaceted data. Our empirical results are reported in Section 5, and Section 6 concludes.

2 Legality of Explicit and Statistical Discrimination

Discrimination in residential real estate lending is policed primarily by two federal statutes, the Fair Housing Act of 1968 (FHA) and the Equal Credit Opportunity Act of 1974 (ECOA).[3] The FHA, which is administered by the Department of Housing and Urban Affairs (HUD), prohibits taking any action affecting the terms of the transaction on the basis of a borrower/applicant's race, religion, national origin, gender, familial status (e.g., family size or marital status), or handicap. The ECOA builds on the FHA by expanding the set of protected classes of borrowers. Under these laws, a lender in the mortgage market may not engage in, among other things, (i) refusing to extend credit, (ii) using different standards in determining whether to extend credit, or (iii) varying the terms of credit offers (e.g., loan amount, interest rate, duration and type of loan) on the basis of one of the above-mentioned protected characteristics.

Under the FHA and the ECOA, either a wronged borrower/applicant or the relevant

[3]Although they lack explicit anti-discriminatory prohibitions, the Community Reinvestment Act of 1977 (CRA) and the Home Mortgage Disclosure Act of 1975 (HMDA) also seek to curb discriminatory lending practices by, respectively, mandating a certain level of lending in low- and moderate-income neighborhoods and requiring public disclosure of mortgage data as it relates to ethnicity in order that the public can monitor for discriminatory patterns.

administrative agency has the authority to bring a civil action against a lender. As in other areas of anti-discrimination law, however, the scope of these anti-discrimination mandates has been determined largely by the legal standards that courts have created for a successful claim. Historically, the primary method by which lenders were found to have engaged in prohibited discrimination was by a showing of *disparate treatment* — specifically, a showing that a lender had treated borrowers or applicants differently because of a protected characteristic. For example, in Watson v. Pathway Financial, 702 F. Supp. 186 (N.D. Ill. 1988), an African-American couple successfully sued a lender under the FHA for rejecting their mortgage application because of delinquent credit card accounts. While the lender's justification was a potentially legitimate basis for denying credit, the court found the lender had violated the FHA because the applicants produced evidence that the lender had approved at least six applications from white borrowers showing similar delinquencies.

Disparate treatment claims have also been used to prohibit traditional redlining, in which a mortgage lender refuses to make loans to entire geographic areas because of their racial composition (see Gano, 2017). Thus, for economists, the two sets of variables that are explicitly illegal under disparate treatment are indicator variables of the protected category (e.g., ethnicity in our case) and geography.

In addition to disparate treatment claims, a private party or governmental agency can bring a claim of lending discrimination under the FHA or ECOA under a *disparate impact* theory. In contrast to disparate treatment claims, claims of disparate impact do not involve any allegation of intentional discrimination in how a lender treats applicants/borrowers but rely instead on the fact that lending practices that are facially neutral in their treatment of different groups nonetheless fall more harshly on a protected category of applicants. For instance, in a joint policy statement on the enforcement of the ECOA and the FHA, the Department of Justice, HUD, and all federal banking regulators provided the following example as an illustration of a lending policy that could give rise to a disparate impact claim:[4]

> *Example.* A lender's policy is to deny loan applications for single-family residences for less than $60,000. The policy has been in effect for ten years. This minimum loan amount policy is shown to disproportionately exclude potential minority applicants from consideration because of their income levels or the value of the houses in the areas in which they live.

Despite agency approval of disparate impact theory, the ability of parties to pursue disparate impact claims has been hindered for two principal reasons. First, the existence of a disparate impact, by itself, is insufficient to prove illegal discrimination. Rather, after a

[4]See Policy Statement on Discrimination in Lending, 59 Fed. Reg. 18,266 (Apr. 15, 1994).

plaintiff demonstrates that a lending practice produces a disparate impact, the lender can defend the practice as justified by a legitimate business necessity, provided no alternative policy or practice could achieve the same goal with less discriminatory effect.

Second, neither statute expressly defines the standard for proving a disparate impact violation. In fact, it was not until the Supreme Court's 2015 decision in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project* that the Court even approved the disparate impact framework under the FHA. While the Court interpreted the FHA to permit disparate impact claims, it also required plaintiffs to establish a "robust" causal connection between a specific practice and an alleged statistical disparity. The Court imposed this requirement to ensure that "'[r]acial imbalance ... does not, without more, establish a prima facie case of disparate impact' and thus protects defendants from being held liable for racial disparities they did not create." Since the decision, however, lower courts have struggled with how plaintiffs can establish proof of causality under this standard for a disparate impact case to move forward (Fulks, 2017).

The economic literature has something to contribute to the question of specifying a standard of proof.[5] Starting in the 1970s (see Aigner and Cain, 1977; Arrow, 1973; Phelps, 1972), economic research shifted discrimination discussions to the statistical theory of discrimination rather than taste-based discrimination associated with Gary Becker. The language of statistical discrimination maps well to the legal theory of disparate impact. However, like courts considering disparate impact claims, economists have struggled with positing causality in statistical discrimination estimation. Criticisms of inadequate data and omitted-variable biases in estimation have plagued this literature.[6] For example, Shafer and Ladd (1981) find some evidence of interracial pricing differences, and Black and Schweitzer (1977) find indications of differences in loan terms. Yet Sandler and Biran (1995) critique any use of statistics in legal proceedings of discrimination claims due to the prevalence of omitted-variable bias and poor identification strategies in these studies. Not surprisingly, the mortgage literature on discrimination parallels these broader patterns, primarily focusing either on the methodological difficulties in providing robust statistical evidence for discrimination (see Black, Schweitzer, and Mandell, 1978; Kaye, 1982; Maddala and Trost, 1982; Rachlis and Yezer, 1993) or on the efficacy of specific types of legislation.[7]

[5]See commentary by Glassman and Verna (2016) and Winfield v. City of N.Y., No. 15CV5236-LTS-DCF, 2016 WL 6208564, at *1 (S.D.N.Y. Oct. 24, 2016); Cty. of Cook v. HSBC N. Am. Holdings Inc., 136 F. Supp.3d 952, 955 (N.D. Ill. 2015).

[6]Cited data omissions in empirical tests of discrimination include omitted information such as the loan to value ratio at origination, the debt to income ratio, all the contracting elements of the mortgage, the property characteristics and exact address, the applicant's ethnicity, gender, credit history and debt burden levels, lender and regulatory characteristics, all at the loan level for successful and unsuccessful applicants.

[7]For example, the Fair Housing Act of 1968, the Equal Credit Opportunity Act of 1974, the Home

However, we think that more can be accomplished. Consistent with the Supreme Court's decision in *Inclusive Communities*, the theory behind statistical discrimination provides guidance as to how one can demonstrate a "robust" causal connection between lending practices and racial or ethnic disparities in lending outcomes. In statistical discrimination, agents (in our case lenders seeking to screen applicants) have limited information and no animus against racial groups. Statistical discrimination arises as a solution to a signal extraction problem. Agents seek to reconstruct hidden information as to the expected creditworthiness of applicants using observable proxies.

We have already established that the use of a protected variable (ethnicity) is illegal under disparate treatment. However, could other variables be legal proxies for hidden information, even if they are correlated with ethnicity? Under the legal theory of disparate impact, statistical discrimination is allowable for "legitimate business necessity." Economic theory guides us that the meaning of this phrase is that a variable legitimately appears as a structural variable that maps the ability of a household to repay a loan to their economic fundamentals. In particular, one can write down a life-cycle model in which cash flow for repayments emerge from the current borrowing position (debt), cost of borrowing (credit score), income (in levels, growth, and risk), wealth, and regular expense levels (cost of living measures).

Legally, limiting the legitimate business necessity defense to structural variables that predict creditworthiness is entirely consistent with legal precedent assessing whether the refusal to extend credit violates federal antidiscrimination law.[8] Further, cases involving "reverse redlining" — the extension of credit to a protected class on inferior terms than those offered to non-protected applicants — also suggest that only structural variables that map to repayment risk should be permitted under the "legitimate business necessity" defense in assessing disparate impact claims in rates, conditional on the extension of credit.[9] Thus, whether in loan rejection or in loan pricing, the use of any of the foregoing structural variables should be considered "legitimate business necessity" under disparate impact theory, even if

Mortgage Disclosure Act of 1975, the Community Reinvestment Act of 1977, and the HMDA amendments to the Financial Institutions Reform Reregulation and Enforcement Act of 1989.

[8]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.'").

[9]See, e.g., 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.")

this variable statistically loads on (punishes) a particular protected category. As a matter of economic theory and legal doctrine, this is legitimate statistical discrimination.

What if a lender cannot see one of these variables (wealth, say), but can see a variable (e.g., the name of the high school attended) that correlates with wealth? Under the theory of disparate impact, this variable should be allowable if it is only disparately impacting the pool of applicants in sorting on wealth. In other words, conditional on latent hidden wealth, high school is orthogonal to loan decisions. A slightly less stringent assumption would also be consistent: conditional on latent hidden wealth, the impact of high school on loan decisions must be the same for one ethnic group as the other (ignorability).

3 Methodology

3.1 Overview

Modern methods to estimate discrimination emerge from the works of Oaxaca (1973) and Blinder (1973). Oaxaca and Blinder established a decomposition method whereby ethnic (or gender) differences in an outcome variable are decomposed into that which can be explained by differences in structural covariates and that which remains unexplained. The first piece is selection; the latter is discrimination. In our context, the outcomes of interest are, iteratively, the loan pricing conditional on acceptance and the accept/reject decision. Structural covariates, (i.e., legitimate sources of variation for lenders to use in scoring) are life-cycle variables that are truly structural sources of repayment risk assessment.

Decomposing selection from discrimination is not as straightforward as simply applying the Oaxaca-Blinder decomposition because, as we discuss in this section, conditional independence is hard to assert. We exploit the institutional framework by which GSEs underwrite loans and price risk to allow us to satisfy the independence conditions.

3.2 Oaxaca-Blinder in an Ethnicity Treatment Frame

Recent innovations cast the Oaxaca-Blinder methodology in the treatment-effect language of the program-evaluation literature (Fortin et al., 2011; Kline, 2011; Słoczyński, 2015). We follow in this tradition. The treatment group in our study is African-American and Hispanic mortgage applicants. Our control group includes all other ethnicities.

Our objective is to estimate the average treatment effect on the treated (ATT); i.e., the difference between the realized loan outcomes for ethnic minority mortgage applicants and the counterfactual outcomes for the same applicants if they were credit scored in the pool of the control ethnicities (predominantly Caucasians). We refer to this counterfactual using

the terminology "under blinded credit scoring" because our control group represents a large majority of mortgage applications. Our first assumption is just that — that we do not suffer from the lack of a base group, a criticism of the Oaxaca-Blinder methods (Oaxaca and Ransom, 1999); namely:

> *Assumption 1 (Simple Counterfactual for Treatment): Credit scoring of an ethnic minority household would be the same as scoring of the control group scoring (blinded scoring) if households were not identified to be in the ethnic treatment group.*

Next we set up the decomposition. Let subscript T stand for the treated (African-American and Hispanics) and B for the control (B denoting blind). Each household i has a potential loan outcome associated with being credit scored as if an ethnic minority Y_{Ti} and being credit scored Y_{Bi} in a blind model. Only one of Y_{Ti} or Y_{Bi} is ever observed for any household i. We are interested in the difference in the average values

$$\Delta = \overline{Y}_T - \overline{Y}_B$$

Next, we make the standard linearity assumption in this literature:

> *Assumption 2 (Linearity in Structural Variables): Outcomes Y_{Ti} and Y_{Bi} are linearly related to the structural variables, denoted X where X is the vector $X_i = [X_{i1}, \ldots, X_{iK}]$:*[10]

$$Y_{Ti} = \beta_{T0} + \sum_{k=1}^{K} X_{iTk}\beta_{Tk} + \varepsilon_{Ti}$$
$$Y_{Bi} = \beta_{B0} + \sum_{k=1}^{K} X_{iBk}\beta_{Bk} + \varepsilon_{Bi}.$$

The difference Δ in outcomes can be decomposed as follows:

$$\Delta = \underbrace{\sum_{k=1}^{K} \left(\overline{X}_{Tk} - \overline{X}_{Bk}\right) \widehat{\beta}_{Bk}}_{explained} + \underbrace{\left(\widehat{\beta}_{T0} - \widehat{\beta}_{B0}\right) + \sum_{k=1}^{K} \overline{X}_{Tk} \left(\widehat{\beta}_{Tk} - \widehat{\beta}_{Bk}\right)}_{unexplained}.$$

Loan outcome differences result from (i) (explained) the difference in the household structural variables, applying the blinded credit scoring, (ii) the difference in the constant shifters

[10] Lenders construct credit scoring models from discriminant analysis of defaults. We anecdotally assume that these models linearly analyze structurally implied variables or step functions of these variables. The scoring models emerge from such analyses, allowing us to approximate with linear terms. We do nonparametric versions of our estimations to show robustness to this linearity assumption.

in outcomes, and (iii) the difference in outcomes due to the treated being scored with different scoring parameters, relating the structural variables to loan outcomes, compared to the blinded scoring parameters. The sum of the two unexplained terms is the total discrimination.

