Racial Discrimination in the Auto Loan Market

ALEXANDER W. BUTLER[1]
ERIK J. MAYER[2]
JAMES P. WESTON[3]

March 31, 2021

Corresponding author:
Erik Mayer
emayer@smu.edu
Cox School of Business, 6212 Bishop Boulevard, Dallas, TX 75275

JEL Codes: D14, D18, G20, G28, G50, J15

Acknowledgements: For helpful comments we thank Carlos Fernando Avenancio-León, Lisa Cook, Tony Cookson, Michael Ferguson, John Griffin, Charles Hadlock, and seminar participants at Indiana University, Rice University, Southern Methodist University, Texas Christian University, American University, Deakin University, University of Otago, University of Wisconsin-Milwaukee, and the University at Buffalo. We also thank conference participants at the Western Finance Association, the European Finance Association, the Lone Star Finance Conference, the FDIC Consumer Research Symposium, the Financial Management Association, the Princeton/Atlanta Fed Conference on Racial Justice and Finance, and the Northeastern University Finance Conference. Any remaining errors are our own.

[1] Jones School of Business at Rice University, alex.butler@rice.edu.
[2] Cox School of Business at Southern Methodist University, emayer@smu.edu.
[3] Jones School of Business at Rice University, westonj@rice.edu.
Declarations of conflicts of interest: none.

Racial Discrimination in the Auto Loan Market

Abstract

We provide evidence of discrimination in auto lending. Combining credit bureau records with borrower characteristics, we find that Black and Hispanic applicants' loan approval rates are 1.5 percentage points lower, even controlling for creditworthiness. In aggregate, discrimination crowds out 80,000 minority loans each year. Results are stronger where racial biases are more prevalent and banking competition is lower. Minority borrowers pay 70 basis point higher interest rates, but default *less* ceteris paribus, consistent with racial bias rather than statistical discrimination. A major anti-discrimination enforcement policy initiated in 2013, but halted in 2018, reduced discrimination in interest rates by nearly 60%.

Auto loans are the most widely used form of installment credit in the U.S. with over 100 million people borrowing as of 2017. Yet, compared to mortgages or student loans, the auto loan market is relatively unstructured, unregulated, and opaque. The lack of transparency makes it harder to monitor whether lenders consider characteristics like race and ethnicity. Indeed, suspicions of discrimination in this market led the Consumer Financial Protection Bureau (CFPB) to issue specific guidance to auto lenders in 2013 on how the Equal Credit Opportunity Act applies to auto loans.

Identifying discrimination requires information on applicant/borrower race and outcomes, but auto lenders are not required to report much data on either applications or loans.[1] Therefore, past studies of auto lending practices are scarce, largely suggestive, and incomplete. In this paper, we build an extensive, novel, and rich dataset to test for discrimination in this market.

Our empirical design links credit bureau records (a 1% nationally representative panel) to the Home Mortgage Disclosure Act (HMDA) data. These databases do not share a common identifier, but mortgages are reported in both places with sufficient granularity that we can uniquely match the majority of mortgages based on loan characteristics. The credit bureau records provide borrowers' financial characteristics and auto loans, while the HMDA data provide demographics. We use our matched panel of roughly 79,000 people per year from 2005 to 2017 to test whether minorities face discrimination in the auto loan market, and find strong evidence that they do.

[1] We use "race" to refer to both race and ethnicity. We limit our samples to people who are White, Black, or Hispanic, and classify people who are Black and/or Hispanic as minorities.

It is difficult to isolate discrimination rooted in preferences (Becker (1957)) or biased beliefs (Bordalo et al. (2016)), from alternatives such as omitted variables and statistical discrimination (Phelps (1972) and Arrow (1973)).[2] To do so, Becker (1957, 1993) proposes an "outcome test" that compares marginal loan profitability. Discrimination based on preferences or biased beliefs should lead to more profitable loans to marginal minority borrowers, because the bar is set higher. Researchers typically use loan performance as a proxy for profitability, and lower default rates for minorities are considered strong evidence of discrimination (Ferguson and Peters (1995)). We test for discrimination in loan approvals, interest rates, and subsequent defaults, and find consistent evidence in all three tests.

Our first tests study loan approval rates, controlling for a broad set of borrower and geographic characteristics (e.g., age, sex, income, ZIP code), and importantly, direct measures of *financial health* (credit score, debt/income, delinquencies, etc.).[3] Few other studies of discrimination have such a rich set of controls. We find that minority applicants have a 1.5 percentage point lower approval rate, comparable to a 26 point credit score reduction (32% of a standard deviation). The difference is 60% larger (2.4 percentage points) for minority applicants with subprime credit, where subjective preferences likely have greater influence. We find large racial differences even in college-educated, high

[2] Statistical discrimination occurs when lenders maximize profits by using race to proxy for aspects of creditworthiness that are unobservable. These attempts are limited by the usefulness of the proxy, particularly whether its use is based on accurate or inaccurate beliefs (Bohren et al. (2019)). We use "statistical discrimination" to refer to profit maximizing decisions based on accurate beliefs. In contrast, we use the term "discrimination" to refer to biased lending decisions, whether they are driven by preferences, or by inaccurate beliefs.

[3] These approval rates reflect both lender approval and borrower take-up. We discuss the nuances in detail below in Section 4.1.

income, and middle-aged applicant subsamples with substantial financial sophistication. Minorities fill out as many applications as White applicants, suggesting shopping effort does not explain the results. Moreover, our estimates likely understate the true magnitudes because our strongest results are in subprime applicants, but our sample consists of *homeowners,* who typically have better than average credit. A conservative back-of-the-envelope calculation suggests that each year more than 80,000 minorities fail to secure loans they would have received if they were White.

We next test whether our results are stronger in states where racial biases are more prevalent. Following Stephens-Davidowitz (2014), we measure states' racial bias using Google Search Volume that includes racial slurs. We find that the reduction in approval rates for minorities is three times larger (2.8 percentage points) in states in the top tercile of racial animus, compared to the remaining states (0.9 percentage points). We also test whether competition among lenders mitigates discrimination. Statistical discrimination should survive competitive pressure, whereas racial bias should be rooted out by competition (e.g., Becker (1957) and Berkovec et al. (1998)). Consistent with biased preferences or beliefs, we find stronger results in low-competition environments.[4]

Perhaps there are racial differences in applicants' overall creditworthiness that lenders observe, but we do not (and these differences correlate with racial biases and competition). If so, we might expect credit card lenders to identify this pattern. However, unlike auto loans, which typically involve personal interaction, most credit card decisions

[4] The distinction between discrimination rooted in biased preferences/beliefs and statistical discrimination provides insight into the economic forces at work. However, it is important to note that statistical discrimination is also illegal in the United States under the Equal Credit Opportunity Act.

are automated and provide less opportunity for discrimination (e.g., Gross and Souleles (2002), Moore (1996), and Tsosie (2016)). We find that the *same* minority applicants who faced reduced access to auto credit are *not* less successful with credit card lenders, *during the same year*. Moreover, the cross-sectional patterns consistent with discrimination in auto lending, are absent from credit card lending. These findings suggest that the human element of auto lending, rather than differences in creditworthiness, leads to the lower approval rates for minorities.

Discrimination also affects an intensive margin of credit provision through higher interest rates. Ceteris paribus, minorities pay 70 basis points more on their auto loans than White borrowers, which is comparable to the effect of a 37 point drop in credit score, and costs the average minority borrower in our sample $410 in present value terms. This result is particularly notable because these minority borrowers faced stricter approval rates for the loans. Moreover, the effect of minority status remains large in subsamples of sophisticated borrowers, and increases to 125 basis points in high racial bias states.

It is unlikely that omitted variable bias is driving our results, because we use an extensive set of controls, and any omitted variable would have difficulty explaining the cross-sectional patterns in discrimination and the results of our credit card falsification test. Nonetheless, such a bias would lead to higher ex post default rates for minorities. Instead, we find *lower* minority default rates in the full sample. For subprime borrowers, who are closer to the extensive margin of credit provision, minority default rates are a statistically significant 2.3 percentage points lower. This finding is consistent with loans to marginal minority borrowers being more profitable than loans to marginal White borrowers, a hallmark of discrimination.[5]

Our final tests evaluate whether regulatory oversight deters discrimination. We exploit a sharp increase in the CFPB's scrutiny of interest rates charged by indirect auto lenders starting in mid-2013. Our differences-in-differences tests show that the additional interest paid by minorities decreased from 84 to 35 basis points in the post-event period. Further tests show that the reduction in discrimination occurred primarily in areas where indirect auto lending is most prevalent, where the CFPB's efforts were focused. Moreover, despite lowering minorities' interest rates, CFPB oversight did *not* lead to a concomitant reduction in minorities' credit access, suggesting that the additional interest minorities are charged is discretionary, rather than necessary to make the loans economically viable.

Collectively, our findings are consistent with biased stereotypes or outright discrimination distorting outcomes for minorities who try to access the auto loan market. Our tests can also rule out some salient alternative explanations. Below, we discuss whether racial differences in financial sophistication, creditworthiness, or recovery rates in the case of default could explain our findings, and conclude that they cannot.

Our paper is related to prior work documenting racial disparities in the market for mortgages (e.g., Munnell et al. (1996)), peer-to-peer loans (Pope and Sydnor (2011)), and small business loans (e.g., Blanchflower, Levine, and Zimmerman (2003)). However, prior studies rarely include default tests, which makes inferences about discrimination difficult.

[5] Section 4.3 provides a detailed discussion of default tests. Based on prior literature and our own tests, we conclude that our default tests are likely conservative (tilted against our findings of discrimination). These tests help isolate discrimination from omitted variables or statistical discrimination, and provide some of the strongest evidence of lending discrimination to date.

For example, evidence from the mortgage market suggests that Black borrowers default more (e.g., Berkovec et al. (1994)), raising questions about whether racial disparities in approvals and interest rates reflect actual bias on the part of lenders. A distinguishing feature of our study is that we provide evidence of discrimination from all three settings—credit approvals, interest rates, and default rates—allowing us to better isolate discrimination from alternative explanations.

The primary contribution of our paper is to provide substantial evidence of discrimination in the U.S. auto loan market. Most prior work in this area focuses on discrimination by automobile salespeople in the form of quoting minority shoppers higher car prices (e.g., Ayers and Siegelman (1995)). Charles, Hurst, and Stephens (2008) document that Black borrowers pay higher rates on auto loans, but their tests do not condition on credit scores. Our study provides the first estimates of the effect of race on auto loan approval, robust estimates of the additional interest minorities pay, and importantly, tests for discrimination in this market based on ex post default rates. Each test provides evidence consistent with discrimination. Finally, our paper provides the first analysis of the impact of recent anti-discrimination enforcement policies in the U.S. auto loan market.

The paper proceeds as follows. Section 2 provides background information on auto lending. Section 3 describes our data and the process for matching credit bureau and HMDA records. Section 4 presents our empirical results. Section 5 discusses the plausibility of alternative explanations. Section 6 provides concluding remarks.

2. Background Information on Auto Lending

In this section we provide some general information about the U.S. auto loan market.[6] In 2017, 91% of U.S. Households had automobiles, and roughly 70% of auto purchases were used vehicles. Automobiles are a major household expenditure and the majority of purchases are financed (85% of new vehicles; 54% of used). Over 100 million U.S. consumers had auto debt as of 2017, with aggregate balances over $1.1 trillion.

Prime borrowers (credit score greater than 660) accounted for 58% of auto loan originations in 2017, with roughly half of these loans financing used cars.[7] Of the remaining 42% (subprime loans), roughly three quarters were for used cars. The average loan amount is around $30,000 for new and $20,000 for used cars. Average interest rates on auto loans range from around 4% for the most creditworthy borrowers, to around 16% for the least creditworthy borrowers.

To understand the structure of the auto lending industry, it is useful to classify lenders into three types: banks (commercial banks, thrifts, credit unions, etc.), auto finance companies, and "buy here pay here" lenders. While banks usually interact directly with consumers (direct lending), finance companies partner with car dealerships to originate loans and do not interact with the consumer (indirect lending).[8] Auto finance companies are either the "captive" financing arm of a major auto manufacturer (e.g., Ford Motor

[6] Unless otherwise specified, statistics in this section come from the National Household Travel Survey or an industry report: https://www.experian.com/assets/automotive/quarterly-webinars/2017-q4-safm.pdf.
[7] The credit score we use throughout the paper is the Vantage Score. The three major consumer credit bureaus developed Vantage Score to rival FICO scores, and it is the second most popular credit score. Vantage Score has the same score range as FICO, and is very similar, which led FICO to sue (unsuccessfully) the credit bureaus for producing such a similar product.
[8] Some banks also have indirect lending programs.

Credit Company) or an independent finance company. "Buy here pay here" lenders are used car dealerships that originate loans on-site, but these dealerships typically do not report their loans to the credit bureaus, and so are not in our analysis. Based on 2017 originations, the market shares are banks (53.3%), finance companies (40.3%, mostly from captives), and "buy here pay here" lenders (6.4%).

For auto loans financed indirectly, the typical scenario begins with the consumer choosing a car and completing an application for credit at the dealership. The dealer then submits this application to an indirect lender. The lender evaluates the application, which does not include the applicant's race, but does include name and address.[9] The lender decides on credit approval, and gives the dealer a minimum interest rate, not seen by the consumer. The dealer may then offer a loan with an interest rate at or higher than the minimum. The difference between these rates is called the "dealer markup." Indirect lenders and car dealerships have agreements specifying any limitations on the size of the dealer markup, and how the profits from the markup are to be shared. If the consumer accepts the credit offer, the dealer sells the loan to the indirect lender within a few days.

In the case of direct lending, it is clear that biases in loan officers' preferences or beliefs about creditworthiness could lead to minorities paying higher interest rates, or having more failed credit searches due to either explicit rejections or bad loan offers. At first glance, indirect lending may appear less prone to discrimination, because lenders do not interact directly with borrowers. However, regulators have actually expressed stronger

---

[9] We note that name and address can be used to proxy for race, and that there is ample evidence that employers screen out job applicants with minority-sounding names (e.g., Bertrand and Mullainathan (2004) and Agan and Starr (2018)). But, we cannot test directly for this specific form of discrimination in auto lending.

concerns about discrimination in indirect lending, where dealership finance officers ultimately play a role similar to a loan officer.

First, dealership finance officers may have the opportunity to advocate for borrowers when engaging with indirect lenders. Selectively advocating for certain types of borrowers could result in discrimination. For example, we include here an excerpt from Rice and Schwartz (2018), a small-sample paired audit study of indirect auto lending at dealerships in Virginia. In this case, the white tester has a credit score of 706, and the finance officer discusses advocating for them to receive the rate corresponding to a credit score of 720. We note that it would also be easy to see this type of advocacy extending to getting marginal borrowers near thresholds approved for loans.

> WHITE TESTER: Is there any way I can like get an idea of -- when in the process will I know what my interest rate is?
>
> FINANCE OFFICER: Once I send it over to the bank. I'm pretty solid you're going to be at 2.99%. I just got to call them and do a little bit of begging.
>
> WHITE TESTER: Begging?
>
> FINANCE OFFICER: Yeah. Get them to give me a bump to the 720, so it's going to depend on what credit unions I go through.

Second, even if we assume minorities receive equal treatment initially, there is a clear opportunity for discrimination when the finance officer sets the dealer markup. At this point, they have an incentive to charge a high markup because indirect lenders share the profits from markups with dealerships. This profit sharing typically incorporates loan performance, and hence, dealerships' incentives are similar to a loan officer's, with some additional incentive to consummate the transaction. With dealership finance officers acting as quasi-loan officers and exerting discretion, it is clear that biased preferences or beliefs about creditworthiness could lead to discrimination in indirect auto lending.

The anecdotal evidence from Rice and Schwartz (2018) is consistent with dealership finance officers exhibiting biases. Despite being more qualified on average, their minority testers were (1) less likely to be taken seriously as buyers and be given financing information, (2) offered more expensive loans, (3) less likely to have policies bent on their behalf, and (4) more likely to be treated disrespectfully. However, audit studies conducted at a few dealerships have clear limitations, and broader empirical studies like ours are necessary to test for market-wide discrimination (Heckman (1998)).

In March of 2013, the CFPB issued a Bulletin signaling its intent to hold indirect auto lenders accountable for discrimination in interest rates. The Bulletin made it clear that indirect lenders are subject to the Equal Credit Opportunity Act even if it is the dealership representative setting the final rate on the indirect lender's behalf. In December 2013, the CFPB issued its first major enforcement action against a large indirect auto lender for discriminatory interest rates. The CFPB ordered Ally Financial (formerly General Motors Acceptance Corporation) to pay over $90 Million in damages and penalties. The CFPB followed with additional enforcement actions.[10] In Section 4.4 we test the effect of the CFPB's increased scrutiny of auto lenders on our measures of discrimination.

[10] For a list of CFPB enforcement actions, see: https://www.consumerfinance.gov/policy-compliance/enforcement/actions/. An important issue in auto lending discrimination cases is that lenders do not keep records of borrowers' race, making it difficult for regulators conducting investigations to arrive at definitive conclusions. For example, in the analysis supporting its enforcement actions, the CFPB relies on borrower names and geographic locations to assign race through Bayesian Improved Surname Geocoding. Moreover, data availability typically constrains regulators to examining only the interest rates on originated loans at lenders under investigation. Our empirical design allows us to test for market-wide interest rate discrimination using linked administrative data on borrower race, and to examine credit access and defaults.

3. Data and Methodology

We merge two data sources to connect peoples' credit histories to their demographics. Mortgage lenders report applicants' race and ethnicity as well as other personal characteristics and loan application information to the Home Mortgage Disclosure Act (HMDA) database. The credit bureau data do not contain much demographic information, but they offer a lens through which to observe a broad set of borrower financial outcomes, and they track borrowers over time. Our approach is to match credit bureau records to HMDA through originated mortgages, which offer a sufficient set of identifying features in both datasets. Section 3.3 describes our matching process in detail. The match leads to the creation of a panel dataset with both demographics and financial outcomes, which we refer to as the Credit Bureau/HMDA Matched Panel.

Our analyses require several data sources in addition to the credit bureau and HMDA data (which we describe in detail below). First, we collect information on racial biases in U.S. states based on the Google Search Volume for racial slurs following the approach of Stephens-Davidowitz (2014). These data are collected from the Google Trends website. Second, we use information on bank branch locations and deposits from the Federal Deposit Insurance Corporation's Summary of Deposits to measure local banking competition. Third, we use the county-level share of non-bank auto lending from Benmelech, Meisenzahl, and Ramcharan (2017). These data sort each county into quartiles based on non-bank auto lending shares as of 2008Q1 using proprietary underlying data. Finally, we use characteristics of the borrowers' ZIP codes as control variables, and these data come from the Census Bureau's American Community Survey.

3.1 Credit Bureau Data

To construct the Credit Bureau/HMDA Matched Panel, we start with a panel dataset of credit bureau records, which is a 1% representative anonymized sample of all U.S. residents with a credit history and Social Security number. The sample is constructed using Social Security numbers ending in an arbitrarily chosen final two-digits. This produces a random sample because the Social Security Administration assigns the last 4 digits of Social Security numbers sequentially, regardless of location. The panel tracks individuals over time, and allows people to enter and exit the sample at the same rate as the target population, ensuring that the sample remains representative over time. This sampling method closely follows that of the Federal Reserve Bank of New York Consumer Credit Panel (see Lee and Van Der Klaauw (2010) for a detailed description of the sampling design and credit bureau data). The full data include annual observations for roughly 2.5 million people per year from 2004-2017, although our final sample of matched observations includes far fewer.

The credit bureau data provide a complete credit history for each individual, including the person's credit score, total debt, debt by category (mortgage, auto, credit card, etc.), past-due debt, new sources of credit opened, and "hard" credit inquiries. These credit inquiries occur when a borrower applies for credit, and the lender checks their credit report. The data also provide the person's age, ZIP code, and starting in 2010, their census tract.

3.2 HMDA Mortgage Application Data

The Home Mortgage Disclosure Act requires nearly all mortgage lenders to report detailed information on the applications they receive, and whether they originate the loan.

Only very small or exclusively rural lenders are exempt from HMDA reporting. Any depository institution must report to the HMDA database if it has at least one branch or office in a metropolitan statistical area (MSA), has at least $44 million in assets (2016 threshold), and originated at least one mortgage in the previous year. Non-depository institutions with assets over $10 million must report if their mortgage originations total at least $25 million (or represent 10% of their loans), and they receive at least five mortgage applications in MSAs. These requirements result in 95% of all first-lien mortgages being reported to the HMDA database (Avery et al. (2017)), and the coverage rate is likely higher for properties in MSAs.

The HMDA data include requested loan size, income, race, and ethnicity as well as the purpose of the loan (purchase, refinancing, improvement), any co-applicants, and the loan's priority (first or second lien). The census tract location of the property is also reported. If a loan is made, any loan sale is reported along with an indicator for sale to any quasi-government entity.

3.3 The Credit Bureau/HMDA Matched Panel

This section describes how we match the databases, analyzes the success rate of the match, and presents summary statistics on the borrowers in the resulting Credit Bureau/HMDA Matched Panel. The credit bureau data and the HMDA data are both anonymized, and there is no unique identifier to link the two datasets. However, the information on originated mortgages is reported at such a granular level in both datasets that the majority of mortgages can be uniquely identified based on a set of their characteristics.

We match mortgages in the credit bureau data to the HMDA data based on the following six characteristics: origination year, census tract location, loan amount, whether the loan is for purchase or refinancing, whether the mortgage is conventional or through the Federal Housing Administration (FHA) or Veterans Administration (VA), and if/to which quasi-government entity the loan is sold. We focus on mortgages originated from 2010-2016, because several matching variables are not available in the credit bureau data prior to 2010. We drop observations that are not uniquely identified. Because the HMDA data contain more than 95% of all originated mortgages, requiring the HMDA mortgage to be unique ensures that any matching mortgage in the credit bureau data identifies the same borrower with near-certainty. Fortunately, 89% of all originated mortgages in the HMDA data are uniquely identified based on the six matching variables.

After identifying unique loans in the HMDA data, we make several additional requirements to improve the quality of the match. We focus on home purchase and refinancing loans (home improvement loans are excluded because they are less well defined in both datasets). We require mortgages to be on owner-occupied homes, so that the property location will match the borrower's location in the credit bureau data. We also require the mortgage to be a first lien, and the property must be located within an MSA (where the HMDA data are most comprehensive). Finally, we require the mortgage to have only one applicant/borrower, so that the demographic data apply directly to the matched person in the credit bureau data.

We apply a similar set of filters to the mortgages from the credit bureau data. We require the mortgage to be the borrower's only first lien mortgage at the time. This filter ensures that the borrower's location in the credit bureau data will match the property location in the HMDA data. We also require that the person live in an MSA directly following the loan origination, and that they were the only applicant on the loan. After combining the filters imposed on the HMDA and credit bureau data, the target population for the matched sample is borrowers taking out a home purchase or refinance loan on their own (no co-applicant), for their primary residence, which is located within an MSA, from 2010-2016.[11]

There are two potential sources of error in our matching. First, a data error in one of the matching variables could create a mismatch, but we expect such errors to be rare because institutions systematically report these data to both the HMDA database and the credit bureaus. A second type of error could occur if a HMDA-reporting lender, and a non-reporting lender, originate identical mortgages that are otherwise unique. The reporting lender's loan could be matched to the credit bureau record of either of the two borrowers. This type of mismatch should be rare because HMDA covers nearly the universe of mortgages. Moreover, which credit bureau record the HMDA loan matches is random, because it will depend on which record is in the 1% random sample of credit bureau data. Therefore, this type of mismatch should not create any bias in our estimates outside of pure noise.

Table 1 presents summary statistics on the match. Panel A presents the match rate, which shows that we find a matching HMDA mortgage for 69% of the mortgages in the

[11] We have also run our tests including borrowers matched based on joint mortgages, and the results are similar. We focus on single applicants because we know the HMDA information applies directly to the borrower (as opposed to a co-applicant).

credit bureau data. Because HMDA includes 95% of mortgages and 89% are unique, the best we can do is roughly 84.5%. In other words, our algorithm found matches in roughly 81.7% (0.69/0.845) of the cases it could have. We view this as a reasonably good matched sample. We require exact matching (rather than nearest matches or propensity scores) in order to keep our matched dataset as accurate as possible.

The summary statistics in Panels B and C of Table 1 examine whether our matched sample is representative of the original population. Panel B shows that the sample of successfully matched home purchase mortgages is broadly representative of the starting population of credit bureau mortgages. One exception is that the matched sample has fewer borrowers with a prior mortgage. The statistics in Panel C show that the matched sample of refinance loans accurately represents the starting sample from the credit bureau data.

[Insert Table 1]

Next, we test whether race influences the match. None of the matching characteristics directly involve race. However, because we study the role of race in financial outcomes, it is important to test whether minorities are underrepresented in the data, and especially if a certain type of minority borrower (e.g., high/low income) is underrepresented. The regressions in Table 2 examine the likelihood that originated mortgages from the HMDA database are matched to our 1% sample of credit bureau records. The results show that borrower race is unrelated to the probability that we are able to match a loan. Furthermore, the interaction terms *Black X Log(Income)* and *Hispanic X Log(Income)* are insignificant, indicating no evidence of selection bias either directly or through the combination of race and income. In these tests, we can only focus on variables in HMDA like race and income. In our auto lending tests below, we can control for variables form both databases, which should mitigate any remaining concerns about selection bias.

[Insert Table 2]

We gather all the data for successfully matched White, Black, and Hispanic borrowers and refer to these data as the Credit Bureau/HMDA Matched Panel. Table 3 Panel A summarizes this dataset, which contains approximately 79,000 people per year, by providing a snapshot of the matched borrowers' characteristics in 2010, and comparing it to a 2010 snapshot of the full credit bureau dataset. Comparing Columns 1 and 2 of Panel A shows that the people in the matched dataset have higher credit scores, are younger, and are more likely to have a mortgage than the average U.S. resident with a credit history. These patterns are not surprising, because people have to either get a new mortgage, or refinance one between 2010 and 2016 to be in the matched panel. Columns 3-5 show that the White borrowers in the matched panel have higher credit scores and incomes than minority borrowers.

[Insert Table 3]

4. Empirical Results

4.1 Applicant Race and Access to Auto Credit

In this section, we test whether race affects access to auto credit. We start by selecting all person-years in the Credit Bureau/HMDA Matched Panel in which someone applies for an auto loan based on the "hard" credit inquiry that appears on their credit file when a lender checks their credit score. We then measure applicants' access to credit using the indicator variable *Credit Approval (Auto)*, which equals one when the person successfully opens a new auto loan during the year. Several recent papers that use credit bureau data construct and validate similar measures of credit access (e.g., Bhutta and Keys (2016), Akey et al. (2018), Akey, Heimer, and Lewellen (2018), Brown, Cookson, and Heimer (2019), and Mayer (2021)). We note that only those who apply for an auto loan will be in our sample. If minorities anticipate lending discrimination, they may be less willing to apply for a loan. If such selection impacts marginally qualified candidates more than better credit quality candidates, then our results may be understated.

The main limitation of this measure of credit access is that we do not observe individual applications being approved/rejected. Instead, we observe whether a person's credit shopping attempt was successful. This issue is unavoidable given that auto lenders do not keep records of applicants' race, and do not report applications to a regulatory authority (ours is the first study to work around this issue by using credit bureau data and information on race). Potential concerns with this measure could arise if, rather than being denied credit, minorities have more failed credit shopping attempts because they have weaker demand for loans (leading them to turn down more offers), or they exert less effort shopping. Fortunately, a strength of the credit bureau data is that it includes peoples' applications across all lenders, so we can observe the total number of auto loan applications for a given person-year. If minorities in fact had weaker demand for auto loans, or exerted less effort shopping, we would expect their failed credit searches to include fewer applications. We find the opposite; Table A.1 shows that minorities' failed credit searches include slightly more applications than White borrowers' failed searches.

Importantly, our measure of credit access is conservative in that if one lender discriminates against a minority applicant, but another lender steps in and makes the loan, we count the episode as a successful credit shopping attempt. This approach puts our empirical tests in a position to test for market discrimination (e.g., Heckman (1998)), rather than discrimination at individual lenders. This is an important distinction, and one that many prior lending discrimination studies ignore (e.g., studies using HMDA applications which are not linked at the person level). Moreover, the approach allows us to estimate the total number of minorities who fail to secure loans each year due to discrimination.

Panel B of Table 3 summarizes the characteristics of the auto loan applicants from 2005-2017. Although borrowers are matched to HMDA based on mortgages originated from 2010-2016, we can observe their auto loan applications in prior years as well. The sample starts in 2005 because we need one prior year to construct lagged controls. Column 1 describes all auto loan applicants in the credit bureau data. Column 2 describes applicants in the matched panel. The applicants in the matched panel have higher credit approval rates and credit scores than the average applicant. Columns 3-5 show that White applicants have higher credit approval rates, credit scores, and incomes than minority applicants.

We test whether race affects access to auto credit by regressing *Credit Approval (Auto)* on *Minority*, an indicator for the person being Black or Hispanic. We control for individual and ZIP code characteristics, as well as state-by-year fixed effects, and indicators for the timing relative to the borrower's credit bureau/HMDA match. Table 4 presents these regression results. We find that minority applicants are 1.5 percentage points less likely to obtain credit than White applicants (Column 2). This unexplained difference in approval rates is roughly the same size as we would see from a 26 point (32% of a standard deviation) reduction in applicant credit score, and a back-of-the-envelope calibration suggests that each year it results in more than 80,000 minority applicants failing to secure loans they would have received if they were White (see Appendix B for the details of the calculation).

Although these estimates are economically large, the difference between Columns 1 and 2 shows the importance of including accurate measures of credit quality. Point estimates on the *Minority* coefficient are three times larger (4.5 percentage points) without these controls. This finding is pertinent for the broader literature on lending discrimination because studies often lack detailed measures of borrower credit quality. For example, the HMDA data do not include credit scores, and attempts to supplement these data have led to small samples, and controversial results (e.g., Munnell et al. (1996) and Day and Liebowitz (1998)).

The results in Column 2 also show slightly higher approval rates for women, but we caution against interpreting these findings as evidence that women receive preferential treatment in auto lending. We suspect this result is due to a sample selection bias, in which women who take out mortgages without a co-applicant (and therefore make it into our sample) are particularly creditworthy. Consistent with this explanation, we find that credit card lenders also extend slightly more credit to the women in our sample. Furthermore, after controlling for the interest rates that auto lenders set to price credit risk, we find similar default rates for men and women—suggesting no differences in loan profitability.

Columns 3-5 of Table 4 show that the reduction in credit approval is insignificantly different for Black versus Hispanic applicants, and that minorities face a larger reduction in approval rates when they are subprime borrowers. This second result is noteworthy because approval for subprime borrowers typically involves more loan officer discretion, lowering the marginal cost of discriminatory decision making. In subsample tests reported in Table A.2, we find large reductions in credit approval for minorities in samples of college-educated, high income, and middle-aged applicants. The fact that we find these results even in subsamples of more sophisticated and experienced borrowers suggests that our results are not driven by racial differences in financial sophistication.

[Insert Table 4]

We conduct two more robustness checks. First, we run tests to mitigate any potential concerns that our credit approval results may be affected by the shortfall in funding experienced by non-bank lenders during the 2008 financial crisis. We repeat the tests from Table 4 on a post-crisis sample (2011-2017) and find similar results (see Table A.3). We also check that our results are robust to using alternate approaches to control for geographic effects. Table A.4 repeats the tests from Table 4 using ZIP code fixed effects and finds similar results.[12]

In our next set of tests, we use the cross-sectional variation in our data to identify where lending discrimination is most prevalent. First, we test whether discrimination is stronger in high racial bias states (similar to studies of labor markets like Charles and

[12] There are over 30,000 ZIP codes in the U.S., and tests employing ZIP code fixed effects only exploit variation within ZIP codes with multiple observations.

Guryan (2008) and work by Dougal et al. (2019) on higher education bond markets). To quantify racial biases, we replicate the approach of Stephens-Davidowitz (2014) and use Google Search Volume for racial slurs. We tabulate our calculation of this state-level measure of racial animus (*Racial Slur GSV*), updated to reflect our 2005-2017 sample period, in Table A.5.[13] In Column 2 of Table 5, we find that the effect of minority status on credit approval is over three times larger (2.8 percentage points) in states in the top tercile of racial animus, compared to the remaining states (0.9 percentage points). We find similar cross-sectional patterns using a continuous version of *Racial Slur GSV*, or using the racial bias index from Levine, Levkov, and Rubinstein (2014) based on interracial marriage rates (see Table A.6).

[Insert Table 5]

We also estimate the reduction in approval rates minorities face in each state. The state level estimates come from a regression similar to those in Table 4 and Table 5, except that the *Minority* indicator is interacted with indicators for each state. In order to consider the $State_i$ *X Minority* coefficient a valid estimate of lending discrimination in the state, we require that our sample contains at least 25 minority applications in the state (excludes 6 states with small minority populations). Figure 1 graphically presents the relation between *Racial Slur GSV* and our state-specific estimates of the reduction in loan approval rates for minorities (also tabulated in Table A.5). The size of the circle plotted for each state is proportional to the number of minority applications in the state, and each state is weighted

[13] See Stephens-Davidowitz (2014) for the search criteria used to construct this measure of racial animus. Google computes search volumes based on a fraction of all Google searches. We collect 50 draws of the data and assign each state its average search volume (we find very little variation across draws).

by the number of minority applications when computing the best fit line in the plot and the correlation between the $State_i$ *X Minority* coefficient and the *Racial Slur GSV*, which is -0.49 (p-value = 0.001). The map in Figure 2 categorizes states based on whether we find a statistically significant reduction in approval rates for minorities, and shows that the strongest evidence of discrimination is in the Deep South, the Ohio River Valley, and parts of the Southwest.