We need two additional assumptions to posit that the decomposition can causally identify discrimination. These assumptions are at the heart of the role GSE grids play in our methodology.

> *Assumption 3 (Overlapping Support): Each possible treated realization of* $X_i = x$ *and* $\varepsilon_i = e$ *must be in the common support; i.e.,* $0 < \Pr[i \subseteq T \mid X_i = x,\ \varepsilon_i = e] < 1$.

> *Assumption 4a (Conditional Independence):* Applicants' unobserved life-cycle characteristics are independent, conditional on observed covariates. $E(\varepsilon \mid X) = 0$

Kline (2011) establishes that a lighter version of Assumption 4a is all that is necessary in this setting:

> *Assumption 4b (Ignorability): Any selection based on unobservables must be the same for the treated and control. Unobservables do not need to be independent of X, as long as their distribution conditional on X is the same for both ethnic groups. Denoting the distribution function by* $g(\cdot)$, $g(\varepsilon \mid X, i \subseteq T) = g(\varepsilon \mid X, i \subseteq B)$.

Neither 4a nor 4b is easy to establish in the setting of applicants selecting to apply for a mortgage. The most worrying concern is the implications from omitted life-cycle variables. Imagine that expected income growth causally relates to repayment risk. However, expected income growth is unobservable. Furthermore, it is correlated with current income, but the intensity of this correlation is different for ethnic minorities as compared to the ethnic majority. All of that seems very plausible. In this case, the omitted-variable bias in the scoring model parameters would lead to a biased Oaxaca-Blinder decomposition. Moreover, the weaker ignorability assumption does not help because the βs would not be biased in the same magnitude.[11] Fortunately, these concerns are greatly attenuated in the context of GSE lending.

[11] Note that this omitted-variable bias may be even larger for studies which proceed with more descriptive models of discrimination.

3.3 Discrimination in Loan Pricing & Error Independence

The GSEs play a very significant role in the pricing of mortgages. Each GSE guarantees the timely payment of principal and interest on its MBS and charges a fee for providing that guarantee. These *guarantee fees* (or G-fees) cover projected credit losses from borrower defaults over the life of the loans, administrative costs, and a return on capital to the GSE. 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).[12]

In March 2008, the GSEs introduced up-front fees based on an 8×9 matrix of LTV and credit score buckets, called the *Loan Level Price Adjustments* (LLPAs). The LLPA schedules are published by the GSEs and reviewed by the Federal Housing Finance Agency on an annual basis (see Federal Housing Finance Administration, 2009, 2010, 2011, 2012, 2013; Fuster and Willen, 2010). Figure 1 depicts a typical LLPA grid. These fees have been adjusted four times between 2008 and 2012 in response to changes in the GSEs' forecasts of house price dynamics and credit losses. 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.

Lenders know the grid at any point in time and thus price consumer-facing interest rates that are sufficient to cover these fees as well as any profits (Fuster et al., 2013). Figure 2 represents a schematic for how a lender would think about the pricing of a loan contract that is being originated with the intent to securitize through a GSE. The figure represents two mortgage applications with the same credit score and the same loan-to-value (LTV) ratio, both of which have been approved by the underwriting system. The mortgage interest rate that a household sees on a mortgage consists of three parts. First, all mortgages within the same LTV/credit score bucket face the same market price, determined by the Base Mortgage Rate, which reflects the primary market interest rate for loans to be securitized by the GSEs. Second, all mortgages within the same credit score/LTV bucket face the same LLPA fee added onto the market price.

The third part of the pricing is discretionary to the originator and is determined by the lender's ability to originate loans at interest rates in excess of the minimum required to cover the latter two components. Lenders determine offered mortgage rates using internally produced *rate sheets*, which are schedules of possible rates that a loan officer (or algorithm programmer) can charge for individuals within a cell of the LLPA grid. They use discretionary pricing in rate sheets to incorporate strategic pricing for dynamic indus-

[12]The actuarially fair pricing of the G-fees 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).

trial organization considerations; i.e., monopoly rents and volume positioning in markets. In addition, higher rates may be simply used to discriminate against either certain types of borrowers who are known to shop around less or certain borrower areas where finance deserts or collusion make sustained higher prices possible.

Imagine some of these differential pricings to the consumer are correlated with ethnicity (they may all be). Again, the legal standard for a disparate impact claim is that a court must ultimately conclude (i) that discrimination occurred and (ii) that no legitimate business necessity mandated the use of a sorting mechanism that discriminated. We translate these requirements into economics in two parts. First, we are able to identify statistical discrimination in interest rates charged to borrowers via the decomposition method. Second, there is no omitted unobservable variable that could be part of a structural model and represents a legitimate business necessity. Rather, any disparities in rates among borrowers reflect profits arising from discretionary pricing rather than from credit risk, because credit risk (which is borne entirely by the GSEs) has been priced through the GSEs' LLPAs. Yet profit margins are unlikely to constitute a legitimate business necessity, as reflected in the reverse-redlining cases noted previously. Thus, all statistical ethnic discrimination that we are able to isolate within the grid represents potentially illegitimate statistical discrimination or explicit bias.

3.4 Discrimination in Loan Rejections & Error Independence

In addition to standardizing loan pricing, the GSEs also play a critical role in determining which mortgage loan applications should be accepted. As is well known, a central activity of the GSEs is to purchase mortgages from originators, packaging mortgages into MBS, and guaranteeing loan payments to MBS investors. We argue that the guarantee of loan payment provides us with exogeneity to unobservables that allows for identification of discrimination in loan acceptance rates. While an originator who retains a mortgage in its portfolio is incentivized to estimate an applicant's credit risk, an originator who is engaged in sourcing and selling loans to the GSEs ordinarily bears no credit risk, except for *put-back risk* discussed below. Moreover, the criteria used by GSEs for determining whether a loan is acceptable for purchase and pooling are based entirely on a specific set of observable characteristics.[13]

[13]The 2010 Sellers' Guide for Fannie Mae states the following condition for the purchase of loans:

> **Loan Qualification**. Lender must ensure that all loans selected for delivery meet Fannie Mae's underwriting and eligibility guidelines and legal requirements and match the terms of the commitment or for MBS loans, the pool purchase contract, including mortgage type, amortization, original term, and pass through rate(s) selected for delivery when the commitment or contract was created.

Within a set loan type (e.g., 30-year, fixed-rate contracts for an owner-occupied, single-family home), Fannie Mae's Sellers' Guide specifies these characteristics to include: Mortgage Eligibility (maximum LTV, allowable loan amortization), Borrower Eligibility (documentation, LTV maximums, credit score minimums), and Occupancy/Property Eligibility (documentation of status).

In addition, the applicant's characteristics are processed for acceptability in an automated underwriter system (*Desktop Underwriter* for Fannie Mae; *Loan Prospector* for Freddie Mac). The underwriter system considers only the following: an applicant's credit report, liquid reserves, total expense ratio, and co-borrowers. For lenders not using the underwriter system, the Sellers' Guide also specifies that manual underwriting should focus on these same observables. Thus, for our purposes, acceptability of an application is based on observables, implying that there would be no point in a lender engaging in statistical discrimination for default risk.

Our setting involves one further complication. The lender remains subject to the residual risk that it may have to repurchase a defaulting loan under its representations and warranty liability (*put-back risk).* The importance of this representation was underscored following the Financial Crisis in 2010, when GSEs put back $4.2 billion of loans (*American Banker*, July 14, 2016).[14] However, the formulaic character of GSE underwriting ensures that both the original accept/reject decision and any subsequent liability for a breach of a representation/warranty will hinge on a borrower/applicant's *observables.* Put-back risk only concerns risks in the documentation and verification of observables (income, debt-to-income, credit score, etc.). Thus, as between two applicants having similar observables, the GSE lending framework suggests that disparity in loan outcomes must relate to the perceived reliability of an applicant's documentation. Yet, conditional on observables, the impact of an applicant's perceived reliability on loan decisions should be the same for one ethnic group as for another. Using race to proxy for an applicant's reliability would therefore have no economic relationship with the repayment risk of an applicant, thus causing illegitimate statistical discrimination.

[14] A large number of put-backs by a lender will induce the GSEs and the MBS market to add a discount to the price for products from this lender. These overall lender adjustments can be controlled for in our Oaxaca-Blinder methods with firm fixed effects, as in Abowd, Kramarz, and Margolis (1999) and Card, Cardoso, and Kline (2016).

4 Data

4.1 Data Sets

A key obstacle in prior empirical mortgage discrimination studies has been their reliance on the Home Mortgage Disclosure Act (HMDA) data,[15] the only data source with loan-level information on applicant ethnicity for both successful and unsuccessful loan applications. HMDA also contains applicant income and (nonstandardized) information on lender name. 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 (credit score, debt-to-income ratio, etc.). A challenge with mortgage loan data in the U.S. has been the lack of a linking of HMDA data with other datasets that do contain loan contracting elements, lender and borrower characteristics, and property characteristics.

Thus, we embarked on a multi-year project of linking loan-level data across four separate data sets using machine learning techniques. The data sets in this large-scale statistical merge are:

- HMDA (30.6 million loans and 11.6 million loan rejections between 2008 and 2012). HMDA data include information on borrower income and ethnicity and lender identifiers for loan-level mortgage origination data. Geography is available only at the level of the census tract.
- Dataquick (30 million loans between 2008 through 2012). Dataquick data provide transaction and assessor information including all mortgage lien recording data, loan performance data (i.e., prepayment and default), lender names and location, but very little information on actual mortgage contract terms other than the loan amount.
- McDash (28.7 million loans between 2008 and 2012). McDash data provide loan-level data compiled by Black Night Financial Services and includes quite comprehensive information on the mortgage contract elements and loan types.
- Equifax (20 million loans between 2008 and 2012). Equifax data provide (for a subset of loans in the McDash data) information on other consumer financing balances that are held by borrowers in addition to their mortgages (this merge was done by Equifax).

We describe our merging algorithm in detail in Appendix A. To standardize our analysis,

[15]The HMDA surveys account for approximately 90% of mortgage origination in the U.S. (see Engel and McCoy, 2011). HMDA reporting is not required for institutions with assets (when combined with the assets of any parent corporation) that are below $10 million on the preceding December 31, or institutions that originate 100 or more home purchase loans (including refinancings of home purchase loans) in the preceding calendar year (see `http://www.ffiec.gov/hmda/pdf/2010guide.pdf`).

we filter the data to focus on 30 year fixed rate, single-family residential loans, securitized by the GSEs, over the period 2008 through 2012. We also filter out CRA zip codes. Finally, because our use of median proxies for LTV and credit score, we trim the support to be LTV between 0.3 and 1 and credit scores to be 630 to 770.

For accepted loans, the merge brings a host of property and applicant characteristics from Dataquick, McDash, and Equifax to the baseline HMDA data, which has ethnicity. However, none of these auxiliary datasets covers loan application rejections. Thus, as explained in more detail in Appendix A, we augment the rejection data with proxies for two key scoring variables that are not included in HMDA. We proxy for the applicant's credit score using the median credit score of the census tract location for all Government National Mortgage Association (GNMA)-securitized loans that were originated each year, as reported in the McDash data set. A census tract is on average 1600 households (4000 inhabitants), designed by the Census Bureau to reflect relatively uniform economic situations. Likewise, although we have the actual applied-for loan amount, we need to proxy for property value to construct a proxy for the loan-to-value ratio. We start with annual snapshots of the property-specific market assessed value and the date of the assessment that are reported in Dataquick for nearly all residential properties in the U.S. Within each census tract, we take the simple average of these market-assessed property values each year. We then use the actual loan application amount as the numerator to construct a proxy for the loan-to-value ratio for a specific property located in that census tract.

4.2 Summary Statistics

In the summary statistics reported in Table 1, FinTech lenders are shown to originate approximately 2.5% of both accepted and rejected loans of GSE 30-year, fixed-rate mortgages. These data were obtained by matching HMDA lender names to the firms that were identified as FinTech in Buchak, Matvos, Piskorski, and Seru (2017). In contrast, 46.6% of the accepted loans were originated by the top 25 lenders by origination volume in their respective loan origination year and 63.8% of the rejected loans were originated by top 25 mortgage originators. The origination volume rankings were determined by matching HMDA lender names with mortgage origination statistics obtained from *Inside Mortgage Finance*. HMDA reports the ethnicity of the borrower, or borrower application. The table shows that 15.9% of the accepted loans were to African-American or Hispanic borrowers whereas 14.2% of the rejected loans were to African-American or Hispanic applicants.[16] Thus, in raw statistics,

[16]HMDA reports three ethnicity categories: White/Asian, African-American/Hispanic or unknown. For those loans in the unknown category, we applied a classifier for ethnic first and last names, as applied in Kerr (2008); Kerr and Lincoln (2010).

our treatment ethnicities have lower rejection rates.

When we run the interest rate estimations, we also want to ensure robustness to debt levels using the Equifax match. Panel C of Table 1 reports the summary statistics for the subsample of 1,121,395 (71% of the total sample of accepted loans). As shown, these loans are quite similar to the loans not matched to Equifax, except that there is a slightly higher share of refinance loans than purchase loans.