[Insert Figure 1 and Figure 2]

In our second cross-sectional analysis, we test whether the effect of race is stronger for applicants living in counties with low levels of banking competition (top tercile of local bank deposits HHI). We find that minorities face a larger reduction in approval rates in low competition environments (Table 5, Column 3). This result is consistent with the gap in approval rates stemming from costly racial biases in preferences/beliefs, which competition should root out (e.g., Becker (1957), Berkovec et al. (1998), and Buchak and Jørring (2017)), as opposed to profitable statistical discrimination.

Our third cross-sectional test is based on the prevalence of non-bank auto lending in the county where the applicant lives. We use data from Benmelech, Meisenzahl, and Ramcharan (2017) who sort counties into quartiles based on the share of non-bank auto lending in 2008 using proprietary data. Nearly 45% of the applicants in our sample live in counties in the top quartile of non-bank auto lending share, because these counties tend to be urban. In Column 4 of Table 5, we find that the effect of race on credit approval is insignificantly different for applicants in counties in the top quartile of non-bank auto lending share, compared to the remaining counties. This finding suggests that the racial disparities in credit approval that we document are not driven by a particular type of lending institution.

In our last cross-sectional test, we consider whether differences in population density contribute to the cross-sectional relationships we document. For example, rural areas may be associated with racial bias and/or low banking competition. However, in Column 5 of Table 5, we find no evidence that the effect of race differs in the less densely populated parts of our sample (bottom tercile based on ZIP code population density).

At this point, we use credit card lending data from the credit bureau to conduct a falsification test. Credit card lenders generally use quantitative algorithms to decide whether to extend credit and how much. This automation reduces the opportunity for discrimination based on biased preferences or beliefs. We examine credit card lenders' willingness to lend to the auto loan applicants in our sample, *during the same borrower-years*. If the lower minority approval rates on auto loans are justified by information available to lenders, but not to us as econometricians, then credit card lenders should also be less willing to lend to these minority applicants.[14] On the other hand, if there is no racial difference in credit card lending, it would suggest that the disparities in auto lending stem from the human component of the lending process.

Table 6 presents our falsification tests. We select the subset of auto loan applicants who applied for credit cards or credit card limit increases that same year.[15] We regress the

[14] If credit card lenders engage in redlining we might expect lower credit card approval rates for minorities, making our falsification test more conservative. See also Cohen-Cole (2011) and Brevoort (2011).

[15] We focus on credit card lending for our falsification tests for three reasons. First, borrowers frequently request this type of credit, generating overlap with our auto loan application sample. Second, we know that these requests are handled remotely by quantitative algorithms. Finally, our credit bureau data include a

dollar increase in the borrower's total credit card limit across all cards (average increase is $3,090) on race, interaction terms, and controls, similar to the tests in Table 5. The results in Column 1 show that credit card lenders extend just as much new credit to minorities as to Whites. Columns 2 and 3 show that the cross-sectional patterns that we observe in auto lending that suggest discrimination are absent from credit card lending. Although these two credit products are different, these credit card results are inconsistent with the idea that the racial disparities in auto lending and their cross-sectional patterns are driven by racial differences in applicants' overall creditworthiness. Put differently, any hypothetical omitted variable that might spuriously create the appearance of discrimination in our auto lending tests must not concern credit card lenders, and yet must generate the cross-sectional patterns we observe.

In sum, we find that minority applicants have significantly lower approval rates when they search for auto loans, even conditional on a rich set of covariates. This result holds in numerous subsamples, and is stronger in areas that exhibit more racial animus, and where the banking market is less competitive. These cross-sectional patterns support an explanation grounded in lenders' racial biases rather than statistical discrimination. Moreover, when minorities' creditworthiness is evaluated by quantitative credit card lending algorithms, these patterns disappear. This set of findings offers strong evidence of discrimination in auto lending.

4.2 Borrower Race and Auto Loan Interest Rates

distinct category for credit card inquiries (as they do for mortgages and auto loans). In contrast, inquiries for other types of credit, such as personal installment loans, are aggregated up into broader categories which likely mix in-person and remote lending, and still remain sparse compared to credit card inquiries.

In this section we test whether auto lenders charge minorities higher interest rates than comparable White borrowers. We construct an auto loan level dataset from the Credit Bureau/HMDA Matched Panel. This dataset contains an observation for each new auto loan to borrowers in the matched panel. We require the loan to be the borrower's only auto loan at origination. We also require information on the borrower's scheduled monthly auto payment and assume a fixed rate loan in order to compute the interest rate. This information is not available from the credit bureau until 2011, and it is missing for 29% of the loans from 2011 going forward. In un-tabulated results, we find no evidence that the loans with missing data are different in terms of loan or borrower characteristics. Our final sample has 25,697 auto loans originated between 2011 and 2017, with 4,874 of these made to minority borrowers.

Table 7 presents summary statistics describing our sample of auto loans. Columns 1-5 show the variable means and standard deviations (in brackets) for all borrowers, White borrowers, minority borrowers, subprime borrowers, and prime borrowers, respectively. White borrowers have higher credit scores and incomes, and pay lower interest rates than minority borrowers on average.

[Insert Table 7 Here]

We test whether minority auto borrowers pay higher interest rates by regressing each auto loan's annual rate on *Minority*, controlling for personal, loan, and ZIP code characteristics, as well as state-by-year fixed effects, and indicators for the timing relative to the borrower's credit bureau/HMDA match and for the calendar month of origination. The results in Column 2 of Table 8 show that minority borrowers pay interest rates 70 basis points higher than can be explained by observable characteristics. This magnitude is comparable to what we would expect from a 37 point decrease in borrower credit score, and is larger than studies have typically found in other consumer credit markets—e.g., Bartlett et al. (2019) find that minorities pay rates 8 basis points higher in the mortgage market.

Several other studies also examine racial disparities in auto loan pricing. Cohen (2012) reports statistics demonstrating that a higher percentage of Black borrowers' loans included dealer markups (and their markups were larger) at several indirect auto lenders targeted in class action lawsuits in the late 1990s and early 2000s.[16] Lanning (2021) finds similar evidence of discriminatory markups using data collected from several indirect lenders during CFPB investigations.[17] Closer to our work on interest rates in the broader market, Charles, Hurst, and Stephens (2008) use data on auto loans to 2,725 White borrowers and 320 Black borrowers from the 1992-2001 waves of the Survey of Consumer Finances (SCF). The authors estimate quantile regressions, and find that race matters primarily at the 75th percentile of the interest rate distribution, where Black borrowers pay 134 basis point higher rates. The authors control for several self-reported measures of a borrower's credit history, but the SCF data do not contain credit scores. In our data, we estimate the additional interest paid by Black borrowers at the 75th percentile of rates to be 100 basis points using our full set of controls, and 139 basis points if we exclude only *Credit Score* (see Table A.7). These findings suggest that even analyses that control for a

[16] These confidential data were accessed as a plaintiff's expert and cannot be used for research purposes.
[17] These data do not contain direct information on borrower race/ethnicity, but the CFPB uses names and geographic locations to assign race through Bayesian Improved Surname Geocoding.

set of credit history variables, but not credit scores, likely significantly overstate the effect of race.[18] Omitting credit history variables altogether (even controlling for age, sex, income, loan characteristics, etc.) leads to estimates that overstate the effect of race by a factor of 2 or more—compare Columns 1 and 2 of Table 8 or see Table A.7.

Next, we take advantage of the rich cross-sectional variation in our data, and test where race has the largest impact on interest rates. The results in Column 3 of Table 8 show that the effect of race is much larger in high racial bias states (top tercile of *Racial Slur GSV*). In these states, minorities pay interest rates 125 basis points (26% of a standard deviation) higher than can be explained by observable characteristics. Perhaps this pattern could partially reflect statistical discrimination, if local racial biases are correlated with unobservable racial differences in creditworthiness. However, we can test this alternative explanation fairly directly by examining ex post default rates, and we find no evidence supporting this alternative. Specifically, we find that while local racial biases predict higher interest rates for minorities, they do not predict higher default rates, suggesting bias rather than statistical discrimination (see Table A.8).

In additional cross-sectional tests, race also appears to have a larger effect on interest rates in areas with low banking competition, although this point estimate is statistically insignificant (Table 8, Column 4). The results in Column 5 show no significant difference in the effect of race on interest rates based on the share of non-bank lending

[18] Direct comparisons of our results to those in Charles, Hurst, and Stephens (2008) should be made with caution in light of the different time periods and imperfect overlap in controls—although we do find similar estimates to theirs when we omit *Credit Score* from the controls. The change in our own estimates when we include/exclude *Credit Score* provides more robust (albeit similar) evidence of its importance. We note that numerous other lending discrimination studies argue that the various (often self-reported) credit history indicators they use should proxy well for the credit scores that lenders actually use.

where the borrower lives, however, the analysis in Section 4.4 will shed light on the CFPB's role in this matter. The test in Column 6 suggests that population density does not play a role in the cross-sectional patterns we find.

[Insert Table 8]

At this point, we consider whether the type of car being purchased (e.g., new versus used), and hence the representative institutions involved in the sale and financing of that type of car, affect the levels of discrimination we find. This analysis is motivated by the fact that automobile dealerships range from large new car dealerships affiliated with manufacturers, to small independent used car dealers. Moreover, in indirect auto lending, employees at car dealerships often help set the interest rate via dealer markup. Admittedly, we cannot directly observe the type of car being purchased, or the institutions involved. However, we do observe the loan size, which (especially in the extremes) is likely a good indicator for whether the car is new versus used. We find the most discrimination for the smallest auto loans (likely used cars). This pattern holds for both prime and subprime borrowers. However, even minorities with prime credit scores buying expensive cars that are almost certainly new, pay rates 18 basis points higher than comparable White borrowers (see Table A.9 for these results).

Additional subsample tests in Table A.10 show that racial differences in financial sophistication are not driving our results—the effect of minority status on interest rates is just as large in the college-educated subsample as the full sample, and also remains large in high income and middle-aged subsamples. In sum, the results in this section show that minorities face discrimination not only at the extensive margin of credit provision (loan approval), but also at an intensive margin (loan pricing).

4.3 Race and Auto Loan Default Rates

4.3.1 Conceptual Framework for Default Tests: Infra-marginality

Becker (1957, 1993) argues that "outcome tests" identify discrimination. In the lending context, this means testing whether loans to marginal minority borrowers are more profitable than those to marginal White borrowers. This test evaluates whether lenders or loan officers set the bar higher when screening minority applicants, due to racial biases—neither omitted variables nor statistical discrimination generate racial differences in marginal loan profitability.

Empirical implementations of the outcome test face at least two challenges. First, researchers do not observe loan profitability directly, but rather, evaluate default rates conditional on loan/borrower characteristics, with fewer defaults for minorities signaling higher profitability (discrimination). Focusing on defaults biases analyses against finding discrimination, because it understates the relative profitability of minority loans by ignoring prepayment risk, which is higher for White borrowers (e.g., Deng and Gabriel (2006)). The second challenge is that researchers do not observe the literal marginal borrower, but rather, examine average default rates conditional on covariates. This is the "infra-marginality problem" (Ayres (2002)).

The infra-marginality problem affects virtually all prior studies of lending discrimination that use default tests (and most research using outcome tests to study discrimination in other contexts).[19] Our default tests condition on a rich set of covariates and focus on the subsample of borrowers nearest the extensive margin of credit provision, allowing us to compare racial differences in average default rates *near the margin*, but not explicitly at the margin. Therefore, we need to consider how the infra-marginality problem is likely to affect our results.

Prior studies point out that the infra-marginality problem biases default tests *against* finding discrimination (e.g., Galster (1993), Brueckner (1996), Ross (1996a), Ross (1996b), Yinger (1996), and Ross and Yinger (1999)). Systemic racial differences in economic outcomes can remain even within the pool of borrowers who make it past an unbiased screening process. The pool of approved minorities may still be less creditworthy than the pool of approved whites. Therefore, even if lenders discriminate by setting a higher threshold for minorities, we still may not find lower *average* default rates for minorities because the entire distribution of creditworthiness is lower.[20]

Focusing on subsamples near the margin and conditioning on covariates may improve the ability of default tests to detect discrimination. Yet, if minorities are less creditworthy on an unobservable dimension, which could bias credit approval and interest rate tests towards finding discrimination, then we should expect default tests to remain biased towards finding minorities default more, and thus against finding discrimination. Ferguson and Peters (1995) formalize this point by showing that concurrent findings of

---

[19] A notable exception is Dobbie et al. (2019), who study discrimination at a high-cost lender in the U.K. The authors use quasi-random variation in the assignment of loan officers to applicants, combined with loan officers' overall strictness, to instrument for loan origination and estimate outcomes at the margin.

[20] For illustrations of this point, see e.g., Ferguson and Peters (1995) and Ross and Yinger (1999). We also show that the entire minority creditworthiness distribution is shifted to the left in our sample (see Figure A.1).

lower average approval rates and equal or lower *average* default rates for minorities constitute evidence of discrimination.[21]

We expect our loan default tests to be conservative in that any finding of lower default rates for minorities should be interpreted as strong evidence of discrimination. Evidence of racial differences in default rates in other markets is mixed. At a high-cost lender in the United Kingdom, Dobbie et al. (2019) find weak evidence of discrimination in default tests, but strong evidence using loan profitability tests. Default rate tests do not provide evidence of discrimination in the mortgage market (e.g., Berkovec et al. (1994, 1998)), even though there is strong evidence of discrimination based on mortgage approvals, interest rates, and audit studies (e.g., Munnell et al. (1996), Bayer, Ferreira, and Ross (2018), Ross et al. (2008), and Hanson et al. (2016)). Hence, finding strong results across all dimensions, including ex post default rates, would seem to provide some of the most compelling evidence in the literature.

4.3.2 Default Test Results

In this section, we implement a version of the outcome test for discrimination (Becker (1957, 1993)), by testing for racial differences in loan performance conditional on loan and borrower characteristics. For these default tests, we need to make two additional requirements to include auto loans in the sample. First, we end the sample with loans originated in 2015, so that we can track the performance of loans for at least two years. Specifically, we examine the loan's status as of December 31 in the year of origination and

[21] Shaffer (1996) extends Ferguson and Peters (1995) to show that an exception can occur if there are significant racial differences in recovery rates conditional upon default. We consider the possibility of differences in recovery rates in Section 5 below.

the following two calendar years.[22] We mark the loan as a default if the borrower is 90 or more days delinquent at any of these three points, or if the automobile has been repossessed during this time. Second, we require auto loans to be originated after their borrower's match to HMDA, i.e., after their mortgage or refinance loan, so that our sample of auto loans is not affected by any forward-looking bias. Without this filter, a forward-looking bias could arise because a recent auto loan default would hurt a borrower's mortgage application, and thus their chances of making it into our matched sample. Requiring auto loans to be originated after the match to HMDA eliminates this concern.

In the tests presented in Table 9, we regress our indicator for default on *Minority*, and controls for personal, loan, and ZIP code characteristics, as well as state-by-year fixed effects, and indicators for the number of years since the borrower's credit bureau/HMDA match and for the calendar month of origination. The results in Column 1 show that in the full sample, minority status has a negative effect on the probability of default, but the point estimate is statistically insignificant. However, a default rate test on the full sample includes many borrowers who are far from the extensive margin of credit provision, weakening the test's ability to detect differences at the margin. Therefore, we run the test on the sample of subprime borrowers, who are close to the margin, and are a more appropriate sample for the outcome test.

In the subprime sample (Column 2), minority status has a negative and statistically significant effect (2.3 percentage points) on the probability of default. This magnitude is

[22] The credit bureau data only allow us to see detailed information on delinquency status as of December 31 each year in our sample. However, we include indicators for calendar month of origination, in order to control for any differences in default rates based on where these December 31 points fall in the life of the loan.

comparable to the effect of a 39 point increase in a borrower's credit score. Column 3 shows that the effect of minority status is insignificant in the sample of prime borrowers. These default rate tests show that loans to minorities near the margin are more profitable than loans to White borrowers near the margin, providing strong evidence that minorities face discrimination in the auto loan market.

[Insert Table 9]

4.4 CFPB Oversight and Auto Lending Discrimination

In this section we test whether more intense regulatory oversight reduces discrimination in the auto loan market. In March 2013 the CFPB conspicuously identified in a Bulletin that it intended to hold indirect auto lenders accountable for discrimination. The CFPB solidified its stance in December 2013, when it issued its first major enforcement action against a large indirect auto lender for discriminatory lending practices, ordering Ally Financial to pay $98 million in damages and penalties.

In our first set of tests, we use a differences-in-differences approach to assess whether the increase in regulatory scrutiny caused a reduction in discrimination. Specifically, we test whether racial disparities in interest rates and credit approval changed after 2013. We use the same samples as our prior tests, and treat 2011-2013 as the pre-intervention period, and 2014-2017 as the post-intervention period.

The differences-in-differences tests for interest rates and credit approval are shown in Columns 1 and 4 of Table 10, respectively. The results in Column 1 show that the additional interest paid by minorities decreased from 84 basis points in the pre period to 35 basis points in the post period—a 58% decrease. This large decline is statistically significant at the 1% level. The results in Column 4 show that the reduction in credit approval that minorities faced declined from 1.8 percentage points to 1.2 percentage points, although this change was statistically insignificant. It may not be surprising that the pressure from the CFPB had less of an impact on approval rates, given that the Bulletin and the Consent Order against Ally Financial focus primarily on interest rates. Yet, importantly, these credit approval results show that pressure to avoid charging minorities disproportionately high dealer markups/rates did not reduce minorities' access to credit, as it would have if the higher rates were necessary to make the loans economically viable.[23] Overall, these tests suggest that the additional interest minorities are charged is largely discretionary, and that the CFPB's actions helped mitigate discrimination.

[Insert Table 10]

Next, we exploit the fact that the CFPB scrutiny fell on indirect auto lenders, e.g., non-bank lenders like manufacturers' financing arms. We use a triple differences approach to test whether the change in discrimination was larger where non-bank auto lending is most prevalent. Column 2 of Table 10 presents our results. Interest rate discrimination dropped significantly more in counties with the most non-bank lending, where lenders faced more scrutiny. In fact, the reduction in discrimination in these areas appears to be driving the overall effect in our differences-in-differences test, as the reduction in the remaining areas is statistically insignificant. The actions taken by the CFPB appear to have reduced discrimination, as opposed to a downward trend in discrimination over time.

[23] We note that it is unlikely that lenders would originate value-destroying loans to minorities in order to avoid CFPB scrutiny, because the CFPB cannot monitor lenders' approval decisions based on race, due to data constraints.

In Column 5 of Table 10, we conduct a similar triple differences test using credit approval as the outcome variable. The results show no significant difference between the trends in discrimination in high versus low non-bank financing areas. This result is not surprising considering the CFPB's focus was on interest rate discrimination. In Columns 3 and 6 of Table 10, we test whether discrimination is decreasing at a different rate in high versus low racial bias states, and find no such evidence.

Figure 3 shows estimates of the additional interest paid by minorities on auto loans each year from 2011-2017. The point estimates come from a regression of interest rates on the full set of controls, where the *Minority* indicator is interacted with indicators for each year. Panel A shows these estimates for the full sample, and the results show that there is no major time trend in the additional interest paid by minorities in the period preceding the CFPB's actions. However, there is a large drop in the additional interest paid by minorities from 2013 to 2014—precisely the time of the CFPB's actions. Panels B and C show the estimates for minorities living in areas with a high versus low share of non-bank auto lending, respectively. The results show a large drop in the additional interest paid by minorities in the high non-bank lending areas that were most affected by the CFPB's actions, and almost no drop in the less-affected areas. These results provide strong evidence that the CFPB's actions led to a reduction in discrimination by non-bank auto lenders.

[Insert Figure 3]

In addition to supporting our previous findings of discrimination, the results in this section are important from a policy perspective. Our paper is the first to analyze the impact of the CFPB's controversial effort to prevent discrimination in auto lending. Our results show that the increase in regulatory scrutiny reduced interest rate discrimination without restricting minorities' access to credit.

5. Discussion

In this section, we provide a framework for evaluating whether a given economic force could explain our findings. We begin by outlining the established empirical results. Then, we demonstrate our framework by considering whether three salient alternatives—racial differences in financial sophistication, creditworthiness, or recovery rates conditional upon default—could explain the patterns we find in the data.

5.1 Empirical Findings

Here, we state five key empirical findings of our study. First, we find that minorities face lower approval rates and higher interest rates on auto loans, even after controlling for credit score, income, loan characteristics, and a broad set of additional covariates (Tables 4 and 8). Additional tests confirm that these results hold in college-educated, high-income, and middle-age subsamples with substantial financial sophistication, as well as for loans of various sizes (Tables A.2, A.9, A.10). We also establish that minorities who tried and failed to obtain credit shopped just as hard as their White counterparts (Table A.1).

Second, we find that the racial disparities in approvals and interest rates correlate positively with local racial biases and negatively with lending competition (Tables 5 and 8). Tests also show that the additional interest minorities pay in high versus low racial bias areas is not grounded in differing minority default risk across these areas, further isolating bias as the mechanism generating rate disparities (Table A.8).

Third, we find that credit card lenders evaluating these same borrowers with algorithms (during the same years that they applied for auto credit) do not deem minorities to be less creditworthy, and the cross-sectional patterns with respect to local racial biases and lending competition are absent (Table 6).

Fourth, we find that, ceteris paribus, minorities are less likely to default on auto loans, consistent with minority loans being more profitable. This result is strongest among subprime borrowers near the extensive margin of credit provision, consistent with racial differences in marginal loan profitability, a hallmark of discrimination (Table 9).

Fifth, we find that an increase in CFPB scrutiny of interest rates led to a 60% reduction in the additional interest that minorities were charged, but did *not* affect their access to credit, suggesting that the additional interest minorities are charged is discretionary, rather than a result of competitive loan pricing necessary to make the loans viable (Table 10).

5.2 Framework for Interpretation

First, we note that racial discrimination based on biased beliefs, stereotypes, or prejudices held by loan officers and dealership finance officers would explain all five of the findings above. No alternative hypothesis we have identified can explain these five findings jointly. The breadth of our tests is an important strength relative to other lending discrimination studies, and it underscores the likelihood that discrimination is an important factor in these results. Yet, we need to consider alternative explanations. We propose that alternatives should (a) explain a significant portion of the five key findings, and (b) not be strongly contradicted by any of these findings. We apply this framework to three salient alternative explanations below.

5.3 Evaluating Alternative Explanations

5.3.1 Racial Differences in Financial Sophistication

Perhaps minority borrowers are on average less financially sophisticated than their White counterparts, and thus exert less effort shopping for credit. Profit maximizing lenders, in turn, might attempt to extract rents from all naïve borrowers (minority and White), resulting in minorities paying higher interest rates and having more failed credit shopping attempts.

Important aspects of our first three findings are inconsistent with this explanation. First, our study is based on a sample of homeowners, and we control for age, income, and credit quality, all of which should reduce any differences in financial sophistication. Second, even in subsamples of the most sophisticated borrowers—people with college educations, high incomes, or experience (middle-aged borrowers)—we find that minorities face reductions in credit access and pay higher interest rates than comparable White borrowers. Third, to explain our results, unobserved racial differences in financial sophistication would have to not only vary systematically with both local racial bias and banking competition, but also would have to affect borrowers' ability to obtain auto credit, but not credit cards. Finally, when we examine failed auto credit shopping attempts, we find that minorities have more applications on average, inconsistent with minorities shopping less. Therefore, we conclude that racial differences in financial sophistication are not a viable explanation for our findings.

5.3.2 Racial Differences in Creditworthiness

Perhaps minority status is correlated with factors that lower creditworthiness, but are difficult for econometricians to observe, such as low job security. If lenders somehow observe these variables directly or use borrowers' race to proxy for the underlying factors, we may attribute to racial bias what is in fact omitted variable bias or statistical discrimination.

Important aspects of our findings run contrary to this explanation. First, the racial differences in creditworthiness would have to exist despite our extensive controls from borrowers' credit histories and within high education and high income subsamples, and would also have to vary systematically with both local racial biases and local banking competition. This high bar for an omitted variable is raised further by the fact that the omitted creditworthiness variable must be pertinent for (secured) auto lending but not for (unsecured) credit card lending. Second, although local racial bias predicts higher interest rates for minorities, it does not predict higher minority default risk (which we would expect it to if rate disparities resulted from omitted variable bias or statistical discrimination). Third, the documented effects of increased CFPB oversight suggest that the additional interest minorities are charged is discretionary, rather than grounded in difficult-to-observe fundamentals. Finally, and most importantly, minorities default less in our tests, contrary to the notion that unobserved factors make minorities worse credit risks.

5.3.3 Racial Differences in Recovery Rates Conditional upon Default

Perhaps lenders' recovery rates conditional on default are lower for minority borrowers than for White borrowers. Recovery rates could be lower if minorities have higher loan-to-value ratios, purchase cars that depreciate faster, or use their car more heavily (lowering resale value). A higher cost of minority defaults might justify fewer approvals and higher interest rates ex ante.

Three aspects of our findings cast doubt on this explanation. First, our tests control for borrower income, credit quality, loan amount, and state of residence, which should capture much of the variation in recovery rates. Second, it seems unlikely that racial differences in recovery rates could explain the cross-sectional geographic variation in our results. Third, if the additional interest paid by minorities were necessary to compensate lenders for lower recovery rates, then the CFPB's actions that reduced this additional interest should have led to lower approval rates for minorities (they did not).

Any remaining concerns about racial differences in recovery rates are difficult to rule out directly because loan-level recovery rate data are not available. However, we can calibrate the potential magnitude of these concerns. Average recovery rates in the case of default are roughly 50% for auto loans (around 58% for prime borrowers and 42% for subprime borrowers).[24] In back-of-the-envelope calculations, we find that even if we make the extreme assumption that the difference in average recovery rates for White and minority borrowers—conditional on our controls—is as large as the raw difference between prime and subprime borrowers (58% versus 42%), lenders could recoup these losses by charging minorities an additional 12 basis points in interest (compared to the additional 70 basis points they actually charge minorities). In fact, even if all White borrowers' defaults had a

[24] Several industry sources report similar statistics, see e.g., *U.S. Auto Loan ABS Tracker: July 2019* from S&P Global.

58% recovery rate, and all minority borrowers' defaults had a 0% recovery rate, it could not explain the magnitude of our results.[25]

5.4 Perspective

Although we argue that our default tests are conservative, even the best studies of lending discrimination are unlikely to contain a single "bulletproof" test. Therefore, studies must rely on the offsetting strengths and weaknesses of their various tests in order to isolate discrimination. The breadth of the tests in our study is a strong point, and we encourage those evaluating alternative explanations to check their compatibility with each of the five findings outlined above.

6. Conclusion

Automobiles are an important asset for U.S. households, and most purchases are financed, which makes auto loans the most widely used form of installment credit with over 100 million borrowers. Despite this credit market's importance, it is relatively unregulated and opaque. This environment, combined with anecdotal evidence of discrimination, has led regulators like the CFPB to express concerns about discrimination in auto lending. However, data limitations have prevented financial economists from conducting market-wide tests for discrimination.

[25] For these back-of-the envelope calculations, we start with a loan amortization schedule based on average minority loan characteristics. We then create White/minority and default/non-default cases, and solve for the interest rate premium lenders would need to charge minorities to equate expected profits across races, given various assumptions about recovery rates. The results are not overly sensitive to other assumptions we make, including that defaults occur at 24 months, and that White and minority borrowers each default at the minority average rate of 3.5%. This assumption is conservative because minorities actually default slightly less than White borrowers, controlling for the same borrower and loan characteristics used in our interest rate tests (even if we exclude the interest rate from the controls in our default tests).

Our study implements a novel merge between two administrative datasets to overcome the data limitations, and our empirical tests provide strong evidence of discrimination in the U.S. auto loan market. Our estimates suggest that racial discrimination crowds out approximately 80,000 minority loans each year, and costs the average minority borrower roughly $410 in present value terms due to higher interest rates. Minorities living in areas where racial biases are more prevalent and lending competition is less intense bear the greatest costs of discrimination. These cross-sectional patterns, a falsification test based on credit cards, and an outcome test (Becker (1957, 1993)) based on auto loan defaults, all point towards racial biases as the primary driver of our findings.

Our study also provides the first analysis of the CFPB's increased anti-discrimination enforcement efforts that started in 2013. We find that the CFPB's efforts led to a 60% decrease in the additional interest that minorities pay on auto loans, with no concomitant decrease in minority credit approval rates. However, CFPB oversight is an area of active debate, and in 2018, Congress passed a joint resolution nullifying the 2013 Bulletin that the CFPB used to spearhead its initiative. Further exploration of the determinants of discrimination in this market, and of the viability of future policy interventions, are promising areas for future research.

References

Agan, Amanda, and Sonja Starr, 2018, Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment, *Quarterly Journal of Economics* 133, 191–235.

Akey, Pat, Christine Dobridge, Rawley Z. Heimer, and Stefan Lewellen, 2018, Pushing Boundaries: Political Redistricting and Consumer Credit, *Available at SSRN: https://ssrn.com/abstract=3031604*.

Akey, Pat, Rawley Z. Heimer, and Stefan Lewellen, 2018, Politicizing Consumer Credit, *Available at SSRN: https://ssrn.com/abstract=3042896*.

Arrow, Kenneth, 1973, *The Theory of Discrimination*. Ed. Orley Ashenfelter and Albert Rees *Discrimination in Labor Markets* (Princeton University Press).

Avery, Robert B., Mary F. Bilinski, Brian K. Bucks, Christine Chai, Tim Critchfield, Ian H. Keith, Ismail E. Mohamed, Forrest W. Pafenberg, Saty Patrabansh, Jay D. Schultz, and Claudia E. Wood, 2017, A Profile of 2013 Mortgage Borrowers: Statistics from the National Survey of Mortgage Originations.

Ayers, Ian, and Peter Siegelman, 1995, Race and Gender Discrimination in Bargaining for a New Car, *American Economic Review* 85, 304–321.

Ayres, Ian, 2002, Outcome Tests of Racial Disparities in Police Practices, *Justice Research and Policy* 4, 131–142.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, 2019, Consumer-Lending Discrimination in the FinTech Era, *Available at SSRN: https://ssrn.com/abstract=3063448*.

Bayer, Patrick, Fernando Ferreira, and Stephen L. Ross, 2018, What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders, *Review of Financial Studies* 31, 175–205.

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

Becker, Gary S., 1993, Nobel Lecture: The Economic Way of Looking at Behavior, *Journal of Political Economy* 101, 385–409.

Benmelech, Efraim, Ralf R. Meisenzahl, and Rodney Ramcharan, 2017, The Real Effects of Liquidity During the Financial Crisis: Evidence From Automobiles, *Quarterly Journal of Economics* 132, 317–365.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan, 1994, Race, Redlining, and Residential Mortgage Loan Performance, *Journal of Real Estate Finance and Economics* 9, 263–294.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan, 1998, Discrimination, Competition, and Loan Performance in FHA Mortgage Lending, *Review of Economics and Statistics* 80, 241–250.

Bertrand, Marianne, and Sendhil Mullainathan, 2004, Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, *American Economic Review* 94, 991–1013.

Bhutta, Neil, and Benjamin J. Keys, 2016, Interest Rates and Equity Extraction During the Housing Boom, *American Economic Review* 106, 1742–1774.

Blanchflower, David G., Phillip B. Levine, and David J. Zimmerman, 2003, Discrimination in the Small-Business Credit Market, *Review of Economics and Statistics* 85, 930–943.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope, 2019, Inaccurate Statistical Discrimination, *NBER Working Paper No. 25935. Available at SSRN: https://ssrn.com/abstract=3401641*.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, 2016, Stereotypes, *Quarterly Journal of Economics* 131, 1753–1794.

Brevoort, Kenneth P., 2011, Credit Card Redlining Revisited, *Review of Economics and Statistics* 93, 714–724.

Brown, James R., J. Anthony Cookson, and Rawley Z. Heimer, 2019, Growing Up Without Finance, *Journal of Financial Economics* 134, 591–616.

Brueckner, Jan K., 1996, Default Rates and Mortgage Discrimination: A View of the Controversy, *Cityscape: A Journal of Policy Development and Research* 2, 65–68.

Buchak, Greg, and Adam Jørring, 2017, Does Competition Reduce Racial Discrimination in Lending?, *Available at SSRN: https://ssrn.com/abstract=2951843*.

Charles, Kerwin Kofi, and Jonathan Guryan, 2008, Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination, *Journal of Political Economy* 116, 773–809.

Charles, Kerwin Kofi, Erik Hurst, and Melvin Stephens, 2008, Rates for Vehicle Loans: Race and Loan Source, *American Economic Review (P&P)* 98, 315–320.

Cohen-Cole, Ethan, 2011, Credit Card Redlining, *The Review of Economics and Statistics* 93, 700–713.

Cohen, Mark A., 2012, Imperfect Competition in Auto Lending: Subjective Markup, Racial Disparity, and Class Action Litigation, *Review of Law and Economics* 8, 21–58.

Day, Theodore E., and S. J. Liebowitz, 1998, Mortgage Lending to Minorities: Where's the bias?, *Economic Inquiry* 36, 3–28.