4.3 Effect of Grid-Leveling on Outcomes

Before turning to the analysis, we present a series of figures showing the effect of leveling applicants/borrowers by analyzing within the LLPA grid. In Figure 3, panels (a), (b), and (c) display histograms of borrower interest rates by treatment (treat=1 is the set of Hispanics and African-Americans). Panel (a) shows that in the raw data, both the treated and the control borrowers seem to have bimodal distributions, with a lot of realized interest rates at 5% and then between 5.75% to 6.75%. However, the ethnic treatment group has more mass in the higher of these modes, resulting in higher means. Therefore, in contrast to the summary statistics that Hispanics and African-Americans have lower rejection rates, the histograms reveal that this treatment group fares worse on interest rates in the raw data.

However, when we level the interest rates within the grid, by subtracting out the year-grid cell mean, Panels (b) and (c) show that the differences in rates become much smaller. The ethnic treatment histograms remain, however, shifted to the right (higher rates). Reassuringly, for our robustness checks, Panels (b) and (c) are very similar. Figures 6 and 7 in Appendix B repeat this exercise, showing differences in the same set of histograms by raw rates versus de-meaned grid rates for splits of FinTech versus traditional lenders and then the Top 25 lenders versus smaller brokers.

Figure 4 presents conventional loan data from the Home Mortgage Disclosure Act surveys for loan applications that were rejected. The figure compares loan application rejection probabilities for the raw data (on the left-hand side of the graph) and demeaned by year-LLPA buckets (on the right-hand side of the graph) for borrowers in the control group (ethnicity $\neq$ African-American/Hispanic) and for the borrowers in the treatment group (ethnicity = African-American/Hispanic). In contrast to the rates histograms, the LLPA pricing schedule is helpful in understanding the rejection rates of loans, since the demeaning leads rejection probabilities for treated and untreated loans to be more similar than they appear in the raw data. Similarly, Figure 8 panels (a) and (b) in Appendix C plot the rejection rates over time for FinTech or Traditional (panel (a)) and Top 25 volume lenders or smaller lenders (panel (b)). FinTech lenders have similar rejection rates as traditional lenders in the grid; however,

smaller-volume lenders reject far fewer applicants than Top 25 lenders.

In sum, the histogram depictions of the data reveal that within-grid pricing seems relatively uniform across types of lenders and borrowers. However, rejection rates vary substantially, even in the within grid framing. This is perhaps not surprising, as the intent of the grid is to level pricing. Before making inferences, however, we need to implement the full Oaxaca-Blinder decomposition in a multivariate setting.

5 Results

5.1 Mortgage rates

Table 5 shows the results of Oaxaca-Blinder regressions of coupon rates on newly issued 30-year fixed-rate mortgages against various right-hand-side variables. Table 5, panel A, includes the full sample of loans, while panel B reports results for the (slightly smaller) sample of loans merged with Equifax data, allowing us to include additional consumer debt variables (the log of the sum of outstanding monthly balances at origination of student loans, car loans, other consumer debt, and HELOCs) on the right-hand side. Panel C adds lender fixed effects to the specification using the Equifax merged data. In panel A, Column (1) shows OLS regressions of the purchase mortgage coupon rate on loan-to-value ratios, credit scores, and log annual income, whereas Column (2) includes month and bucket fixed effects. Columns (3)–(5) show results of Oaxaca (1973) regression of purchase mortgage coupon rate on the same explanatory variables. Panels B and C report the same set of specifications with the inclusion of the log of additional borrower debt.

As shown in Column (3) of panels A, B, and C in Table 5, the overall difference in the mean mortgage interest rate between African-American and Hispanic borrowers ranges between .17% and .19%. As shown in Panels A and B, Column (5), half of this total difference is due to discrimination. Specifically, the unexplained differences in rates is about .091% in Panel A without controls for other debt and about .083% in Panel B using the smaller Equifax other debt controls. Adding lender fixed effects, as reported in Panel C, reduces the differential effect of discrimination on the mortgage interest rates of African-American and Hispanic borrowers to 7.7 basis points. The interpretation of these results is that conditional on being given a loan, African-American and Hispanic borrowers pay an average of 0.08% more than other borrowers for their purchase mortgage.

Despite the seemingly small amount, in the context of residential mortgage loans this pricing differential is economically meaningful for both borrowers and originators. For borrowers, imposing an additional 0.08% per year in interest for a 30-year fixed rate amortizing

loan has a present value of approximately 1% of the loan balance assuming a a benchmark annual interest rate of 3.5%. Likewise for originators, sourcing mortgages with an 8 basis premium on the coupon would represent a material increase in the expected servicing income for the loan. Given that the majority of interest payments on a mortgage go to compensate MBS investors as well as the need to cover the G-fee, the typical loan servicing fee for conventional mortgages is just 25 basis points—the minimum required by Fannie Mae and Freddie Mac (Aldrich, Greenberg, and Payner, 2001). Thus, our finding of an 8 basis point price differential suggest pricing discrimination may constitute a significant fraction of loan servicing income available for a mortgage, as well as a possible source for "excess" servicing income (i.e., servicing income in excess of 25 basis points) that can be monetized through an excess servicing IO trust.

In Column (5) of panels A, B, and C we reported the differential weighting of the life-cycle covariates in setting the mortgage contract interest rate. As shown, there appears to be a statistically significant underweighting of the log of the borrower's income when lenders consider African-American and Hispanic borrower applications. Again, this differential underweighting of observable life-cycle variables also leads to disparate impact in the form of higher rejection rates.

Table 6 shows results for traditional vs. FinTech lenders. Panel A reports the results for an OLS regression in Column (1) and Columns (3)–(5) present the results of Oaxaca (1973) regressions of mortgage rate for purchase mortgages against life-cycle characteristics including month and bucket fixed effects for traditional lenders. Columns (6)–(8) present the same OLS and Oaxaca (1973) regressions for the FinTech lenders. The OLS specifications, reported in Column (1), indicate that .89% of the mortgage interest rate of the traditional mortgage lenders is not associated with life-cycle factors. The Oaxaca decompositions, however, indicate, as shown in Column (2), that the traditional lenders have nearly a .19% higher mortgage interest rate for African-American and Hispanic borrowers. Of that differential interest rate, 9.8 basis points are justified by differences in the levels of the life-cycle characteristics of African-American and Hispanic borrowers, whereas 9.1 basis points are due to discrimination that is not justified by observable life-cycle factors. As reported in Panel B, the inclusion of lender fixed effects, leaves the overall differential mortgage interest rate unchanged at .189%, which is essentially the same as the value reported for the full sample in Table 5, whereas the discrimination factor is now reduced to about 7 basis points, again similar to the full sample results. We again find, as reported in Column (5) that lenders appear to underweight income in their interest rate decisions; however, these weight differentials are quite small.

The OLS results reported in Table 6, Column 5, indicate about a 7 basis point differential

mortgage interest rate for purchase mortgages that is not determined by life-cycle observables. The results of Oaxaca (1973) regressions for the FinTech lender are reported in Panels A and B, Columns (6)–(8). As shown, the FinTech lenders have a smaller overall differential mortgage interest rate for African-American and Hispanic borrowers of .14%. However, the fraction of that differential that is unexplained by the differences in African-American and Hispanic observables is larger at .08% without lender fixed effects and .07% when lender fixed effects are included. FinTech lenders again underweight the credit score and the log borrower annual income in there mortgage interest rate decisions, although the differential magnitudes are very small.

Table 7 shows the corresponding results for large top 25 versus smaller non FinTech lenders. The OLS specifications for both the top 25 and the smaller non FinTech lenders, reported in Columns (1) and (5), indicate a mean increment in the mortgage interest rate that is not a function of life-cycle factors of 8–9 basis points for both lender types. From the Oaxaca regressions, reported in Columns (2)–(4) and Columns (6)–(8), the overall difference in the mean mortgage interest rate between African-American and Hispanic and Caucasian and Asian borrowers is about .2%. The unexplained component of that differential is about .09% for both lender types. The top 25 lenders appear to systematically overweight the loan-to-value ratio and the credit score and underweight the log of the borrowers income. Again, as previously discussed, these differential weightings of life-cyle characteristics are also discriminatory, since the weights should be equal. The addition of the lender fixed effects to these specifications, slightly lessens the discriminatory component of the mortgage interest rate differential for African-American and Hispanic borrowers to about 8 basis points for the large lenders and about 7 basis points for the smaller non FinTech lenders. The same pattern of statistically significant differential weightings on the life-cycle factors for African-American and Hispanic borrowers still evident in these specifications.

In Tables 8–10, we re-run the purchase mortgage regressions that are reported in Tables 5–7 using a sample of refinanced mortgages. Tables 8 reports the main results for the refi mortgage sample for both the OLS and Oaxaca specifications. As shown in all three panels for Column (3), the overall difference in the mortgage interest rates for African-American and Hispanic borrowers is about 19 basis points. As reported in Column (4), the explained component of this differential is about 16 basis points. The unexplained discriminatory component of the mortgage interest rate differential for African-American and Hispanic borrowers, reported in Column (5), is about 3 basis points which is less than half the discriminatory differential found for purchase mortgage interest rates. This difference could be due to differences in competitive factors in the refi mortgage market. It could also reflect the fact that a purchase mortgage is a tied purchase of a mortgage and a house and it is likely

that the lenders/brokers for this market segment are more specialized thus making lender search more difficult for borrowers. We again see the same pattern of statistically significant overweighting of the loan-to-value ratio and the credit score and the underweighting of the log income of African-American and Hispanic borrowers. These weighting differentials on observable characteristics for these borrowers are also discriminatory.

Table 9 reports the results of the OLS and the Oaxaca decomposition regressions for refi mortgage rates on legitimate life-cycle characteristics for traditional and FinTech lenders. We again find smaller differential effects in the mortgage interest rate for the refi mortgage sample than those found for the purchase mortgage sample reported in Table 6. The overall differential in the rates is about 19 basis points for the traditional lenders, whereas the unexplained illegitimate discriminatory differential in interest rates for African-American and Hispanic borrowers is about 4.5 basis points. In contrast, the FinTech lenders have a differential mortgage interest rate of only 7 basis points overall, and the unexplained discriminatory component of that is only 1.3 basis points or nearly three times smaller than that of the traditional lenders. Interestingly, as shown in Columns (2) and (6) of Table 9, FinTech lenders do not appear to use differential weights on life-cycle factors, whereas the traditional lenders overweight the loan-to-value ratio and credit scores and underweight the log of borrower income. Both the unexplained mortgage interest assigned to African-American and Hispanic borrowers and the differential weighting of the observable life-cycle factors of these borrowers should be interpreted as discriminatory on the part of traditional lenders leading to disparate impacts on these borrowers.

We again consider the differences in top 25 and smaller non FinTech lenders in Table 10. As shown in both Panel A and B, the differences between the two lender types is not large. The overall differential in the mortgage interest rates for mortgages originated by the top 25 lenders in terms of origination volume is about 16 basis points once we include lender fixed effects. The unexplained discriminatory component of the differential is about 3 basis points similar to the overall results reported in Tables 8. The smaller non FinTech lenders have a larger overall differential interest rate for African-American and Hispanic borrowers of 26 basis points. However, the unexplained decomposition is only about 3.3 basis points whereas 21 basis points is attributable to differences in observables for these borrowers. We again find that both large and small lenders tend to overweight the loan-to-value and the credit score and underweight the log of the borrower income. These differences would also lead to disparate impact in the form of reduced access to the refi mortgage market for African-American and Hispanic borrowers.

Overall, our results for interest rate discrimination indicate that the discipline of the LLPA's leaves little scope for interest rate discrimination among lenders, but discrimination

nevertheless exists. On average, as shown, ethnic minority borrowers pay about 0.08% higher mortgage interest rates for purchase mortgages and these differences cannot be justified by observables. The discriminatory differential mortgage interest rates for refi mortgages originated to African-American and Hispanic borrowers are 0.03% higher which is less than half as large as the differential for purchase mortgages. The reasons for these differences are likely to be associated with local competitive factors and perhaps to frictions associated with the tied purchase structure of purchase mortgages involving negotiations both for a house and a loan. In future work we intend to explore these channels in more depth.

Another important result from our Oaxaca decomposition regressions is that the unexplained differential interest rate charged to African-American and Hispanic borrowers was about 7 basis points for purchase mortgages originated by FinTech lenders and this differential is similar to that of traditional lenders for purchase mortgages. However, in contrast, the FinTech lenders discriminate significantly *less* for refi mortgages given our finding that the differential interest rate charged for African-American and Hispanic borrowers was only about 1.3 basis points compared to about 3 basis points of discriminatory differential that was charged by traditional refi lenders. In addition to this difference, FinTech refi lenders were also found to *equally* weight the legitimate observable factors of minority and majority borrowers. We posit that the differential mortgage interest rates charges on the part of traditional lenders as well as their differential weighting of observable characteristics could be legally problematic, since they are unrelated to the life-cycle characteristics of African-American and Hispanic borrowers.

5.2 Rejection Rates

As discussed above, Figure 4 reveals no immediate sign of discriminatory rejection rates, since there is no consistent relationship between either raw or de-meaned rejection rates for the control and treated groups over time. Rejection rates are higher for the control group in some years, and for the treatment group in others. However, because borrower or lender characteristics might differ significantly between the treated and untreated groups, which would make the unconditional comparisons misleading, we now look at rejection rates in more detail.