Deng, Yongheng, and Stuart Gabriel, 2006, Risk-Based Pricing and the Enhancement of Mortgage Credit Availability among Underserved and Higher Credit-Risk Populations, *Journal of Money, Credit and Banking* 38, 1431–1460.

Dobbie, Will, Andres Liberman, Daniel Paravisini, and Vikram Pathania, 2019, Measuring Bias in Consumer Lending, *Available at SSRN: https://ssrn.com/abstract=3258699*.

Dougal, Casey, Pengjie Gao, William J. Mayew, and Christopher A. Parsons, 2019, What's in a (School) Name? Racial Discrimination in Higher Education Bond Markets, *Journal of Financial Economics* 134, 570–590.

Ferguson, Michael F., and Stephen R. Peters, 1995, What Constitutes Evidence of Discrimination in Lending?, *Journal of Finance* 50, 739–748.

Galster, George C., 1993, The Facts of Lending Discrimination Cannot Be Argued Away by Examining Default Rates, *Housing Policy Debate* 4, 141–146.

Gross, David B., and Nicholas S. Souleles, 2002, Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data, *Quarterly Journal of Economics* 117, 149–185.

Hanson, Andrew, Zackary Hawley, Hal Martin, and Bo Liu, 2016, Discrimination in mortgage lending: Evidence from a correspondence experiment, *Journal of Urban Economics* 92, 48–65.

Heckman, James J., 1998, Detecting Discrimination, *Journal of Economic Perspectives* 12, 101–116.

Lanning, Jonathan A., 2021, Testing Models of Economic Discrimination Using the Discretionary Markup of Indirect Auto Loans, *Working Paper*.

Lee, Donghoon, and Wilbert Van Der Klaauw, 2010, An Introduction to the FRBNY Consumer Credit Panel, *Federal Reserve Bank of New York Staff Reports*.

Levine, Ross, Alexey Levkov, and Yona Rubinstein, 2014, Bank Deregulation and Racial Inequality in America, *Critical Finance Review* 3, 1–48.

Mayer, Erik J., 2021, Big Banks, Household Credit Access, and Intergenerational Economic Mobility, *Available at SSRN: https://ssrn.com/abstract=3816308*.

Moore, Mary, 1996, Credit Scoring's Uses Expand as It Gains Acceptance, *The American Banker* 4A.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney, 1996, Mortgage Lending in Boston: Interpreting HMDA Data, *American Economic Review* 86, 25–53.

Phelps, Edmund S., 1972, The Statistical Theory of Racism and Sexism, *American Economic Review* 62, 659–661.

Pope, Devin G., and Justin R. Sydnor, 2011, What's in a Picture? Evidence of Discrimination from Prosper.com, *Journal of Human Resources* 46, 53–92.

Rice, Lisa, and Erich Schwartz Jr., 2018, Discrimination When Buying a Car: How the Color of Your Skin Can Affect Your Car-Shopping Experience, *National Fair Housing Alliance Report*.

Ross, Stephen L., 1996, Mortgage Lending Discrimination and Racial Differences in Loan Default, *Journal of Housing Research* 7, 117–126.

Ross, Stephen L., 1996, Flaws in the Use of Loan Defaults To Test for Mortgage Lending Discrimination, *Cityscape: A Journal of Policy Development and Research* 2, 41–48.

Ross, Stephen L., Margery Austin Turner, Erin Godfrey, and Robin R. Smith, 2008, Mortgage Lending in Chicago and Los Angeles: A Paired Testing Study of the Pre-application Process, *Journal of Urban Economics* 63, 902–919.

Ross, Stephen L., and John Yinger, 1999, The Default Approach to Studying Mortgage Discrimination: A Rebuttal, in Margery Austin Turner and in Felicity Skidmore ed.: *Mortgage Lending Discrimination: A Review of Existing Evidence*.

Shaffer, Sherrill, 1996, Evidence of Discrimination in Lending: An Extension, *Journal of Finance* 51, 1551–1554.

Stephens-Davidowitz, Seth, 2014, The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data, *Journal of Public Economics* 118, 26–40.

Tsosie, Claire, 2016, How Your Credit Card Issuer Determines Your Credit Limit, *U.S. News & World Report*.

Yinger, John, 1996, Why Default Rates Cannot Shed Light on Mortgage Discrimination, *Cityscape: A Journal of Policy Development and Research* 2, 25–31.

Table 1: Summary of the Credit Bureau/HMDA Match

This table summarizes the match between mortgages in the credit bureau data and mortgages in the Home Mortgage Disclosure Act (HMDA) data. This match ultimately leads to our panel dataset of credit bureau records with information on financial outcomes and borrower race/ethnicity (from HMDA). The starting sample of credit bureau mortgages contains both home purchase mortgages and refinance loans originated from 2010-2016. The borrower is required to apply for the loan on their own (i.e. joint applications are excluded), and to live (after the loan is originated) within a metropolitan statistical area, and the mortgage must be the borrower's only first-lien mortgage (i.e. mortgages for second homes are excluded). The matching between credit bureau mortgages and HMDA mortgages is done based on the following characteristics: whether the loan is for home purchase or refinancing, the loan origination year, the census tract of the property, the loan amount, whether the mortgage is conventional or through the Federal Housing Administration (FHA) or Veterans Administration (VA), and whether the loan is purchased by Fannie Mae or Freddie Mac on the secondary market. Only the mortgages in the HMDA data that are unique based on these matching variables are used as potential matches. Panel A shows the success rate of the matching approach. Panel B summarizes loan and borrower characteristics for the home purchase mortgages in the credit bureau data, the subsample that were successfully matched to HMDA, and the unmatched loans. The final two columns show the normalized difference and the result of a t-test comparing the mean of the matched sample to the mean of the unmatched sample. Panel C provides similar summary statistics for refinance loans.

Panel A: Match Rate					
	Credit Bureau Sample	Matched to HMDA	Match Rate		
Home Purchase Mortgages	107,085	66,345	61.96%		
Refinance Loans	65,046	52,115	80.12%		
All Loans	172,131	118,460	68.82%		
Panel B: Home Purchase Mortgages					
	Credit Bureau Sample	Matched to HMDA	Unmatched	Matched vs. Unmatched	
	(N = 107,085)	(N = 66,345)	(N = 40,740)	Norm. Diff	t-stat
Match Criteria					
Conventional Loan	0.631	0.623	0.643	-0.03	-4.70
FHA Loan	0.289	0.293	0.283	0.01	2.47
VA Loan	0.080	0.084	0.074	0.03	5.38
Fannie Mae	0.243	0.251	0.231	0.03	5.89
Freddie Mac	0.149	0.158	0.134	0.05	9.98
Loan Amount	192,142	193,758	189,508	0.02	3.99
Non-Match Characteristics					
Credit Score $_{t-1}$	717	719	715	0.04	7.78
Age	42.0	41.1	43.3	-0.12	-21.97
Have Mortgage $_{t-1}$	0.310	0.254	0.401	-0.23	-33.37
Total Debt $_{t-1}$	78,802	66,519	98,895	-0.19	-24.01
Past Due Debt $_{t-1}$	311	283	356	-0.02	-3.43
Auto Debt $_{t-1}$	8,176	8,145	8,227	-0.00	-1.00
Panel C: Refinance Loans					
	Credit Bureau Sample	Matched to HMDA	Unmatched	Matched vs. Unmatched	
	(N = 65,046)	(N = 52,115)	(N = 12,931)	Norm. Diff	t-stat
Match Criteria					
Conventional Loan	0.815	0.814	0.821	-0.01	-1.85
FHA Loan	0.125	0.125	0.124	0.00	0.27
VA Loan	0.060	0.061	0.055	0.02	2.74
Fannie Mae	0.307	0.308	0.301	0.01	1.49
Freddie Mac	0.202	0.210	0.171	0.07	9.86
Loan Amount	196,062	193,971	204,491	-0.06	-7.21
Non-Match Characteristics					
Credit Score $_{t-1}$	738	738	739	-0.01	-1.04
Age	49.4	49.6	48.7	0.05	7.61
Have Mortgage $_{t-1}$	1.00	1.00	1.00	.	.
Total Debt $_{t-1}$	214,145	212,926	219,054	-0.03	-3.97
Past Due Debt $_{t-1}$	233	229	250	-0.00	-0.62
Auto Debt $_{t-1}$	8,128	8,058	8,409	-0.02	-2.67

Table 2: Does Borrower Race Affect the Credit Bureau/HMDA Match?

This table presents regressions that examine the determinants of whether a mortgage in the Home Mortgage Disclosure Act (HMDA) data is matched to a credit bureau record through the process described in Section 3.3. The sample includes all home purchase mortgages and refinance loans in the HMDA data that are first liens on owner-occupied properties located in metropolitan statistical areas, originated from 2010-2016. The loans are also required to have only one applicant (i.e. joint applications are excluded). Through the matching process described in Section 3.3, these mortgages from HMDA are matched to mortgages reported in a nationally representative 1% sample of credit bureau records. For the regressions in this table, the outcome variable is an indicator for whether the HMDA mortgage was matched to a credit bureau record, and the explanatory variables are loan and borrower characteristics from the HMDA data. Columns 1, 2, and 3 present the results for the full sample, the sample of home purchase mortgages, and the sample of refinance loans, respectively. The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of being matched). The standard errors are clustered by census tract–year.

	Full Sample	Home Purchase Mortgages	Refinance Loans
	Matched	Matched	Matched
	(1)	(2)	(3)
Match Criteria			
FHA Loan	0.008	-0.116^{***}	0.005
	(0.006)	(0.008)	(0.010)
VA Loan	0.057^{***}	-0.025^{**}	0.021
	(0.009)	(0.012)	(0.016)
Purchased by Fannie Mae	0.107^{***}	0.212^{***}	0.093^{***}
	(0.005)	(0.010)	(0.006)
Purchased by Freddie Mac	0.130^{***}	0.281^{***}	0.114^{***}
	(0.006)	(0.013)	(0.007)
Log(Loan Amount)	0.026^{***}	-0.016^{*}	-0.003
	(0.005)	(0.009)	(0.006)
Non-Match Characteristics			
Black	-0.157	-0.167	-0.345
	(0.154)	(0.225)	(0.215)
Hispanic	-0.013	-0.320^{*}	0.137
	(0.129)	(0.184)	(0.188)
Black X Log(Income)	0.012	0.012	0.031
	(0.014)	(0.020)	(0.019)
Hispanic X Log(Income)	0.001	0.027	-0.011
	(0.012)	(0.017)	(0.017)
Log(Income)	-0.137^{***}	-0.189^{***}	-0.060^{***}
	(0.004)	(0.008)	(0.005)
Census Tract-by-Year FE	Yes	Yes	Yes
R-Squared	0.022	0.044	0.042
Observations	18,085,605	8,921,824	9,141,794

Table 3: Summary Statistics from the Credit Bureau/HMDA Matched Panel

This table presents summary statistics describing the Credit Bureau/HMDA Matched Panel (see Section 3.3 for information about the construction of this dataset). Panel A provides a snapshot of the matched dataset in 2010, and compares it to a 2010 snapshot of the full credit bureau dataset for reference. Column 1 presents the sample means and standard deviations (in brackets) for the full credit bureau dataset, Column 2 shows these statistics for the Credit Bureau/HMDA Matched Panel, and Columns 3-5 show the statistics for the White, Black, and Hispanic borrowers in the matched dataset, respectively. The *Income* and *Debt to Income* variables are only available for borrowers in the matched dataset because they use HMDA reported income. Panel B shows similar summary statistics for the person-years in which individuals apply for auto loans from 2005-2017.

Panel A: 2010 Snapshot	Full Credit Bureau Sample (N = 2,597,877)	Matched Sample (N = 78,932)	White (N = 65,207)	Black (N = 6,338)	Hispanic (N = 7,387)
Credit Score $_{t-1}$	669	707	715	660	678
	[113]	[87.2]	[84.0]	[94.9]	[89.9]
Age	49.8	42.3	42.6	42.8	39.9
	[18.9]	[13.9]	[14.1]	[13.5]	[12.9]
Have Mortgage $_{t-1}$	0.295	0.552	0.577	0.431	0.428
	[0.456]	[0.497]	[0.494]	[0.495]	[0.495]
Total Debt $_{t-1}$	67,475	123,552	129,415	92,478	98,034
	[164,108]	[166,047]	[170,688]	[125,459]	[148,536]
Past Due Debt $_{t-1}$	1,890	805	654	1,609	1,457
	[12,611]	[4,750]	[4,319]	[6,797]	[5,991]
Auto Debt $_{t-1}$	3,665	6,587	6,468	7,161	7,152
	[8,917]	[11,019]	[10,958]	[11,065]	[11,478]
Income	.	73,295	75,805	62,686	60,239
	.	[83,244]	[88,953]	[37,173]	[51,847]
Debt to Income $_{t-1}$	.	1.86	1.89	1.54	1.82
	.	[2.64]	[2.42]	[2.30]	[4.30]

Panel B: Auto Loan Applicants (2005-2017)	Full Credit Bureau Sample (N = 4,406,635)	Matched Sample (N = 218,476)	White (N = 175,911)	Black (N = 18,408)	Hispanic (N = 24,157)
Credit Approval (Auto)	0.722	0.832	0.847	0.783	0.757
	[0.448]	[0.374]	[0.360]	[0.412]	[0.429]
Credit Score $_{t-1}$	663	697	705	655	673
	[105]	[82.4]	[79.8]	[88.6]	[82.1]
Age	43.2	41.7	42.0	42.2	39.7
	[14.9]	[13.1]	[13.2]	[12.9]	[12.3]
Have Mortgage $_{t-1}$	0.401	0.643	0.661	0.560	0.569
	[0.490]	[0.479]	[0.473]	[0.496]	[0.495]
Total Debt $_{t-1}$	102,200	152,308	158,553	120,910	130,351
	[193,180]	[185,190]	[192,014]	[132,993]	[162,920]
Past Due Debt $_{t-1}$	1,667	639	521	1,269	1,027
	[8,360]	[4,725]	[4,663]	[4,779]	[5,066]
Auto Debt $_{t-1}$	9,170	10,986	10,880	10,814	11,906
	[15,190]	[15,752]	[15,748]	[15,159]	[16,191]
Income	.	78,395	81,578	65,480	65,061
	.	[97,191]	[104,641]	[38,979]	[64,490]
Debt to Income $_{t-1}$	.	2.18	2.18	1.98	2.31
	.	[2.74]	[2.47]	[2.21]	[4.48]

Table 4: The Effect of Applicant Race on Auto Credit Approval

The tests in this table regress a measure of auto credit approval on race, individual characteristics, and ZIP code characteristics. The outcome variable is an indicator for the person successfully opening a new auto loan, and the sample includes all person-years in which individuals apply for auto loans. Columns 4 and 5 restrict the sample to applicants with subprime, and prime credit scores, respectively. The individual level data consist of credit bureau records that have been matched to Home Mortgage Disclosure Act records (see Section 3.3 for details). The dataset includes credit bureau records for the years 2005-2017. The coefficients are reported in terms of percentage points (i.e., a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of credit approval). The standard errors are clustered by state-year.

	Full Sample			Subprime Borrowers	Prime Borrowers
	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)
	(1)	(2)	(3)	(4)	(5)
Demographics					
Minority	-4.465^{***}	-1.480^{***}	-1.661^{***}	-2.375^{***}	-0.840^{***}
	(0.289)	(0.259)	(0.332)	(0.399)	(0.271)
Minority X Hispanic			0.328		
			(0.410)		
Female	1.598^{***}	1.115^{***}	1.126^{***}	1.492^{***}	1.133^{***}
	(0.173)	(0.169)	(0.169)	(0.352)	(0.180)
Age	0.042^{***}	-0.067^{***}	-0.066^{***}	0.017	-0.072^{***}
	(0.008)	(0.008)	(0.008)	(0.015)	(0.009)
Log(Income)	3.886^{***}	1.704^{***}	1.711^{***}	4.586^{***}	0.736^{***}
	(0.184)	(0.180)	(0.181)	(0.407)	(0.199)
Credit Characteristics					
Credit Score $_{t-1}$		0.057^{***}	0.057^{***}	0.161^{***}	0.013^{***}
		(0.002)	(0.002)	(0.004)	(0.003)
Log(Total Debt $_{t-1}$)		0.866^{***}	0.866^{***}	0.403^{***}	0.868^{***}
		(0.053)	(0.053)	(0.070)	(0.077)
Debt to Income $_{t-1}$		-0.032	-0.032	0.040	-0.220^{***}
		(0.062)	(0.062)	(0.119)	(0.079)
Log(Past Due Debt $_{t-1}$)		-1.179^{***}	-1.178^{***}	-0.745^{***}	-0.413^{***}
		(0.051)	(0.051)	(0.061)	(0.066)
ZIP Code Characteristics					
Log(Personal Income Per Capita)	1.087^{*}	-0.095	-0.076	0.573	-0.350
	(0.629)	(0.611)	(0.611)	(1.088)	(0.701)
Log(Population Density)	-0.014	0.009	0.010	0.067	0.037
	(0.065)	(0.065)	(0.065)	(0.142)	(0.072)
Bachelors Degree	5.108^{***}	1.406	1.373	3.907	1.765
	(1.254)	(1.236)	(1.238)	(2.374)	(1.372)
Commute Using Car	12.020^{***}	10.569^{***}	10.533^{***}	12.663^{***}	8.640^{***}
	(1.194)	(1.149)	(1.146)	(2.317)	(1.276)
State-by-Year FE	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes
R-Squared	0.047	0.085	0.085	0.105	0.047
Observations	218,300	214,534	214,534	68,494	146,036

Table 5: Where Does Applicant Race Have the Largest Impact on Auto Credit Approval?

The tests in this table regress a measure of auto credit approval on race, individual characteristics, and ZIP code characteristics. The outcome variable is an indicator for the person successfully opening a new auto loan, and the sample includes all person-years in which individuals apply for auto loans. The explanatory variables of interest are indicators for the applicant belonging to a racial minority, and the interaction of *Minority* with indicators for living in a state in the top tercile of racial bias (based on Google Search Volume for racial slurs), living in a county in the top tercile of the Herfindahl index for bank deposits (*Low Banking Competition*), living in a ZIP code in the bottom tercile of population density (*Rural*), or living in a county in the top quartile in terms of the share of non-bank auto lending (*High Non-Bank Financing*). These county quartile assignments come from Benmelech et. al. (2017) who compute them as of 2008Q1 using proprietary data. The individual level data are from the matched dataset of credit bureau records and Home Mortgage Disclosure Act records (see Section 3.3 for details). The dataset includes credit bureau records for the years 2005-2017. The included individual controls and ZIP code controls are the same as those reported in Table 4. The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of credit approval). The standard errors are clustered by state-year.

	Credit Approval (Auto) (1)	Credit Approval (Auto) (2)	Credit Approval (Auto) (3)	Credit Approval (Auto) (4)	Credit Approval (Auto) (5)
Minority	-1.480^{***}	-0.906^{***}	-1.268^{***}	-1.259^{***}	-1.509^{***}
	(0.259)	(0.254)	(0.255)	(0.298)	(0.246)
Minority X High Racial Bias State		-1.910^{***}			
		(0.443)			
Minority X Low Banking Competition			-0.728^{*}		
			(0.424)		
Low Banking Competition			0.214		
			(0.207)		
Minority X High Non-Bank Financing				-0.351	
				(0.401)	
High Non-Bank Financing				-0.782^{***}	
				(0.241)	
Minority X Rural					0.117
					(0.461)
Rural					-0.124
					(0.303)
Individual Controls	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes
R-Squared	0.085	0.085	0.085	0.085	0.085
Observations	214,534	214,534	214,534	214,534	214,534

Table 6: Falsification Test – Race and Credit Card Lending

The tests in this table examine racial differences in credit card lending as a falsification test—we expect these lenders' automated approach to reduce discrimination. Specifically, we select the subset of auto loan applicants who applied for credit cards (or credit card limit increases) during the same year as their auto loan application. We then regress the dollar increase in the borrower's total credit card limit across all cards (average increase = $3,090) on race, individual characteristics, and ZIP code characteristics. Following prior tests, the explanatory variables of interest are indicators for the applicant belonging to a racial minority, and the interaction of *Minority* with indicators for living in a state in the top tercile of racial bias (based on Google Search Volume for racial slurs), living in a county in the top tercile of the Herfindahl index for bank deposits (*Low Banking Competition*), living in a ZIP code in the bottom tercile of population density (*Rural*), or living in a county in the top quartile in terms of the share of non-bank auto lending (*High Non-Bank Financing*). These county quartile assignments come from Benmelech et. al. (2017) who compute them as of 2008Q1 using proprietary data. The individual level data are from the matched dataset of credit bureau records and Home Mortgage Disclosure Act records (see Section 3.3 for details). The dataset includes credit bureau records for the years 2005-2017. The included individual controls and ZIP code controls are the same as those reported in Table 4. The standard errors are clustered by state-year.

	CC Limit Inc. (1)	CC Limit Inc. (2)	CC Limit Inc. (3)	CC Limit Inc. (4)	CC Limit Inc. (5)
Minority	38.23	−10.44	110.36	94.97	7.12
	(73.09)	(84.07)	(85.54)	(101.61)	(83.03)
Minority X High Racial Bias State		181.61			
		(154.99)			
Minority X Low Banking Competition			−234.78		
			(145.53)		
Low Banking Competition			74.16		
			(72.00)		
Minority X High Non-Bank Financing				−121.48	
				(135.60)	
High Non-Bank Financing				179.29**	
				(84.55)	
Minority X Rural					121.35
					(155.43)
Rural					50.05
					(105.87)
Individual Controls	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes
R-Squared	0.075	0.075	0.075	0.075	0.075
Observations	124,601	124,601	124,601	124,601	124,601

Table 7: Summary Statistics on the Auto Loans in the Credit Bureau/HMDA Matched Panel

This table presents summary statistics on the auto loans in the Credit Bureau/HMDA Matched Panel. The sample is constructed at the auto loan level and includes one observation for each new auto loan originated from 2011-2017 (the time period over which interest rates are available). To be included, the loan must be the borrower's only outstanding auto loan at origination, so that the loan's performance can be tracked in the credit bureau data. For *Auto Loan Default*, the statistics are based on the 2011-2015 subsample, because we need 2 years after origination to compute this variable. Column 1 presents the sample means and standard deviations (in brackets) for the full sample. Columns 2-5 present these statistics for the subsamples of White, minority, subprime, and prime borrowers respectively.

	Full Sample (N = 25,697)	White Borrowers (N = 20,823)	Minority Borrowers (N = 4,874)	Subprime Borrowers (N = 6,115)	Prime Borrowers (N = 19,574)
Demographics					
Female	0.425	0.422	0.437	0.407	0.430
	[0.494]	[0.494]	[0.496]	[0.491]	[0.495]
Age	43.5	43.7	42.7	40.2	44.5
	[13.7]	[13.8]	[13.0]	[12.5]	[13.9]
Income	67,354	69,276	59,144	59,396	69,847
	[40,075]	[41,296]	[33,143]	[32,920]	[41,758]
Auto Loan Variables					
Auto Loan Default	0.017	0.013	0.035	0.055	0.004
	[0.130]	[0.114]	[0.184]	[0.228]	[0.066]
Auto Loan Rate	0.060	0.057	0.077	0.100	0.048
	[0.048]	[0.045]	[0.058]	[0.063]	[0.034]
Auto Loan Amount	21,233	21,017	22,157	20,058	21,603
	[10,201]	[10,178]	[10,244]	[9,897]	[10,266]
Auto Loan to Income Ratio	0.389	0.373	0.455	0.400	0.385
	[0.248]	[0.238]	[0.279]	[0.248]	[0.249]
Auto Loan Term (Months)	65.1	64.6	67.5	66.9	64.6
	[13.2]	[13.2]	[12.7]	[13.6]	[13.0]
Credit Characteristics					
Credit Score $_{t-1}$	717	724	685	604	752
	[78.4]	[75.3]	[83.3]	[44.7]	[47.0]
Total Debt $_{t-1}$	129,567	133,584	112,407	96,101	140,053
	[123,667]	[125,091]	[115,843]	[115,554]	[124,269]
Debt to Income $_{t-1}$	2.08	2.09	2.06	1.73	2.20
	[1.78]	[1.75]	[1.90]	[1.87]	[1.73]
Past Due Debt $_{t-1}$	308	237	609	1,129	51
	[1,312]	[1,157]	[1,800]	[2,366]	[483]
Auto Debt Share	0.278	0.270	0.315	0.397	0.241
	[0.311]	[0.306]	[0.329]	[0.372]	[0.279]

Table 8: The Effect of Borrower Race on Auto Loan Interest Rates

The regressions in this table examine the effect of borrower race on auto loan interest rates. The sample is constructed at the auto loan level from the matched dataset of credit bureau records and Home Mortgage Disclosure Act records (see Section 3.3 for details). The sample includes one observation for each new auto loan originated from 2011-2017 (the time period over which interest rates are available), and we require the loan to be the borrower's only outstanding auto loan at origination. The explanatory variables of interest are indicators for the borrower belonging to a racial minority, and the interaction of *Minority* with indicators for living in a state in the top tercile of racial bias (based on Google Search Volume for racial slurs), living in a county in the top tercile of the Herfindahl index for bank deposits (*Low Banking Competition*), living in a ZIP code in the bottom tercile of population density (*Rural*), or living in a county in the top quartile in terms of the share of non-bank auto lending (*High Non-Bank Financing*). These county quartile assignments come from Benmelech et. al. (2017) who compute them as of 2008Q1 using proprietary data. The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the interest rate). The standard errors are clustered by state-year.

	Rate (1)	Rate (2)	Rate (3)	Rate (4)	Rate (5)	Rate (6)
Demographics and Interaction Terms						
Minority	1.600***	0.704***	0.442***	0.614***	0.648***	0.691***
	(0.169)	(0.117)	(0.084)	(0.110)	(0.137)	(0.120)
Minority X High Racial Bias State			0.805***			
			(0.166)			
Minority X Low Banking Competition				0.293		
				(0.208)		
Low Banking Competition				0.052		
				(0.065)		
Minority X High Non-Bank Financing					0.083	
					(0.175)	
High Non-Bank Financing					0.197**	
					(0.093)	
Minority X Rural						0.056
						(0.223)
Rural						−0.023
						(0.078)
Female	−0.397***	−0.259***	−0.265***	−0.260***	−0.259***	−0.259***
	(0.052)	(0.039)	(0.038)	(0.039)	(0.039)	(0.039)
Age	−0.014***	0.012***	0.012***	0.012***	0.012***	0.012***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log(Income)	−0.228	0.400***	0.396***	0.396***	0.397***	0.400***
	(0.143)	(0.130)	(0.130)	(0.130)	(0.130)	(0.130)
Auto Loan Characteristics						
Auto Loan Term Indicators	Yes	Yes	Yes	Yes	Yes	Yes
Log(Auto Loan Amount)	−2.922***	−2.674***	−2.674***	−2.669***	−2.677***	−2.674***
	(0.137)	(0.143)	(0.143)	(0.142)	(0.143)	(0.143)
Auto Loan to Income Ratio	−0.326	0.458	0.450	0.443	0.457	0.458
	(0.276)	(0.281)	(0.280)	(0.281)	(0.281)	(0.282)
Credit Characteristics						
Credit Score $_{t-1}$		−0.019***	−0.019***	−0.019***	−0.019***	−0.019***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(Total Debt $_{t-1}$)		−0.129***	−0.129***	−0.129***	−0.129***	−0.129***
		(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Debt to Income $_{t-1}$		−0.038*	−0.038*	−0.037*	−0.038*	−0.038*
		(0.019)	(0.019)	(0.020)	(0.020)	(0.019)
Log(Past Due Debt $_{t-1}$)		0.336***	0.335***	0.335***	0.337***	0.336***
		(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Auto Debt Share		0.595***	0.594***	0.593***	0.597***	0.594***
		(0.142)	(0.142)	(0.142)	(0.143)	(0.142)
ZIP Code Characteristics						
Log(Personal Income Per Capita)	0.031	0.071	0.035	0.085	0.034	0.072
	(0.244)	(0.187)	(0.182)	(0.191)	(0.175)	(0.186)
Log(Population Density)	−0.023	0.010	0.007	0.013	−0.003	0.006
	(0.031)	(0.022)	(0.021)	(0.022)	(0.021)	(0.031)
Bachelors Degree	−2.422***	−0.902**	−0.861**	−0.916**	−0.841**	−0.914**
	(0.535)	(0.399)	(0.390)	(0.404)	(0.374)	(0.386)
Commute Using Car	−1.252***	−0.713**	−0.680**	−0.690**	−0.701**	−0.731***
	(0.358)	(0.293)	(0.292)	(0.289)	(0.304)	(0.271)
Origination Month Indicators	Yes	Yes	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.255	0.440	0.441	0.441	0.441	0.440
Observations	25,531	25,523	25,523	25,523	25,523	25,523

Table 9: Borrower Race and Auto Loan Default Rates

The regressions in this table test whether borrower race affects the likelihood of auto loan default. The sample is constructed at the auto loan level from the matched dataset of credit bureau records and Home Mortgage Disclosure Act records (see Section 3.3 for details). The sample includes one observation for each new auto loan originated from 2011-2015 (the period over which we can compute both interest rates and our indicator for default). The auto loans are required to be originated after the match between the credit bureau and HMDA records, and the loan must be the borrower's only outstanding auto loan at origination. The outcome variable is an indicator for whether the borrower became 90 or more days delinquent on the loan during the year of origination or the following two calendar years. Column 1 shows the results for the full sample, and Columns 2 and 3 show the results for borrowers with subprime and prime credit scores, respectively. The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the default rate). The standard errors are clustered by state-year.

	Full Sample	Subprime Borrowers	Prime Borrowers
	Auto Loan Default	Auto Loan Default	Auto Loan Default
	(1)	(2)	(3)
Demographics			
Minority	-0.237	-2.337^{**}	0.288
	(0.397)	(1.125)	(0.345)
Female	0.122	0.619	-0.081
	(0.216)	(1.118)	(0.132)
Age	0.016^{*}	0.020	0.006
	(0.009)	(0.049)	(0.008)
Log(Income)	-0.601	-1.734	-0.514
	(0.450)	(1.847)	(0.378)
Auto Loan Characteristics			
Auto Loan Term Indicators	Yes	Yes	Yes
Log(Auto Loan Amount)	1.653^{***}	4.824^{**}	0.595^{*}
	(0.436)	(2.104)	(0.358)
Auto Loan to Income Ratio	-1.697	-3.826	-0.564
	(1.045)	(4.028)	(0.902)
Auto Loan Rate	45.656^{***}	72.553^{***}	16.548^{***}
	(6.616)	(15.369)	(5.820)
Credit Characteristics			
Credit Score $_{t-1}$	-0.014^{***}	-0.060^{***}	-0.005^{**}
	(0.003)	(0.019)	(0.002)
Log(Total Debt $_{t-1}$)	-0.309^{**}	-0.707^{*}	-0.035
	(0.145)	(0.408)	(0.068)
Debt to Income $_{t-1}$	0.261^{*}	0.894^{*}	0.032
	(0.157)	(0.494)	(0.091)
Log(Past Due Debt $_{t-1}$)	0.492^{***}	0.191	0.224^{*}
	(0.117)	(0.168)	(0.118)
Auto Debt Share	2.890^{***}	6.579	0.623
	(1.064)	(4.255)	(0.604)
ZIP Code Characteristics			
Log(Personal Income Per Capita)	-0.568	-3.414	0.085
	(0.802)	(3.807)	(0.511)
Log(Population Density)	-0.001	0.260	-0.049
	(0.099)	(0.431)	(0.054)
Bachelors Degree	0.123	0.086	-0.764
	(1.643)	(8.553)	(0.911)
Commute Using Car	-0.492	13.956	-3.155^{**}
	(2.176)	(8.669)	(1.539)
Origination Month Indicators	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes
R-Squared	0.096	0.173	0.054
Observations	10,509	2,005	8,480

Table 10: The 2013 CFPB Intervention and Racial Disparities in Auto Credit

This table examines the effect of the 2013 CFPB Intervention on racial disparities in auto loan interest rates and approval rates. Columns 1-3 examine the interest rates on auto loans from our Credit Bureau/HMDA Matched Panel that were originated from 2011-2017 (the time period over which interest rates are available). The explanatory variables of interest are indicators for the person belonging to a racial minority, and the interaction of *Minority* with indicators for the application occurring in 2014 or later (*Post*), for the person living in a county in the top quartile of non-bank auto lending share (*High Non-Bank Financing*), and for the person living in a state in the top tercile of racial bias based on Google Search Volume for racial slurs (*High Racial Bias State*). Column 1 presents a differences-in-differences test for whether the CFPB intervention affected the additional interest minorities' are charged on auto loans, and Columns 2 and 3 present triple-differences tests for whether the CFPB intervention had a larger effect in certain areas (note that several of the interaction terms are subsumed by the State-by-Year FE). Columns 4-6 present similar tests examining the effect of the CFPB intervention on auto credit approval. In these tests, the outcome variable is an indicator for the person successfully opening a new auto loan, and the sample includes all person-years in our data in which individuals apply for auto loans from 2011-2017. The control variables included in the tests in this table are the same as those reported in previous tables. The coefficients are reported in terms of percentage points, i.e., a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the interest rate, or in the probability of credit approval. The standard errors are clustered by state-year.