Focusing first on the full sample of loans for which we have complete data, Table 2 compares regressions of lenders' accept and reject origination decisions on borrowers' loan-to-value ratios, their credit scores, and the log of their annual incomes. In Column (1), we show the ordinary least squares (OLS) regression results with lender fixed effects and in Column (2) we show the same OLS regression with the inclusion of month, bucket and lender fixed

effects. The results for both of the OLS specifications indicate that the effect of the intercept is positive. These results suggest that, in addition to legal life-cycle screening components associated with the rejection rates of borrowers, on average non-life-cycle screening elements, possibly associated with protected characteristics, increase the likelihood of a borrower's loan application being rejected.

Similar to the OLS results, Columns (3)–(5) show results of the Oaxaca (1973) decomposition. As shown, the Oaxaca results suggest that overall African-American and Hispanic borrowers are 15% more likely to be rejected for a mortgage. Of this total difference in the mean rejection rate, from row 1 of Column (4) 10.3% the difference in rejection rates is due to average differences in legitimate observable characteristics of African-American and Hispanic borrowers. However, from row 1 Column (5), 4.91% of the differential rejection rate is *not* explainable by observables and is thus attributable to illegitamate discrimination. As discussed in Section 3 above, the formulaic character of GSE underwriting means that two applicants with similar observable structural variables, and who would be assigned to the same LLPA grid, should face the same likelihood of loan approval. This non-zero treatment effect is clear evidence of discrimination against African-American and Hispanic borrowers. In addition, as shown in Column (5), there appear to be statistically significant differences in the weights that lenders attach to the evaluation of legitimate life-cycle covariates. These results suggest that lenders overweight the importance of the credit scores of African-American and Hispanic borrowers. This differential overweight on credit score represents another component of illegitimate screening, or discrimination, since the legitimate use of life-cycle characteristics should receive equal weighting on average between the two borrower groups.

Tables 3 and 4 show corresponding results for traditional versus FinTech lenders and for top-25 versus smaller non FinTech lenders, respectively. It can be seen in Column (3) of Table 3 that the illegitimate discrimination rate for traditional lenders is about the same as for the overall sample (5.08%), whereas FinTech lenders are nearly 1% (92 basis points) *less* likely to reject an African-American or Hispanic loan applicant for reasons unrelated to observable life-cycle covariates. Here again, as shown in Column (3), traditional lenders appear to overweight the credit scores of African-American and Hispanic borrowers in their rejection decisions. Interestingly, as shown in Column (6), FinTech lenders also positively overweight the borrowers' loan-to-value ratios as well as their credit scores in their rejection decisions. However, given the 1% greater likelihood that FinTech lenders will not reject a loan application for African-American and Hispanic borrowers, the observed overweighting of these life-cycle factors appears not to lead to increased rejection levels for minority borrowers, in fact, the overweighting appears to benefit these borrowers.

As reported in Table 4 in Column (1) and (4), the overall difference of mean in the rejection rate of African-American and Hispanic borrowers is 14.3% and 16.1% respectively for the larger (top 25) and the smaller (non-top 25 lenders) non-FinTech lenders respectively. The unexplained illegitimate discrimination, not based on observable life-cycle factors, is 5.67% for larger lenders and 3.85% for smaller lenders. The results reported in Column (3) indicate that the top 25 lenders in terms of origination volume appear to equally weight the importance of the loan-to-value ratio for African-American and Hispanic borrowers. However, there is a statistically significant overweighting of credit scores and an underweighting of the log of borrower income. In contrast, smaller lenders have a statistically significant overweighting on the importance of the credit scores of African-American and Hispanic borrowers and an overweight of the log of the borrower's income in their decisions to reject loan applications. The results on the presence of large differentials in the weights that both large and small lenders attach to legitimate life-cycle characteristics is another concerning indication that these lenders are more likely to discriminate against African-American and Hispanic borrowers leading to disparate impact on mortgage market access for these borrowers.

The regressions in reported in Tables 2 through 4 allow different FICO/LTV buckets to have different constants, but otherwise all parameters are assumed to be the same across buckets. To investigate differences across the buckets in more detail, we also run 72 separate Oaxaca (1973) regressions, one per FICO/LTV bucket, and then plot the discrimination coefficient results in Figure 5. In Panel a, each group of bars representing one FICO bucket, with each bar within one group corresponding to a different LTV bucket. In Panel b, each group of bars representing one LTV bucket, with each bar within one group corresponding to a different FICO bucket. It can clearly be seen here that there is very significant discrimination in rejection rates for low-FICO borrowers. For FICO scores below 680, after controlling for observable characteristics, African-American and Hispanic borrowers are between 4% and 8% more likely to have their loans rejected than other borrowers, depending on LTV. As FICO increases beyond 680, the extent of this discrimination decreases, and all-but disappears for FICO scores above 740.

Overall, these results suggest that African-American and Hispanic borrowers do not have equal access to the mortgage market and that this differential access is due to illegitimate discrimination leading to disparate impact. Interestingly, the algorithmic underwriting practices of the FinTech lenders indicate a significantly lower incidence of discriminatory credit scoring and a greater reliance on similarly weighted and legitimate life-cycle screening factors in their origination decisions for African-American and Hispanic borrowers. In contrast for the largest lenders, the top 25 in terms of underwriting volume, we find an illegitimate discriminatory treatment effect that is positive and statistically significant of more than 5.67%

and a smaller effect of 3.85% among smaller non-FinTech lenders. These results imply that, after controlling for observable differences, African-American and Hispanic borrowers face significantly more illegitimate discrimination among, traditional lenders and large and small non-FinTech lenders. This result is in stark contrast with the -1% chance of being rejected by FinTech lenders. These results indicate an increased fairness in purchase mortgage lending access from algorithmic underwriting by FinTech lenders and potentially from concerns about exposure to put-back risks on the part of large lenders.

6 Conclusions

To investigate whether the use of machine-learning-based algorithmic underwriting has led to statistical discrimination among protected classes of borrowers, we analyze a mortgage data set that includes never-before-linked information at the loan-level on income, ethnicity, debt-to-income ratios, loan-to-value ratios, all contract terms, and indicators for whether the lender-of-record primarily used algorithmic underwriting. After controlling for observable differences, we find that conditional on obtaining a loan, African-American and Hispanic pay a slightly (0.08%) higher interest rate for purchase mortgages and about 0.03% higher mortgage interest rate for refinance mortgages. These differences are less pronounced among refi lenders who utilize algorithmic underwriting, for whom the differential is only 0.01% We additionally find that these borrowers are almost 5% more likely to be rejected for a mortgage than other borrowers.

We make important methodological contributions to current legal debates concerning the "robust statistical" measurement of *disparate impact*, as required by the courts. We also make contributions to debates within the economic literature concerning the identification of illegitimate discriminatory use of protected characteristics in credit screening. Our identification strategy relies on the unique screening and grid pricing strategies of the GSEs, and allows us to test for disparate impact even in a setting in which borrowers select to apply for a mortgage and where not all life-cycle variables, such as wealth, are observable. Since the legal standard for a disparate impact claim is that a court must ultimately conclude (i) that discrimination occurred and (ii) that no legitimate business necessity justifies the use of a sorting mechanism that discriminated, our empirical strategy successfully translates these requirements into economics in two parts. First, we successfully identify statistical discrimination in mortgage interest rates charged to borrowers via our application of the Oaxaca decomposition method. Second, we establish that the GSE underwriting and pricing structure insures that there is no omitted unobservable variable whose inclusion in a structural model is a legitimate business necessity. We then establish, again through appli-

cation of the Oaxaca decomposition, that lenders' accept/reject decisions, where credit risk is a legitimate business necessity based on the life-cycle characteristics of borrowers, also exhibits an important discriminatory component that is unexplained by legitimate life-cycle characteristics. This component is composed of differential treatment of ethnic minorities on average and on illegitimate differential weighting of the similar life-cycle variables between African-American/Hispanic borrowers and Causian/Asian borrowers.[17]

Our findings for more algorithmic FinTech mortgage lending suggest that, in addition to the efficiency gains of these innovations, they may also serve to make the mortgage lending markets more accessible to African-American and Hispanic borrowers and provide these borrowers with fairer pricing. This positive evaluation of the net benefits of algorithmic underwriting bodes wells for the needed expansion of U.S. residential mortgage markets as they continue to recover from the 2009 crisis. On a more cautionary note, however, the discipline imposed by the GSE's underwriting and pricing requirements are likely to be an important reason for the observed small differentials in mortgage-rejection rates between ethnic minorities and majority borrower populations among FinTech lenders. 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, outside of mortgage lending, the possibility remains that, lacking formal underwriting and pricing standards, lending algorithms might proxy for unobservable life-cycle variables by relying on observable characteristics (e.g., the name of an applicant's high school) that produce illegitimate statistical discrimination.

[17]Under the Supreme Court's Inclusive Communities framework, one can view our Oaxaca decomposition as providing a means to satisfy the Court's requirement that a plaintiff first establish a "robust" causal connection between a lending practice and an alleged statistical disparity before shifting the burden to the lender to establish a legitimate business necessity for the practice. While we lack information as to what specific practices are producing these discriminatory effects in GSE lending, the fact that the GSE underwriting and pricing structure relies entirely on observable life-cycle variables also permits the conclusion that, whatever practices produced these outcomes, they cannot be justified as lenders' attempts to proxy for unobservable life-cycle variables. Accordingly, the practices producing these discriminatory effects must be lacking in a legitimate business necessity, to the extent courts limit this defense to factors related to a borrower/applicant's credit risk.

Figure 1: LLPA Adjustments from Sample Rate Sheet

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: GSE Framework Providing Comparable Risk Pricing to Lenders for within-Grid Mortgages

The figure represents two mortgage applications with the same maturity, the same credit score and the same loan-to-value (LTV) ratio. The interest rate a household sees on a mortgage consists of three parts. First, all mortgages within the same credit score-LTV face the same market price, which emerges from the Base Mortgage Rate coming from the primary market for loans to be securitized by the GSEs. Second, all mortgages within the same maturity-credit score-LTV face the same G-fee added onto the market price. The G-fee, or guarantee fee, is the fee that the GSEs assign to insure the mortgage against default. The third part of the pricing is discretionary to the originator and is determined by the lenders revenue model and profits from their proprietary credit scoring.

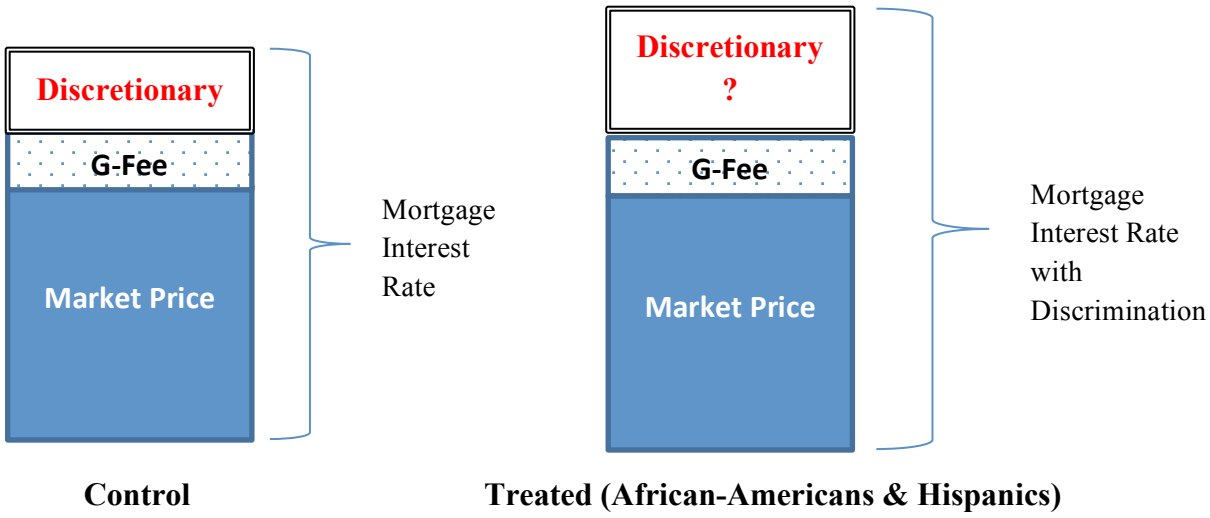
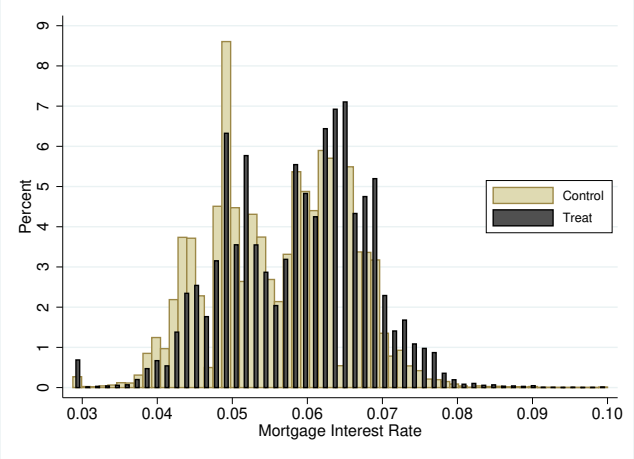
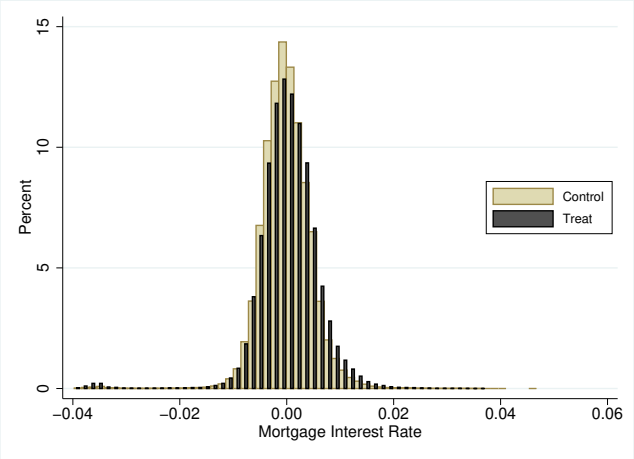