	Outcome Var = Auto Loan Rate			Outcome Var = Credit Approval (Auto)		
	(1)	(2)	(3)	(4)	(5)	(6)
Minority	0.838^{***}	0.614^{***}	0.538^{***}	-1.813^{***}	-2.097^{***}	-1.118^{*}
	(0.132)	(0.205)	(0.135)	(0.491)	(0.544)	(0.605)
Minority X Post	-0.490^{***}	-0.156	-0.401^{**}	0.607	1.451^{*}	0.951
	(0.163)	(0.233)	(0.175)	(0.618)	(0.751)	(0.730)
Minority X Post X High Non-Bank Financing		-0.625^{**}			-1.526	
		(0.293)			(1.073)	
Minority X High Non-Bank Financing		0.401^{*}			0.552	
		(0.242)			(0.811)	
Post X High Non-Bank Financing		0.021			-0.739	
		(0.150)			(0.541)	
High Non-Bank Financing		0.139			-0.269	
		(0.109)			(0.389)	
Minority X Post X High Racial Bias State			-0.312			-1.085
			(0.307)			(1.149)
Minority X High Racial Bias State			0.950^{***}			-2.270^{***}
			(0.238)			(0.862)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Auto Loan Characteristics	Yes	Yes	Yes	n/a	n/a	n/a
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origination Month Indicators	Yes	Yes	Yes	n/a	n/a	n/a
State-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.398	0.398	0.399	0.057	0.057	0.057
Observations	25,523	25,523	25,523	130,867	130,867	130,867

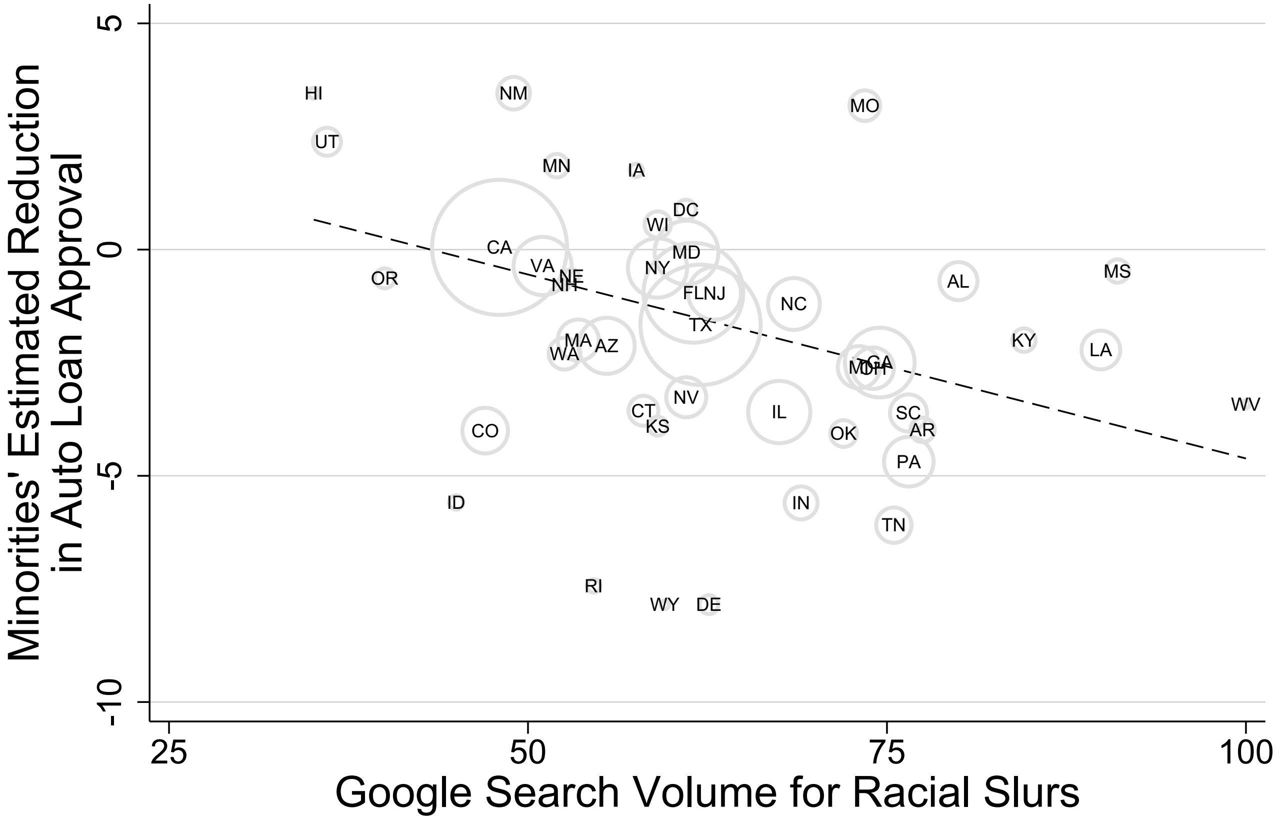

Figure 1
Auto Lending Discrimination and Racial Biases
This figure plots our point estimates of the reduction in auto loan approval rates that minorities face in each U.S. state against the prevalence of racial biases in the state measured using the Google Search Volume for racial slurs (following Stephens-Davidowitz (2014)). The point estimates come from a regression of auto loan approval on controls, similar to the regression reported in Column 2 of Table 4, except that the *Minority* indicator is interacted with indicators for each state and the District of Columbia. We require that our sample contains at least 25 minority applications in a state in order to report the *State*$_i$ *X Minority* coefficient estimate (excludes 6 states with small minority populations). The size of the circle plotted for each state is proportional to the number of minority applications in the state. Each state is weighted by the number of minority applications when computing the best fit line in the plot, and the correlation between the *State*$_i$ *X Minority* coefficient and the *Racial Slur GSV*, which is -0.49 (p-value = 0.001).

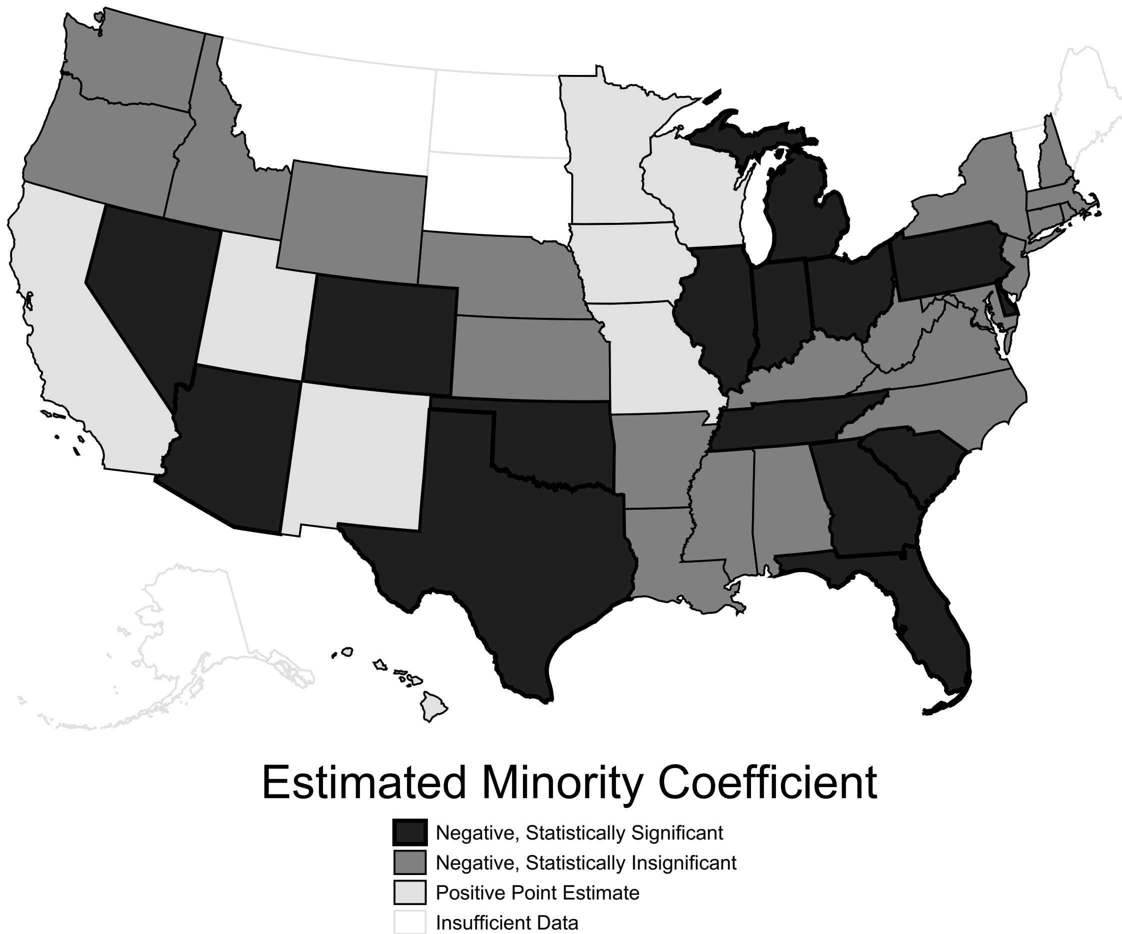

Figure 2

Where is the Evidence of Auto Lending Discrimination Strongest?

This figure presents a map categorizing U.S. states based on whether we find statistically significant evidence that minorities face reduced access to auto credit in the state. Our estimates of whether minorities face reduced access to credit come from a regression of auto loan approval on controls, similar to the regression reported in Column 2 of Table 4, except that the *Minority* indicator is interacted with indicators for each state and the District of Columbia. We require that our sample contains at least 25 minority applications in a state in order to make any inferences about discrimination in the state based on the $State_i$ *X Minority* coefficient (this excludes 6 states with small minority populations). In the states shaded black, we find statistically significant evidence (p-value ≤ 0.1) that minorities face a reduced auto loan approval rate. In the dark gray states, we find negative but statistically insignificant $State_i$ *X Minority* coefficients, and in the light gray states we find positive coefficients.

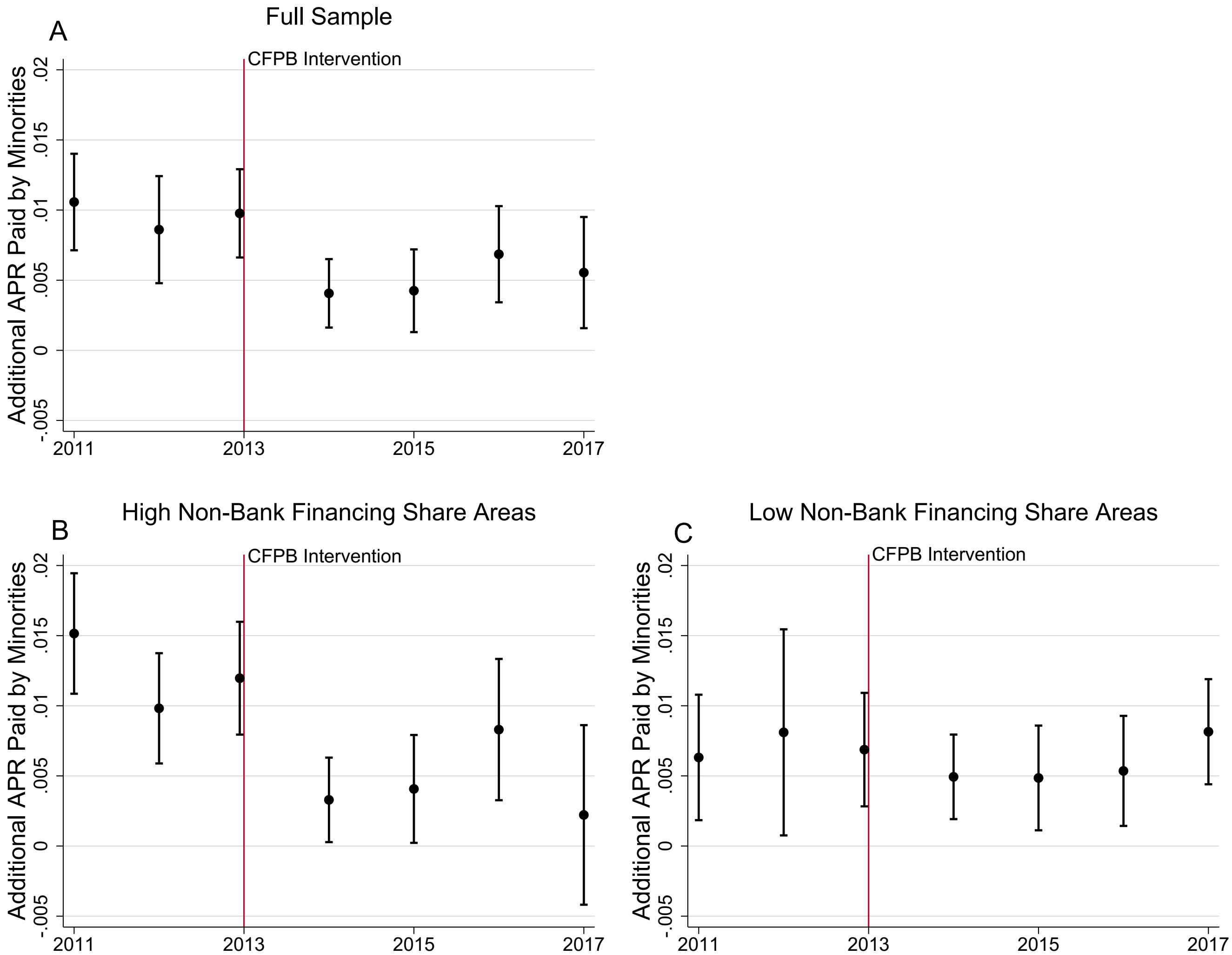

Figure 3
The 2013 CFPB Intervention and Racial Disparities in Auto Loan Interest Rates
This figure shows estimates of the additional interest that minorities pay on auto loans each year from 2011-2017. The top plot shows the estimates for the full sample, and the bottom left (right) plot shows estimates for minorities living in areas where a high (low) share of loans are financed by non-bank lenders. Each set of point estimates comes from a regression of interest rates on the full set of individual, loan, and ZIP code level controls, similar to the regression in Column 2 of Table 8, except that the *Minority* indicator is interacted with indicators for each year. The plots show these *Minority X Year* coefficient estimates and 90% confidence intervals. Over the course of 2013, the Consumer Financial Protection Bureau signaled to indirect auto lenders (primarily non-bank lenders) that it would increase its efforts to hold them accountable for discrimination in the interest rates they charge. The CFPB signaled this intent with a Bulletin in March of 2013, and especially with its first major enforcement action against a large indirect auto lender (Ally Financial) in December of 2013. The vertical line in the plots denotes the cutoff between the pre (2011-2013) and post (2014-2017) periods we use to examine the effect of heightened CFPB scrutiny on lending discrimination.

Internet Appendix

Appendix A — Supplementary Tables and Figures

Table A.1: Number of Auto Loan Applications in Failed Credit Searches

This table presents statistics on the number of auto loan applications filed by people whose credit search failed. Specifically, we tabulate the number of "hard" auto credit inquiries (which occur when a lender checks the applicant's credit score) for person-years where the person had at least one inquiry but did not open a new auto loan. We winsorize the number of applications at 7 (the 99th percentile) when computing means. In the full sample and the prime/subprime subsamples, difference-in-means tests show that minorities file significantly more applications in failed searches (in each case, p-value < 0.01).

	Full Sample		Subprime Borrowers		Prime Borrowers	
	White (N=26,918)	Minority (N=9,864)	White (N=11,563)	Minority (N=6,147)	White (N=14,862)	Minority (N=3,400)
Number of Apps						
Mean	1.61	1.79	1.83	1.90	1.45	1.58
1	65.58%	60.35%	59.61%	57.41%	70.19%	65.62%
2	20.77%	20.89%	20.90%	20.82%	20.68%	21.26%
3	7.01%	9.03%	8.80%	10.02%	5.63%	7.38%
4	3.18%	4.30%	4.67%	4.88%	2.02%	3.18%
5 or more	3.46%	5.43%	6.02%	6.87%	1.49%	2.56%

Table A.2: Race and Auto Credit Approval – Subsample Tests

This table presents regressions of auto credit access (an indicator for the person successfully opening a new auto loan) on applicant race, individual characteristics, and ZIP code characteristics (like the tests in Table 4). The sample in Column 1 includes all person-years where the individual applies for an auto loan during the year. Column 2 restricts the sample to college-educated applicants (those with a record of student loan debt on their credit report at some point during our sample). Column 3 examines high income applicants (those with income above the sample median of $58,000). Column 4 examines middle-aged applicants (ages 40 to 64). The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of credit approval). The standard errors are clustered by state-year.