Figure 3: Borrower interest rates by treatment

(a) Raw data



(b) De-meaned by subtracting year-grid cell mean



(c) De-meaned by subtracting year-grid cell mean (debt sample)

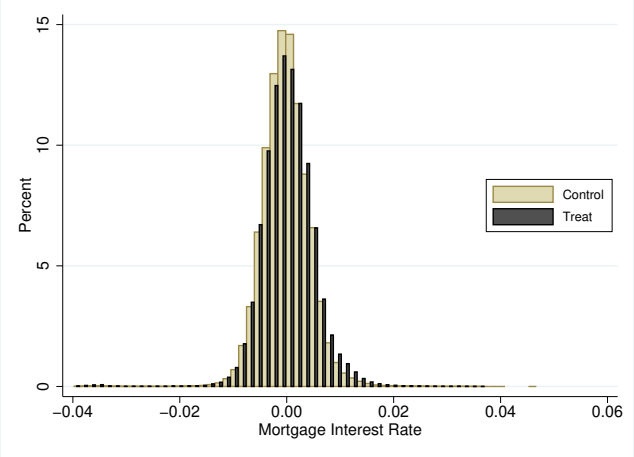


Figure 4: Comparison of loan application rejection probabilities for the raw and de-meaned to year buckets: for the control group (ethnicity $\neq$ African-American/Hispanic) versus the treatment group (ethnicity = African-American/Hispanic).

The figure compares the yearly loan application rejection probabilities (estimates for the raw and the demeaned to year buckets) for the full sample control group (ethnicity $\neq$ African-American/Hispanic) versus the treatment group (ethnicity = African-American/Hispanic).

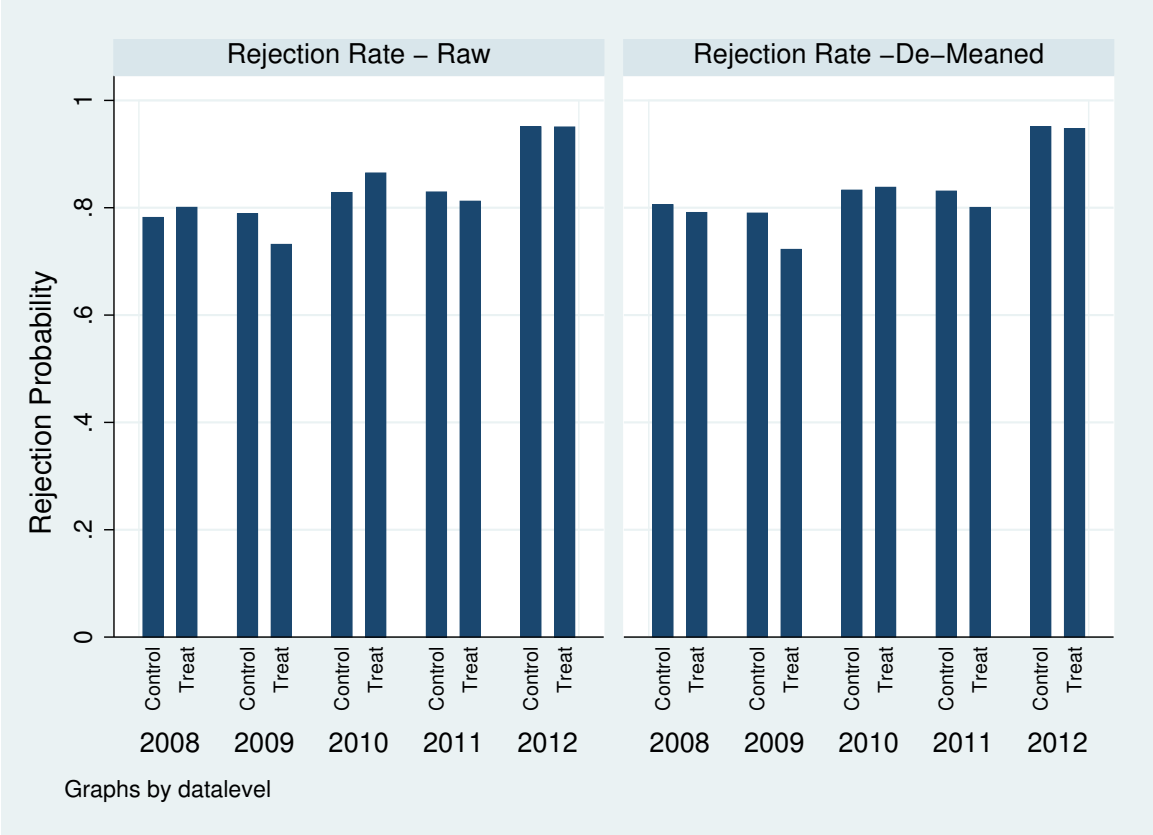
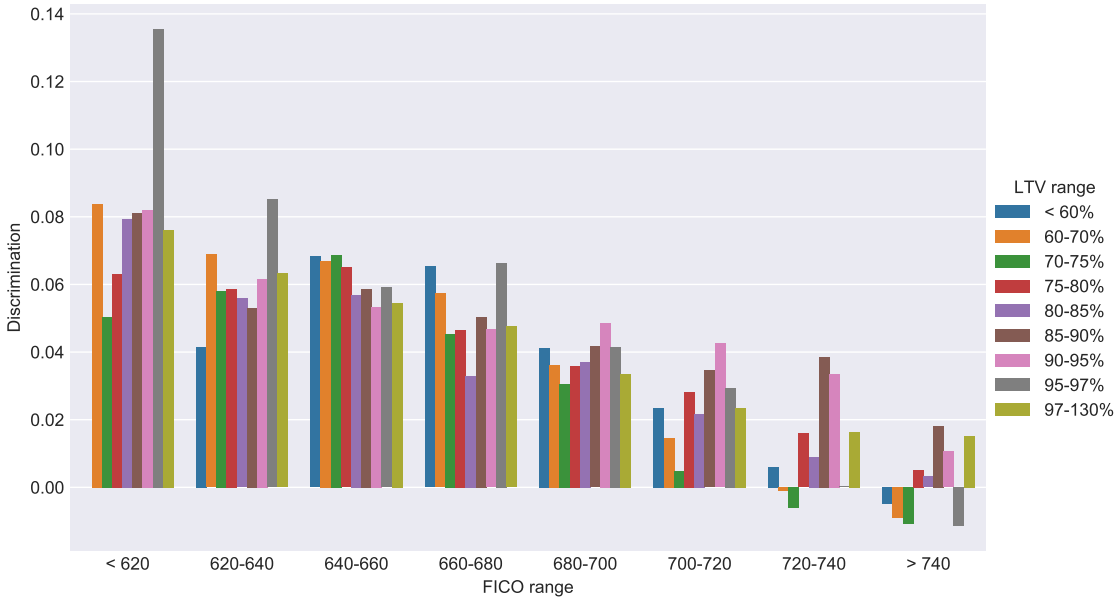
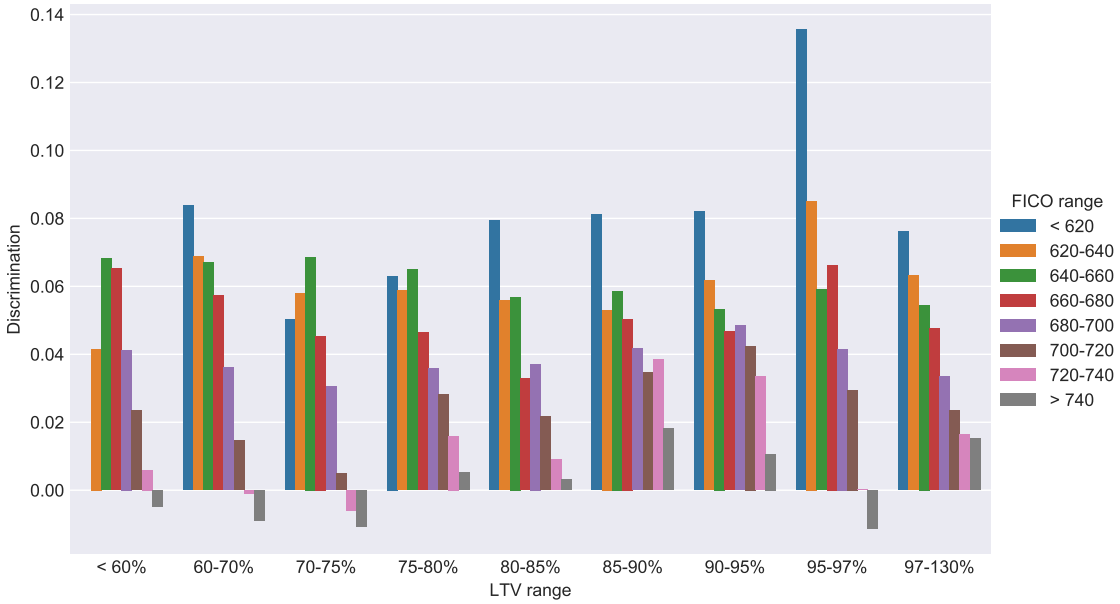


Figure 5: Discrimination in rejection rates by FICO and LTV

The figure show the discrimination coefficient from running Oaxaca (1973) regressions of purchase mortgage rejection rates on on various explanatory variables for each of 72 LTV/FICO buckets.



(a) Versus FICO



(b) Versus LTV

Table 1: Summary Statistics

Summary statistics for the originated and rejected loan samples. The information for the rejected loans was obtained from HMDA. As discussed in the Appendix, the credit scores for the rejected loans are proxied as the median vantage 3.0 score by census tract, as reported for all loans in the McDash data set. The loan-to-value ratios for rejected loan applications are proxied by the HMDA reported loan amount requested divided by the median house price in the loan census tract. The median house prices are measured as the assessed market value of all houses by census tract as reported in the Dataquick Assessor files by year. The originated loan data are from a merge of HMDA, Dataquick and McDash data sets, as discussed in the Appendix. The top 25 lenders are identified from annual loan origination data reported by *Inside Mortgage Finance*, for the years 2008 through 2012. The FinTech lenders were identified by lender name and were classified as FinTech due to their focus on on-line and algorithmic underwriting practices following the taxonomy introduced by Buchak et al. (2017).

Panel A: Rejected Applications (N = 2,970,922)					
	Mean	Standard Deviation	Minimum	Median	Maximum
Interest Rate %	--	--	--	--	--
Loan Amount $	197,163	122,407	20,000	169,000	793,000
Applicant Income $	85,268	69,700	10,000	69,000	999,000
Loan-to-Value	0.859	0.201	0.388	0.876	1.300
Credit Score	689	36	359	690	825
FinTech	0.041				
Top 25	0.497				
African_American/Hispanic	0.220				
Purchase=1; Refinance=0	0.157				

Panel B: Accepted Applications without Equifax match (N = 3,570,267)					
	Mean	Standard Deviation	Minimum	Median	Maximum
Interest Rate %	0.052	0.008	0.029	0.050	0.100
Loan Amount $	225,194	117,659	7,000	203,000	793,000
Applicant Income $	104,867	76,137	10,000	87,000	999,000
Loan-to-Value	0.764	0.196	0.190	0.785	1.300
Credit Score	717	75	359	725.5	850
FinTech	0.029				
Top 25	0.501				
African_American/Hispanic	0.132				
Purchase=1; Refinance=0	0.335				

Panel C: Accepted Applications with Equifax match (N = 819,424)					
	Mean	Standard Deviation	Minimum	Median	Maximum
Interest Rate %	0.053	0.008	0.029	0.051	0.100
Loan Amount $	230,216	110,809	10,000	210,000	600,000
Applicant Income $	101,370	71,745	10,000	85,000	999,000
Loan-to-Value	0.743	0.150	0.300	0.777	1.000
Credit Score	724	35	631	731	770
FinTech	0.028				
Top 25	0.505				
African_American/Hispanic	0.140				
Purchase=1; Refinance=0	0.385				
Debt (Other) Outstanding	41,995	60,399	0	22,645	3,177,989

Table 2: Rejection Discrimination: Main Results

Column (1) shows OLS regression of the mortgage rejection rate on the loan-to-value ratio, the credit score, and the log income of the borrower; Column (2) includes month and bucket fixed effects. Columns (3)–(5) show results of Oaxaca (1973) regression of purchase mortgage coupon rate on the same explanatory variables.

	Dependent Variable: Rejection				
	(1)	(2)	(3)	(4)	(5)
Model	OLS	OLS	Oaxaca Decomposition		
			Difference in Means	Explained	Unexplained
Discrimination	0.0484***	0.0423***	0.152***	0.103***	0.0491***
	[0.000470]	[0.000450]	[0.000524]	[0.000345]	[0.000458]
Loan-to-Value	0.400***	0.451***		0.0178***	0.003
	[0.000865]	[0.00282]		[0.000162]	[0.00635]
Credit Score	-0.00113***	-0.000237***		0.0151***	1.486***
	[3.22e-06]	[1.02e-05]		[0.000247]	[0.0185]
Log Income	-0.0956***	-0.0746***		0.0159***	-0.00217
	[0.000282]	[0.000270]		[7.86e-05]	[0.00790]
Observations	6,177,674	6,177,670	6,177,670		
R-squared	0.27	0.348			
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		
Lender FE	Y	Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Rejection Discrimination: Results by Traditional Lenders vs. FinTech Lenders

Columns (1)–(3) show results of Oaxaca (1973) regression of the mortgage rejection rate for the traditional (non-FinTech) lenders on the loan-to-value ratio, the credit score, the log income of the borrower and year and bucket fixed effects. Columns (4)–(6) show results of Oaxaca (1973) regression of purchase mortgage rejection rate for the FinTech lenders on the same explanatory variables.

	Dependent Variable: Rejection					
	(1)	(2)	(3)	(4)	(5)	(6)
	Traditional Lenders			FinTech Lenders		
	Oaxaca Decomposition			Oaxaca Decomposition		
Model	Difference in Means	Explained	Unexplained	Difference in Means	Explained	Unexplained
Discrimination	0.157***	0.107***	0.0508***	-0.00682**	0.00238	-0.00920***
	[0.000531]	[0.000350]	[0.000463]	[0.00335]	[0.00208]	[0.00273]
Loan-to-Value		0.0179***	0.00269		0.0100***	0.0949**
		[0.000167]	[0.00642]		[0.000678]	[0.0396]
Credit Score		0.0138***	1.444***		0.0256***	1.646***
		[0.000253]	[0.0187]		[0.00127]	[0.122]
Log Income		0.0161***	-0.0000351		0.00504***	0.047
		[8.04e-05]	[0.00798]		[0.000241]	[0.0517]
Observations	5,951,291			226,379		
Year FE	Y			Y		
Bucket FE	Y			Y		
Lender FE	Y			Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Rejection Discrimination: Results by Top 25 Lenders vs. Smaller Non-FinTech Lenders

Columns (1)–(3) show results of Oaxaca (1973) regression of the mortgage rejection rate for the top 25 lenders on the loan-to-value ratio, the credit score, the log income of the borrower and year and bucket fixed effects. Columns (4)–(6) show results of Oaxaca (1973) regression of purchase mortgage rejection rate for the smaller lenders on the same explanatory variables.

	Dependent Variable: Rejection					
	(1)	(2)	(3)	(4)	(5)	(6)
	Top 25 Lenders			Small Lenders		
Model	Oaxaca Decomposition			Oaxaca Decomposition		
	Difference in Means	Explained	Unexplained	Difference in Means	Explained	Unexplained
Discrimination	0.145***	0.0885***	0.0567***	0.161***	0.123***	0.0385***
	[0.000709]	[0.000404]	[0.000646]	[0.000749]	[0.000567]	[0.000616]
Loan-to-Value		0.0202***	0.00332		0.0147***	0.00177
		[0.000237]	[0.00904]		[0.000210]	[0.00837]
Credit Score		0.0257***	1.454***		0.00825***	1.497***
		[0.000389]	[0.0281]		[0.000301]	[0.0232]
Log Income		0.0172***	-0.104***		0.0154***	0.139***
		[0.000115]	[0.0111]		[0.000104]	[0.0106]
Observations	3,263,880			3,192,159		
Year FE	Y			Y		
Bucket FE	Y			Y		
Lender FE	Y			Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Interest Rate Discrimination: Main Results for Purchase Loans

Panel A: Column (1) shows OLS regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower; Column (2) includes month and bucket fixed effects. Columns (3)–(5) show results of Oaxaca (1973) regression of purchase mortgage coupon rate on the same explanatory variables. Panel B: Reports the same specifications with the addition of a control for other consumer debt held by the borrower using a sample that merges the Equifax debt data; Panel C: Adds lender fixed effects to the Panel B specifications.

Panel A: Full Sample

	Dependent Variable: Mortgage Interest Rate				
	(1)	(2)	(3)	(4)	(5)
Model	OLS	OLS	Oaxaca Decomposition		
			Difference in Means	Explained	Unexplained
Discrimination	0.00153***	0.000899***	0.00185***	0.000941***	0.000911***
	[2.64e-05]	[1.59e-05]	[2.69e-05]	[2.18e-05]	[1.60e-05]
Loan-to-Value	0.00183***	-0.00238***		-6.48e-05***	8.07E-06
	[6.40e-05]	[0.000145]		[4.15e-06]	[0.000402]
Credit Score	-1.25e-05***	2.59e-06***		-3.49e-05***	6.26E-05
	[1.06e-07]	[3.12e-07]		[4.14e-06]	[0.000616]
Log Income	-0.000492***	-0.000443***		7.65e-05***	-0.00507***
	[1.44e-05]	[8.35e-06]		[1.87e-06]	[0.000295]
Observations	866,882	848,148	848,148		
R-squared	0.027	0.692			
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		

Panel B: Sample that Merges with Equifax Debt Data

	Dependent Variable: Mortgage Interest Rate				
	(1)	(2)	(3)	(4)	(5)
Model	OLS	OLS	Oaxaca Decomposition		
			Difference in Means	Explained	Unexplained
Discrimination	0.00127***	0.000822***	0.00167***	0.000836***	0.000833***
	[3.98e-05]	[2.32e-05]	[4.05e-05]	[3.33e-05]	[2.32e-05]
Loan-to-Value	0.00226***	-0.00213***		-4.26e-05***	-0.00119
	[0.000113]	[0.000298]		[6.96e-06]	[0.000775]
Credit Score	-4.38e-05***	-8.41e-06***		7.04e-05***	0.00321
	[4.23e-07]	[1.07e-06]		[9.09e-06]	[0.00235]
Log Income	-0.000585***	-0.000549***		8.99e-05***	-0.00486***
	[2.43e-05]	[1.38e-05]		[3.05e-06]	[0.000453]
Log Debt	0.000259***	7.18e-05***		-1.53e-05***	7.87E-05
	[7.18e-06]	[4.07e-06]		[1.19e-06]	[0.000103]
Observations	315,257	311,911	311,911		
R-squared	0.052	0.707			
Debt Covariates	Y	Y	Y		
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		

Panel C: Sample that Merges with Equifax Debt Data including Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate				
	(1)	(2)	(3)	(4)	(5)
Model	OLS	OLS	Oaxaca Decomposition		
			Difference in Means	Explained	Unexplained
Discrimination	0.00132***	0.000758***	0.00186***	0.00109***	0.000769***
	[2.53e-05]	[1.64e-05]	[2.76e-05]	[2.24e-05]	[1.66e-05]
Loan-to-Value	0.00238***	-0.00146***		-4.96e-05***	0.00131***
	[6.15e-05]	[0.000148]		[4.57e-06]	[0.000397]
Credit Score	-1.12e-05***	2.24e-06***		-3.06e-05***	8.82E-05
	[1.14e-07]	[3.19e-07]		[4.38e-06]	[0.000629]
Log Income	-0.000584***	-0.000389***		6.75e-05***	-0.00428***
	[1.41e-05]	[8.82e-06]		[1.94e-06]	[0.000309]
Observations	785,899	785,899	785,899		
R-squared	0.197	0.696			
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		
Lender FE	Y	Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Interest Rate Discrimination: Results for Purchase Loans by Traditional Lenders vs. FinTech Lenders

Panel A: Column (1) shows OLS regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects. Columns (2)–(4) show the results of Oaxaca (1973) regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects for the traditional (non-FinTech) lenders. Columns (5)–(7) reports the same regressions for the FinTech lenders. Panel B: Repeats the same specifications for traditional (non-FinTech) and FinTech lenders including lender fixed effects.

Panel A: Without Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Traditional Lenders				FinTech Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000899***	0.00189***	0.000980***	0.000913***	0.000793***	0.00140***	0.000596***	0.000805***
	[1.66e-05]	[2.79e-05]	[2.26e-05]	[1.66e-05]	[0.000102]	[0.000140]	[9.97e-05]	[0.000104]
Loan-to-Value	-0.00168***		-5.45e-05***	0.00104***	-0.00184*		-6.91e-05*	2.16E-06
	[0.000151]		[4.61e-06]	[0.000404]	[0.00103]		[3.98e-05]	[0.00283]
Credit Score	2.57e-06***		-3.45e-05***	0.000348	-3.04E-06		4.22E-06	-0.0113**
	[3.26e-07]		[4.44e-06]	[0.000635]	[2.17e-06]		[2.29e-05]	[0.00463]
Log Income	-0.000427***		7.30e-05***	-0.00508***	-0.000709***		0.000143***	-0.00736***
	[8.86e-06]		[1.96e-06]	[0.000310]	[5.87e-05]		[1.61e-05]	[0.00189]
Observations	768,408	768,408			17,491	17,491		
R-Square	0.691				0.571			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		

Panel B: With Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Traditional Lenders				FinTech Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000758***	0.00189***	0.00113***	0.000768***	0.000699***	0.00140***	0.000685***	0.000717***
	[1.67e-05]	[2.79e-05]	[2.28e-05]	[1.68e-05]	[0.000102]	[0.000141]	[0.000100]	[0.000104]
Loan-to-Value	-0.00144***		-4.88e-05***	0.00139***	-0.00159		-5.76E-05	-0.00013
	[0.000150]		[4.57e-06]	[0.000401]	[0.00102]		[3.95e-05]	[0.00282]
Credit Score	2.45e-06***		-3.25e-05***	0.000415	-3.13E-06		5.78E-06	-0.0113**
	[3.22e-07]		[4.42e-06]	[0.000636]	[2.15e-06]		[2.27e-05]	[0.00461]
Log Income	-0.000382***		6.60e-05***	-0.00421***	-0.000658***		0.000133***	-0.00685***
	[8.93e-06]		[1.95e-06]	[0.000313]	[5.85e-05]		[1.58e-05]	[0.00191]
Observations	768,408	768,408			17,491	17,491		
R-Square	0.696				0.578			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		
Lender FE	Y	Y			Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Interest Rate Discrimination: Results for Purchase Loans by Top 25 Lenders vs. Smaller Lenders

Panel A: Column (1) shows OLS regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects. Columns (2)–(4) show the results of Oaxaca (1973) regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects for the large (top 25) lenders. Columns (5)–(7) reports the same regressions for the smaller (non-top-25) lenders. Panel B: Repeats the same specifications for the large (top 25) lenders and the smaller (non-top-25) lenders including lender fixed effects.

Panel A: Without Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 25 Lenders				Small Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000890***	0.00224***	0.00133***	0.000908***	0.000911***	0.00151***	0.000585***	0.000921***
	[2.28e-05]	[3.83e-05]	[3.06e-05]	[2.29e-05]	[2.41e-05]	[4.06e-05]	[3.32e-05]	[2.42e-05]
Loan-to-Value	-0.00247***		-8.03e-05***	0.00132**	-0.000692***		-2.27e-05***	0.00056
	[0.000205]		[6.30e-06]	[0.000549]	[0.000220]		[6.80e-06]	[0.000596]
Credit Score	3.28e-06***		-4.54e-05***	0.00187**	2.30e-06***		-3.05e-05***	-0.00179*
	[4.36e-07]		[6.90e-06]	[0.000878]	[4.89e-07]		[5.39e-06]	[0.000929]
Log Income	-0.000385***		6.09e-05***	-0.00515***	-0.000456***		8.42e-05***	-0.00487***
	[1.22e-05]		[2.49e-06]	[0.000427]	[1.30e-05]		[3.08e-06]	[0.000453]
Observations	412,530	412,530			355,886	355,886		
R-Square	0.689				0.694			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		

Panel B: With Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 25 Lenders				Small Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000802***	0.00224***	0.00142***	0.000819***	0.000685***	0.00151***	0.000812***	0.000694***
	[2.28e-05]	[3.83e-05]	[3.07e-05]	[2.29e-05]	[2.44e-05]	[4.07e-05]	[3.37e-05]	[2.46e-05]
Loan-to-Value	-0.00227***		-7.56e-05***	0.00163***	-0.000358		-1.47e-05**	0.000993*
	[0.000204]		[6.26e-06]	[0.000547]	[0.000219]		[6.72e-06]	[0.000590]
Credit Score	2.15e-06***		-2.80e-05***	0.00162*	2.40e-06***		-3.04e-05***	-0.00163*
	[4.28e-07]		[6.90e-06]	[0.000890]	[4.83e-07]		[5.34e-06]	[0.000923]
Log Income	-0.000373***		5.95e-05***	-0.00471***	-0.000388***		7.30e-05***	-0.00344***
	[1.22e-05]		[2.49e-06]	[0.000428]	[1.31e-05]		[3.05e-06]	[0.000459]
Observations	412,530	412,530			355,886	355,886		
R-Square	0.691				0.702			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		
Lender FE	Y	Y			Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Interest Rate Discrimination: Main Results for Refi Mortgages

Panel A: Column (1) shows OLS regression of the refi mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower; Column (2) includes month and bucket fixed effects. Columns (3)–(5) show results of Oaxaca (1973) regression of purchase mortgage coupon rate on the same explanatory variables. Panel B: Reports the same specifications with the addition of a control for other consumer debt held by the borrower using a sample that merges the Equifax debt data; Panel C: Adds lender fixed effects to the Panel B specifications.

Panel A: Full Sample					
	Dependent Variable: Mortgage Interest Rate				
	(1)	(2)	(3)	(4)	(5)
				Oaxaca Decomposition	
Model	OLS	OLS	Difference in Means	Explained	Unexplained
Discrimination	0.00154***	0.000453***	0.00205***	0.00160***	0.000456***
	[1.99e-05]	[1.27e-05]	[2.13e-05]	[1.72e-05]	[1.27e-05]
Loan-to-Value	0.00222***	-0.000374***		-8.09e-07***	0.000313**
	[2.94e-05]	[5.86e-05]		[2.28e-07]	[0.000144]
Credit Score	-1.31e-05***	5.33e-06***		-4.51e-05***	0.00613***
	[7.53e-08]	[2.41e-07]		[2.45e-06]	[0.000454]
Log Income	-0.00177***	-0.000816***		0.000138***	-0.00190***
	[1.01e-05]	[6.39e-06]		[1.67e-06]	[0.000248]
Observations	1,461,437	1,402,610	1,402,610		
R-squared	0.059	0.655			
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		

Panel B: Sample that Merges with Equifax Debt Data					
	Dependent Variable: Mortgage Interest Rate				
	(1)	(2)	(3)	(4)	(5)
				Oaxaca Decomposition	
Model	OLS	OLS	Difference in Means	Explained	Unexplained
Discrimination	0.00134***	0.000381***	0.00193***	0.00154***	0.000382***
	[3.03e-05]	[1.84e-05]	[3.19e-05]	[2.63e-05]	[1.85e-05]
Loan-to-Value	0.00327***	-0.000372***		2.02e-06**	-0.00013
	[6.29e-05]	[0.000130]		[8.89e-07]	[0.000295]
Credit Score	-5.25e-05***	-1.56e-05***		0.000104***	0.00086
	[3.02e-07]	[7.96e-07]		[6.14e-06]	[0.00188]
Log Income	-0.00171***	-0.000783***		0.000129***	-0.00108***
	[1.71e-05]	[1.04e-05]		[2.60e-06]	[0.000370]
Log Debt	0.000225***	0.000137***		-1.05e-05***	-0.000101
	[5.43e-06]	[3.32e-06]		[1.11e-06]	[9.92e-05]
Observations	504,167	503,029	503,029		
R-squared	0.103	0.669			
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		

Panel C: Sample that Merges with Equifax Debt Data including Lender Fixed Effects					
	Dependent Variable: Mortgage Interest Rate				
	(1)	(2)	(3)	(4)	(5)
				Oaxaca Decomposition	
Model	OLS	OLS	Difference in Means	Explained	Unexplained
Discrimination	0.00101***	0.000290***	0.00197***	0.00168***	0.000289***
	[1.87e-05]	[1.26e-05]	[2.11e-05]	[1.71e-05]	[1.28e-05]
Loan-to-Value	0.00259***	0.000585***		3.01e-06***	0.000350**
	[2.77e-05]	[5.59e-05]		[4.41e-07]	[0.000139]
Credit Score	-9.57e-06***	4.27e-06***		-3.28e-05***	0.00580***
	[7.42e-08]	[2.36e-07]		[2.33e-06]	[0.000451]
Log Income	-0.00127***	-0.000717***		0.000122***	-0.00132***
	[9.66e-06]	[6.46e-06]		[1.59e-06]	[0.000249]
Observations	1,354,399	1,354,399	1,354,399		
R-squared	0.24	0.659			
Year FE	N	Y	Y		
Bucket FE	N	Y	Y		
Lender FE	Y	Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Interest Rate Discrimination: Results for Refinance Mortgages by Traditional Lenders vs. FinTech Lenders

Panel A: Column (1) shows OLS regression of the refinance mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects for the traditional (non-FinTech) lenders. Columns (2)–(4) show the results of Oaxaca (1973) regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects for the traditional (non-FinTech) lenders. Columns (5)–(7) reports the same regressions for the FinTech lenders. Panel B: Repeats the same specifications for traditional (non-FinTech) and FinTech lenders including lender fixed effects.

Panel A: Without Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Traditional Lenders				FinTech Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000443***	0.00199***	0.00155***	0.000446***	0.000165***	0.000737***	0.000576***	0.000161***
	[1.29e-05]	[2.15e-05]	[1.73e-05]	[1.29e-05]	[6.02e-05]	[9.75e-05]	[7.88e-05]	[6.09e-05]
Loan-to-Value	0.000385***		1.83e-06***	0.000457***	0.00201***		6.26E-06	0.00106
	[5.78e-05]		[4.08e-07]	[0.000143]	[0.000282]		[5.20e-06]	[0.000719]
Credit Score	5.14e-06***		-4.26e-05***	0.00632***	-7.60E-07		2.82E-06	-0.000469
	[2.47e-07]		[2.52e-06]	[0.000459]	[1.02e-06]		[3.88e-06]	[0.00234]
Log Income	-0.000775***		0.000134***	-0.00187***	-0.00101***		8.83e-05***	0.000142
	[6.59e-06]		[1.69e-06]	[0.000252]	[3.15e-05]		[8.49e-06]	[0.00128]
Observations	1,300,245	1,300,245			54,154	54,154		
R-Square	0.649				0.634			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		

Panel B: With Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Traditional Lenders				FinTech Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000297***	0.00199***	0.00170***	0.000296***	0.000136**	0.000737***	0.000605***	0.000132**
	[1.29e-05]	[2.15e-05]	[1.74e-05]	[1.30e-05]	[5.95e-05]	[9.75e-05]	[7.90e-05]	[6.02e-05]
Loan-to-Value	0.000525***		2.80e-06***	0.000332**	0.00212***		6.67E-06	0.00102
	[5.71e-05]		[4.46e-07]	[0.000142]	[0.000278]		[5.53e-06]	[0.000709]
Credit Score	4.57e-06***		-3.65e-05***	0.00595***	-1.36E-06		4.84E-06	-0.000902
	[2.42e-07]		[2.46e-06]	[0.000460]	[1.01e-06]		[4.11e-06]	[0.00231]
Log Income	-0.000710***		0.000123***	-0.00140***	-0.000881***		7.67e-05***	6.58E-05
	[6.60e-06]		[1.63e-06]	[0.000254]	[3.15e-05]		[7.50e-06]	[0.00128]
Observations	1,300,245	1,300,245			54,154	54,154		
R-Square	0.657				0.643			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		
Lender FE	Y	Y			Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Interest Rate Discrimination: Results for Refinance Mortgages by Top 25 Lenders vs. Smaller Lenders

Panel A: Column (1) shows OLS regression of the refinance mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects for the large (top 25) lenders. Columns (2)–(4) show the results of Oaxaca (1973) regression of the purchase mortgage coupon rate on the loan-to-value ratio, the credit score, and the log income of the borrower including month and bucket fixed effects for the large (top 25) lenders. Columns (5)–(7) reports the same regressions for the smaller (non-top-25) lenders. Panel B: Repeats the same specifications for the large (top 25) lenders and the smaller (non-top-25) lenders including lender fixed effects.

Panel A: Without Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 25 Lenders				Small Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000368***	0.00165***	0.00128***	0.000373***	0.000463***	0.00257***	0.00210***	0.000467***
	[1.58e-05]	[2.58e-05]	[2.04e-05]	[1.59e-05]	[2.16e-05]	[3.73e-05]	[3.03e-05]	[2.18e-05]
Loan-to-Value	0.000519***		4.03e-06***	0.000485***	0.000230**		1.80E-07	0.000242
	[7.32e-05]		[8.02e-07]	[0.000173]	[9.28e-05]		[1.77e-07]	[0.000249]
Credit Score	2.62e-06***		-1.55e-05***	0.00570***	8.05e-06***		-8.24e-05***	0.00768***
	[3.07e-07]		[3.24e-06]	[0.000576]	[4.00e-07]		[4.91e-06]	[0.000753]
Log Income	-0.000714***		0.000129***	-0.00133***	-0.000809***		0.000130***	-0.00240***
	[8.30e-06]		[2.15e-06]	[0.000303]	[1.07e-05]		[2.63e-06]	[0.000441]
Observations	739,530	739,530			560,718	560,718		
R-Square	0.622				0.682			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		

Panel B: With Lender Fixed Effects

	Dependent Variable: Mortgage Interest Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 25 Lenders				Small Lenders			
		Oaxaca Decomposition				Oaxaca Decomposition		
Model		Difference in				Difference in		
	OLS	Means	Explained	Unexplained	OLS	Means	Explained	Unexplained
Discrimination	0.000268***	0.00165***	0.00138***	0.000269***	0.000338***	0.00257***	0.00223***	0.000336***
	[1.58e-05]	[2.58e-05]	[2.05e-05]	[1.59e-05]	[2.18e-05]	[3.74e-05]	[3.08e-05]	[2.22e-05]
Loan-to-Value	0.000567***		4.64e-06***	0.000385**	0.000400***		3.67E-07	0.000164
	[7.24e-05]		[8.13e-07]	[0.000171]	[9.22e-05]		[3.15e-07]	[0.000247]
Credit Score	2.03e-06***		-9.63e-06***	0.00499***	7.21e-06***		-7.36e-05***	0.00673***
	[3.00e-07]		[3.20e-06]	[0.000578]	[3.83e-07]		[4.72e-06]	[0.000755]
Log Income	-0.000657***		0.000119***	-0.00107***	-0.000760***		0.000123***	-0.00188***
	[8.26e-06]		[2.07e-06]	[0.000304]	[1.08e-05]		[2.57e-06]	[0.000451]
Observations	739,530	739,530			560,718	560,718		
R-Square	0.629				0.691			
Year FE	Y	Y			Y	Y		
Bucket FE	Y	Y			Y	Y		
Lender FE	Y	Y			Y	Y		

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

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.

Arrow, K., 1973, Higher education as a filter, *Journal of Public Economics* 2, 193–216.

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, University of Chicago.

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., 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.

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

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.

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.

Kline, P., 2011, Oaxaca-Blinder as a reweighting estimator, *American Economic Review: Papers & Proceedings* 101, 532–537.

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

Oaxaca, R., 1973, Male-female wage differentials in urban labor markets, *International Economic Review* 14, 693–709.

Oaxaca, R. L., and M. R. Ransom, 1999, Identification in detailed wage decompositions, *Review of Economics and Statistics* 81, 154–157.

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 mortgage markets, *Journal of Housing Research* 4, 315–336.

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).

Słoczyński, T., 2015, The Oaxaca-Blinder unexplained component as a treatment effects estimator, *Oxford Bulletin of Economics and Statistics* 77, 588–604.

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

A 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 CoreLogic DataQuick dataset, which provides detailed property transaction and ownership information in addition to a time series history of all recorded mortgage lien events such new mortgage originations, prepayments, REO, foreclosure, short sales, and arms-length sales and loan payoffs. The algorithm developed to match these two data sets relies on matching distinct loan characteristics, e.g. origination date, loan amount, and termination and distress events so as to precisely match each loan with its counterpart in the other dataset using a modified k-nearest neighbor classifier (see Hastie, Tibshirani, and Friedman, 2009; James, Witten, Hastie, and Tibshirani, 2015).

A.1 Merge process for newly originated fixed rate GSE loans

There are two steps to our merging process. The first step is to serialize the Dataquick data into a record format in which each property is assigned a full event history string for each mortgage lien and the priority and performance of these mortgage loan positions. Dataquick provides very comprehensive geographic coverage for mortgage originations and terminations at the property-level for most five digit zip codes and most census tracks in the U.S. It accounts for about 90 percent of property Assessor's Pin Numbers and all the mortgage and lien recording for each property all sourced from public records. The second stage of the merging process is to employ functions for k-nearest neighbor algorithms using *sklearn.neighbors* in Python to fit radial kernels using *BallTree*. The k-neighbors classifier implements learning based on the 25 nearest neighbors in the corresponding zip code within the McDash mortgage data that also records the loan contract features and a loan-level string of performance characteristics. 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 are assigned a *category subscore* that is normalized to a float value between 0 and 1. Each sub-score is them squared to achieve a greater penalty for matches on key elements such as the loan amount. The category subscore is then

scaled by a *category factor* which represents the categories importance to the match quality relative to other elements uses in the match. Each category factor is an integer between 0 and100 and the sum of the category factors is equal to 100. Our scoring algorithm (*get.score* in Python's sklearn) takes into account the 13 different elements of each matched pair of loans in order 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 good quality matches with scores of 2000 or less. The merge rates for the Dataquick to McDash data sets using the modified $k - nearestneighbor$ algorithm for loans originated between 2008 and 2012 is 14.76 million loans. of these 8.8 million loans are GSE loans 90% of those are good merges.

Our prior machine learning strategy is less applicable for the merge of the HMDA data to McDash data, because we have only origination data in HMDA and as well as a greatly reduced set of loan characteristics at origination including: 1) the regulator type, 2) the loan type (conventional), 3) property type (1 to 4 single family residential properties, 4) loan purpose (refinance or purchase), 5) occupancy status, 6) original loan amount, 7) MSA, state, county and census tract, 8) self reported borrower and co-borrower ethnicity (African-American, Hispanic, Asian, Caucasian, unknown), 9) borrower/coborrower gender, 10) borrower annual income, 11) year of origination, 12) denial reason if loan application is rejected, and 13) lender name. For this merge, we instead we standardize the lenders names between Dataquick and HMDA and then merge these data sets using lender names, loan amount, lien type, and the loan purpose fields. Of the 30.6 million originations in the HMDA data sets and about 10.4 million GSE loans, we successfully merged 60% of these loans. We then merged the Dataquick to McDash merged data to the Dataquick to HMDA merged data using the crosswalk developed with the k-nearest neighbor algorithm and we obtained a final data set of 3.47 million loans that are single family fixed rate GSE loans originated between 2008 and 2012.

A.2 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/Dataquick/McDash data set of fixed rate GSE loans that were originated between 2008 and 2012 to a random sample of 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.

A.3 The HMDA sample of rejected conventional loans

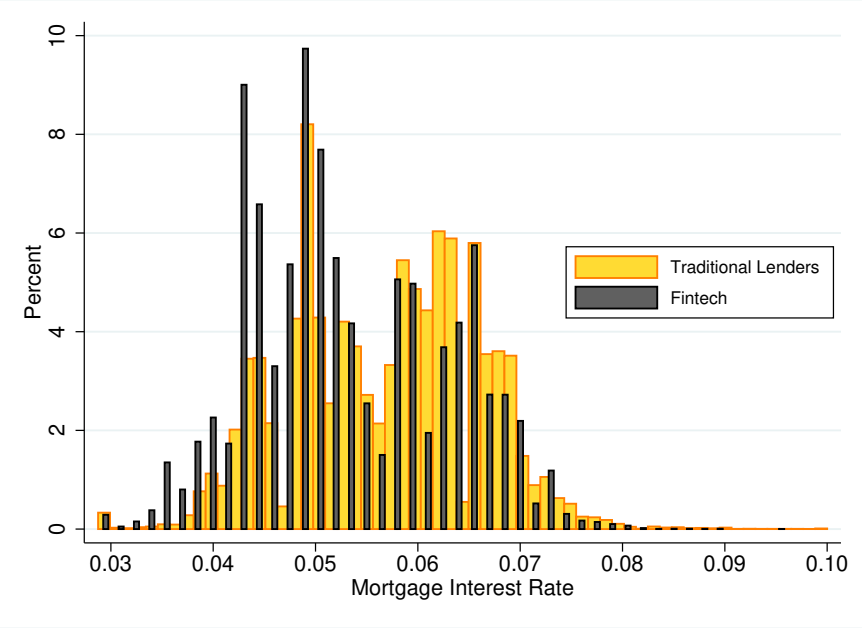
The second important class of loans in our data set includes all of the conventional conforming loans in HMDA for 1-4 family residential borrowers whose loan applications were either denied by the originator, approved but not accepted by the borrower, withdrawn by the applicant, or the loan application file was closed for incompleteness. These data include information on: 1) the regulator type, 2) the loan type (conventional), 3) property type (1 to 4 single family residential properties, 4) loan purpose (refinance or purchase), 5) occupancy status, 6) original loan amount, 7) MSA, state, county and census tract, 8) self reported borrower and co-borrower ethnicity (African-American, Hispanic, Asian, Caucasian, unknown), 9) borrower/co-borrower gender, 10) borrower annual income, 11) year of origination, 12) denial reason if loan application is rejected, and 13) lender name. They also include information on demographic and minority representation in the census tract in which the collateral on the loan is located. These variables include: the Federal Financial Institutions Examination Council (FFIEC) tract median family income to MSA median family income as a percentage, median family income for tract in thousands of dollars, tract population in thousands, tract minority population as a percentage, tract number of owner occupied units in thousands,tract number of 1- to 4-Family units in thousands.

HMDA does not include information on the credit score or the loan-to-value ratio of the rejected loan application files. For this reason, we proxy for the loan-to-value ratio by computing the mean assessed market value of all houses in each census tract in the U.S. using the panel of assessor's data from Dataquick. For each assessment we also have the year of the assessment. We then compute the ratio of the requested loan balance to the median value of all the homes in the appropriate census tract and year as reported by Dataquick to compute an estimated loan-to-value ratio for the rejected loan application. We proxy for the applicant's unobserved credit scores by using the McDash Vantage 3.0 score to compute the median credit score for each census track reported in McDash. The median census tract credit score is applied as a proxy for the credit score of the rejected loan.

B Robustness: Loan distribution by lender type

Figure 6: Borrower interest rates by treatment, FinTech vs. traditional lenders

(a) Raw data



(b) De-meaned by subtracting year-grid cell mean

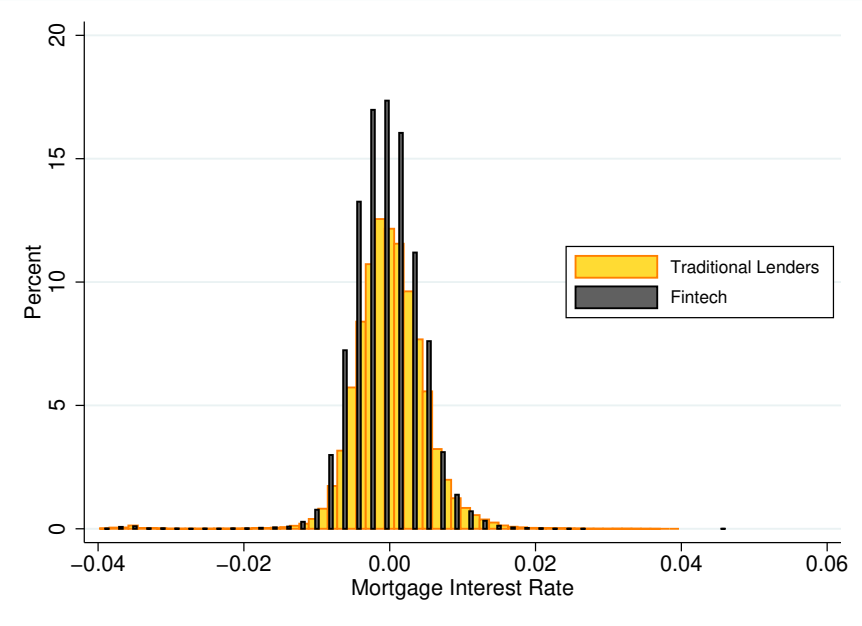
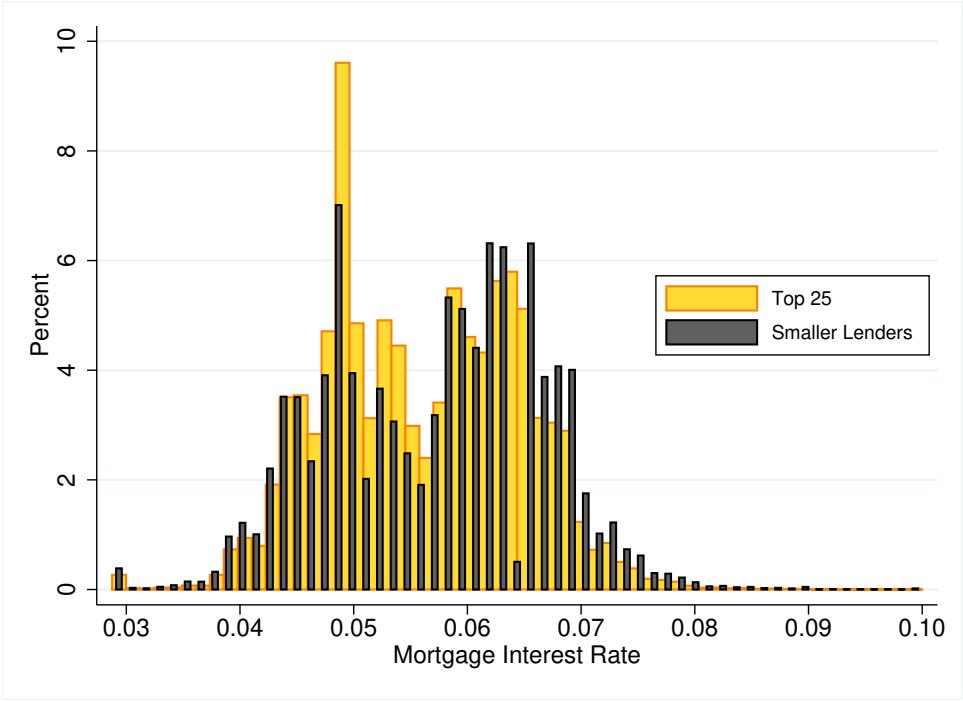
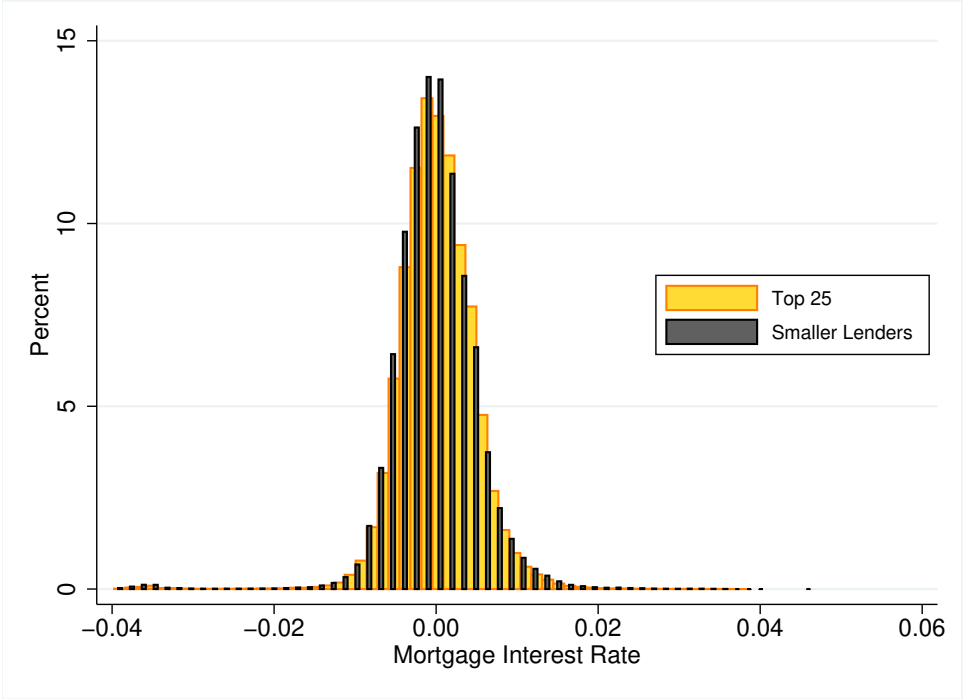


Figure 7: Borrower interest rates by treatment, large vs. small lenders

(a) Raw data



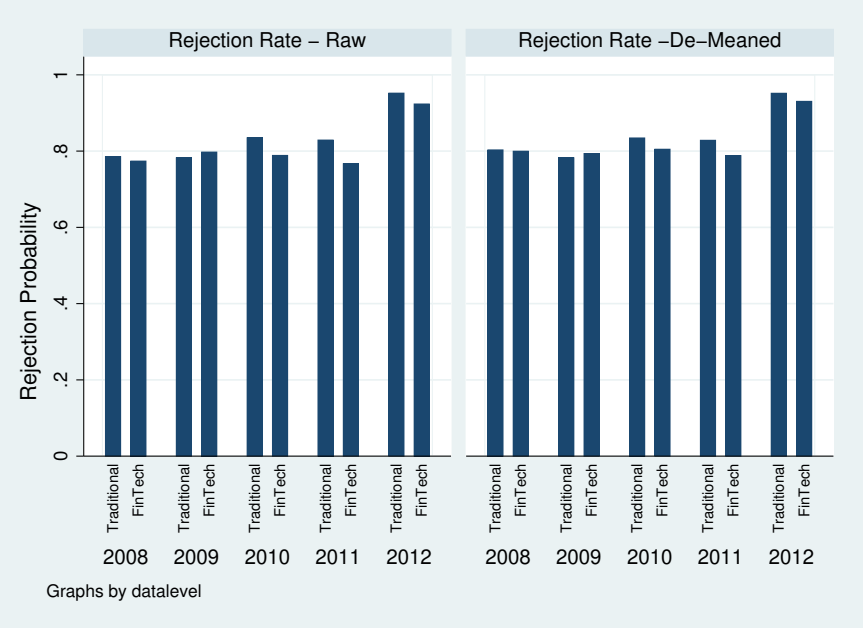
(b) De-meaned by subtracting year-grid cell mean



C Robustness: Rejection rates by lender type

Figure 8: Rejection rates by lender type

(a) FinTech versus traditional



(b) Large versus small