	Full Sample	College-Educated	High Income	Middle-Aged
	Credit Approval (Auto) (1)	Credit Approval (Auto) (2)	Credit Approval (Auto) (3)	Credit Approval (Auto) (4)
Minority	-1.480^{***}	-1.711^{***}	-0.688^{**}	-1.064^{***}
	(0.259)	(0.338)	(0.301)	(0.352)
Individual Controls	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes
R-Squared	0.085	0.088	0.073	0.073
Observations	214,534	90,287	119,353	102,374

Table A.3: Race and Auto Credit Approval – Post Financial Crisis Sample

This table repeats the tests shown in Table 4, except on a post financial crisis sample (2011-2017). The tests regress a measure of auto credit approval on race, individual characteristics, and ZIP code characteristics. The outcome variable is an indicator for the person successfully opening a new auto loan, and the sample includes all person-years in which individuals apply for auto loans. Columns 4 and 5 restrict the sample to applicants with subprime, and prime credit scores, respectively. The individual level data consist of credit bureau records that have been matched to Home Mortgage Disclosure Act records (see Section 3.3 for details). The coefficients are reported in terms of percentage points (i.e., a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of credit approval). The standard errors are clustered by state-year.

	Full Sample			Subprime Borrowers	Prime Borrowers
	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)
	(1)	(2)	(3)	(4)	(5)
Demographics					
Minority	-3.868^{***}	-1.445^{***}	-1.857^{***}	-2.364^{***}	-0.852^{***}
	(0.338)	(0.319)	(0.409)	(0.553)	(0.316)
Minority X Hispanic			0.740		
			(0.508)		
Female	1.074^{***}	0.784^{***}	0.807^{***}	0.889^{*}	0.932^{***}
	(0.225)	(0.217)	(0.217)	(0.462)	(0.220)
Age	0.006	-0.062^{***}	-0.061^{***}	-0.007	-0.063^{***}
	(0.010)	(0.009)	(0.009)	(0.020)	(0.010)
Log(Income)	3.400^{***}	1.634^{***}	1.649^{***}	4.704^{***}	0.629^{**}
	(0.211)	(0.215)	(0.216)	(0.508)	(0.253)
Credit Characteristics					
Credit Score $_{t-1}$		0.049^{***}	0.049^{***}	0.156^{***}	0.009^{***}
		(0.002)	(0.002)	(0.006)	(0.003)
Log(Total Debt $_{t-1}$)		0.841^{***}	0.842^{***}	0.302^{***}	0.951^{***}
		(0.071)	(0.071)	(0.095)	(0.102)
Debt to Income $_{t-1}$		0.020	0.019	0.242	-0.213^{**}
		(0.081)	(0.081)	(0.175)	(0.097)
Log(Past Due Debt $_{t-1}$)		-1.072^{***}	-1.070^{***}	-0.681^{***}	-0.408^{***}
		(0.060)	(0.060)	(0.074)	(0.080)
ZIP Code Characteristics					
Log(Personal Income Per Capita)	0.124	-0.780	-0.734	0.022	-1.105
	(0.758)	(0.744)	(0.743)	(1.432)	(0.816)
Log(Population Density)	-0.017	-0.041	-0.039	-0.141	0.057
	(0.082)	(0.080)	(0.081)	(0.190)	(0.091)
Bachelors Degree	5.267^{***}	2.015	1.933	4.309	2.551
	(1.485)	(1.470)	(1.471)	(3.086)	(1.584)
Commute Using Car	10.382^{***}	10.100^{***}	10.019^{***}	11.637^{***}	8.654^{***}
	(1.483)	(1.480)	(1.482)	(3.412)	(1.645)
State-by-Year FE	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes
R-Squared	0.029	0.057	0.057	0.079	0.030
Observations	132,113	130,867	130,867	38,068	92,796

Table A.4: Race and Auto Credit Approval – ZIP Code Fixed Effects

This table repeats the tests shown in Table 4, except that the tests here use ZIP code fixed effects rather than ZIP code control variables. The tests regress a measure of auto credit approval on race and individual credit characteristics. The outcome variable is an indicator for the person successfully opening a new auto loan, and the sample includes all person-years in which individuals apply for auto loans. Columns 4 and 5 restrict the sample to applicants with subprime, and prime credit scores, respectively. The individual level data consist of credit bureau records that have been matched to Home Mortgage Disclosure Act records (see Section 3.3 for details). The dataset includes credit bureau records for the years 2005-2017. The coefficients are reported in terms of percentage points (i.e., a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of credit approval). The standard errors are clustered by state-year.

	Full Sample			Subprime Borrowers	Prime Borrowers
	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)
	(1)	(2)	(3)	(4)	(5)
Demographics					
Minority	-3.965^{***}	-1.269^{***}	-1.546^{***}	-2.180^{***}	-0.617^{*}
	(0.313)	(0.300)	(0.373)	(0.513)	(0.330)
Minority X Hispanic			0.492		
			(0.472)		
Female	1.698^{***}	1.210^{***}	1.223^{***}	1.522^{***}	1.220^{***}
	(0.185)	(0.183)	(0.183)	(0.413)	(0.205)
Age	0.045^{***}	-0.060^{***}	-0.060^{***}	0.038^{**}	-0.073^{***}
	(0.009)	(0.009)	(0.009)	(0.019)	(0.010)
Log(Income)	3.963^{***}	1.991^{***}	2.003^{***}	4.778^{***}	0.966^{***}
	(0.193)	(0.194)	(0.195)	(0.471)	(0.228)
Credit Characteristics					
Credit Score $_{t-1}$		0.056^{***}	0.056^{***}	0.154^{***}	0.014^{***}
		(0.002)	(0.002)	(0.005)	(0.003)
Log(Total Debt $_{t-1}$)		0.758^{***}	0.757^{***}	0.245^{***}	0.789^{***}
		(0.056)	(0.056)	(0.073)	(0.083)
Debt to Income $_{t-1}$		-0.013	-0.013	-0.080	-0.187^{**}
		(0.066)	(0.066)	(0.136)	(0.083)
Log(Past Due Debt $_{t-1}$)		-1.148^{***}	-1.147^{***}	-0.755^{***}	-0.393^{***}
		(0.053)	(0.053)	(0.068)	(0.071)
ZIP Code FE	Yes	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes
R-Squared	0.116	0.147	0.147	0.232	0.130
Observations	216,376	212,596	212,596	65,950	143,879

Table A.5: Minorities' Estimated Reduction in Auto Loan Approval by State

This table presents in Column 1 our estimates of the reduction in auto loan approval rates that minorities face in each of the 50 states and the District of Columbia. The estimates come from a regression of auto loan approval on controls, similar to the regression reported in Column 2 of Table 4, except that the *Minority* indicator is interacted with indicators for each state and the District of Columbia. We require that our sample contains at least 25 minority applications in a state in order to report the $State_i$ *X Minority* coefficient estimate (excludes 6 states with small minority populations). In Column 2 we report a measure of the prevalence of racial biases in each state (*Racial Slur GSV*), which is based on Google Search Volume for racial slurs, following Stephens-Davidowitz (2014). The state-level search volume data are normalized by Google so that the state with the highest proportion of searches fitting the criteria has a search volume of 100. Google computes search volumes based on a fraction of all Google searches, so we collect 50 draws of the data and assign each state its average search volume (we find very little variation across draws). For reference, Columns 3 and 4 report the share of minorities in our sample of applicants, and in the overall population, for each state.

State	Estimated Reduction in Auto Loan Approval (%) (1)	Racial Slur GSV (2)	Minority Share of Loan Applicants (%) (3)	Minority Share of State Population (%) (4)
Delaware	-7.85	62.6	22.1	30.1
Wyoming	-7.85	59.5	9.6	10.0
Rhode Island	-7.43	54.6	8.7	18.4
Tennessee	-6.09	75.5	11.9	21.8
Indiana	-5.60	69.0	8.6	15.8
Idaho	-5.59	45.0	7.2	12.1
Pennsylvania	-4.69	76.5	11.0	16.9
Oklahoma	-4.06	72.0	13.5	17.1
Colorado	-4.00	47.0	15.0	25.0
Arkansas	-3.99	77.5	14.5	22.3
Kansas	-3.90	59.0	8.1	17.1
South Carolina	-3.61	76.5	16.9	33.5
Illinois	-3.59	67.5	19.6	30.7
Connecticut	-3.56	58.0	14.3	23.5
West Virginia	-3.42	100.0	2.8	5.2
Nevada	-3.26	61.0	27.8	35.1
Ohio	-2.62	74.0	8.9	16.1
Michigan	-2.60	73.0	8.7	19.2
Georgia	-2.50	74.5	28.6	39.7
Washington	-2.30	52.5	9.8	15.5
Louisiana	-2.22	89.9	19.8	36.7
Arizona	-2.13	55.5	21.2	34.0
Kentucky	-2.00	84.5	10.2	11.5
Massachusetts	-2.00	53.5	12.6	16.3
Texas	-1.66	62.0	32.6	49.6
North Carolina	-1.20	68.5	19.4	30.4
New Jersey	-0.97	63.0	17.9	31.1
Florida	-0.95	61.5	28.1	38.3
New Hampshire	-0.78	52.6	4.1	4.2
Alabama	-0.70	80.0	20.0	30.4
Oregon	-0.63	40.0	7.3	14.0
Nebraska	-0.59	53.0	6.9	14.3
Mississippi	-0.47	91.1	24.8	40.1
New York	-0.41	59.0	14.5	32.7
Virginia	-0.36	51.0	23.8	27.9
Maryland	-0.06	61.0	36.7	38.2
California	0.05	48.0	33.2	44.0
Wisconsin	0.55	59.0	8.4	12.7
District of Columbia	0.88	61.0	51.5	60.1
Iowa	1.75	57.5	3.9	8.5
Minnesota	1.85	52.0	5.8	10.6
Utah	2.38	36.0	8.1	14.3
Missouri	3.18	73.5	9.4	15.8
Hawaii	3.46	35.1	16.9	10.9
New Mexico	3.46	49.0	34.2	48.4
Vermont	N/A	60.1	0.8	2.8
North Dakota	N/A	56.4	0.5	3.5
South Dakota	N/A	53.5	2.1	4.4
Maine	N/A	52.6	0.7	2.8
Alaska	N/A	61.5	5.2	9.5
Montana	N/A	52.5	1.5	3.6

Table A.6: Race and Auto Credit Approval – Alternate Definitions of Local Racial Bias

This table repeats the cross-sectional test focusing on local racial bias from Table 5, except using alternate definitions of local racial bias. Columns 1 and 2 replicate the results from Table 5. Column 3 interacts *Minority* with a continuous and standardized version of *Racial Slur GSV* (instead of using the indicator for being in the top tercile of *Racial Slur GSV*). Columns 4 and 5 use an alternate measure of states' racial bias—the bias index from Levine, Levkov, and Rubinstein (2014) based on interracial marriage rates. Column 4 uses an indicator for being in a state in the top tercile of the bias index, and Column 5 uses a continuous and standardized version of the index. The tests in this table are otherwise similar to those in Tables 4 and 5. The outcome variable is an indicator for the person successfully opening a new auto loan, and the sample includes all person-years in which individuals apply for auto loans. The individual level data are from the matched dataset of credit bureau records and Home Mortgage Disclosure Act records (see Section 3.3 for details). The dataset includes credit bureau records for the years 2005-2017. The included individual controls and ZIP code controls are the same as those reported in Table 4. The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the probability of credit approval). The standard errors are clustered by state-year.

	Credit Approval (Auto) (1)	Credit Approval (Auto) (2)	Credit Approval (Auto) (3)	Credit Approval (Auto) (4)	Credit Approval (Auto) (5)
Minority	-1.480^{***}	-0.906^{***}	-1.555^{***}	-1.326^{***}	-1.523^{***}
	(0.216)	(0.254)	(0.217)	(0.230)	(0.217)
Minority X High Racial Bias State (GSV)		-1.910^{***}			
		(0.443)			
Minority X Racial Slur GSV			-0.954^{***}		
			(0.215)		
Minority X High Racial Bias State (Index)				-1.216^{**}	
				(0.617)	
Minority X Racial Bias Index					-0.563^{***}
					(0.210)
Individual Controls	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes
R-Squared	0.085	0.085	0.085	0.085	0.085
Observations	214,534	214,534	214,534	214,534	214,534

Table A.7: Borrower Race and Auto Loan Interest Rates – Quantile Regressions

The quantile regressions in this table estimate the effect of borrower race at the 75th percentile of auto loan interest rates. The sample of auto loans is constructed from the Credit Bureau/HMDA Matched Panel, and has an observation for each new auto loan originated from 2011-2017 (the time period over which interest rates are available). We require the loan to be the borrower's only auto loan at origination, so that loan characteristics can be accurately measured. To make our sample more similar to Charles, Hurst, and Stephens (2008), these tests focus on only White and Black borrowers. The coefficients are reported in terms of percentage points (i.e. a coefficient of one indicates that a unit increase in the explanatory variable predicts a one percentage point increase in the interest rate).

	Rate 75th Percentile	Rate 75th Percentile	Rate 75th Percentile
	(1)	(2)	(3)
Demographics			
Black	1.004^{***}	1.387^{***}	2.785^{***}
	(0.164)	(0.183)	(0.234)
Female	-0.281^{***}	-0.307^{***}	-0.515^{***}
	(0.063)	(0.056)	(0.087)
Age	0.012^{***}	-0.006^{**}	-0.013^{***}
	(0.002)	(0.002)	(0.003)
Log(Income)	0.594^{***}	0.667^{***}	0.114
	(0.149)	(0.159)	(0.169)
Auto Loan Characteristics			
Auto Loan Term Indicators	Yes	Yes	Yes
Log(Auto Loan Amount)	-2.622^{***}	-2.870^{***}	-3.118^{***}
	(0.157)	(0.157)	(0.162)
Auto Loan to Income Ratio	0.739^{**}	0.550^{**}	0.488
	(0.300)	(0.269)	(0.367)
Credit Characteristics			
Credit Score $_{t-1}$	-0.019^{***}		
	(0.000)		
Log(Total Debt $_{t-1}$)	-0.186^{***}	-0.310^{***}	
	(0.027)	(0.023)	
Debt to Income $_{t-1}$	-0.000	0.144^{***}	
	(0.029)	(0.022)	
Log(Past Due Debt $_{t-1}$)	0.598^{***}	0.943^{***}	
	(0.030)	(0.022)	
Auto Debt Share	0.647^{***}	0.648^{***}	
	(0.103)	(0.164)	
ZIP Code Characteristics			
Log(Personal Income Per Capita)	0.545^{**}	0.617^{***}	0.431^{***}
	(0.217)	(0.129)	(0.155)
Log(Population Density)	0.020	0.018	-0.017
	(0.021)	(0.018)	(0.034)
Bachelors Degree	-1.360^{***}	-2.234^{***}	-2.608^{***}
	(0.409)	(0.342)	(0.348)
Commute Using Car	-0.826	-0.556	-0.974^{**}
	(0.528)	(0.654)	(0.383)
Origination Month Indicators	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes
Pseudo R-Squared	0.308	0.273	0.177
Observations	22,850	22,850	22,850

Table A.8: Local Racial Bias Predicts Higher Interest Rates for Minorities, but Not Higher Defaults

The regressions in this table test whether local racial biases affect the racial differences in interest rates and default rates. As in the default tests in Table 9, the sample includes one observation for each loan in our data originated between 2011 and 2015 (the period over which we can compute both interest rates and our indicator for default). The auto loans are required to be originated after the match between the credit bureau and HMDA records, and the loan must be the borrower's only outstanding auto loan at origination. In Column 1, the outcome variable is the interest rate (see Table 8, Column 3 for a version of this test on the less restricted 2011-2017 sample). In Column 2, the outcome variable is an indicator for whether the borrower became 90 or more days delinquent on the loan during the year of origination or the following two calendar years. This test is similar to the default tests in Table 9, except that we intentionally omit the interest rate from the controls, since we are testing whether racial biases are correlated with racial differences in creditworthiness, rather than assessing loan profitability. The coefficients are reported in terms of percentage points, and the standard errors are clustered by state-year.

	Auto Loan Rate	Auto Loan Default
	(1)	(2)
Minority	0.312^{**}	-0.136
	(0.129)	(0.484)
Minority X High Racial Bias State	0.768^{***}	0.477
	(0.176)	(0.885)
Individual Controls	Yes	Yes
Auto Loan Characteristics (Excluding Int. Rate)	Yes	Yes
ZIP Code Controls	Yes	Yes
Origination Month Indicators	Yes	Yes
State-by-Year FE	Yes	Yes
Time Relative to Match Indicators	Yes	Yes
R-Squared	0.388	0.080
Observations	10,509	10,509

Table A.9: Auto Loan Size and Racial Disparities in Interest Rates

The regressions in this table examine the effect of borrower race on auto loan interest rates for loans of various sizes. The sample of auto loans is constructed from the Credit Bureau/HMDA Matched Panel and contains one observation for each new auto loan originated from 2011-2017 (the time period over which interest rates are available). We also require the loan to be the borrower's only outstanding auto loan at origination. The tests in Columns 1-4 of Panel A show the effect of belonging to a racial minority (*Minority*) on interest rates for the subsample of loans in each quartile (Q1-Q4) of the loan amount distribution. The average loan amount for loans in the quartile is listed for reference. Panels B and C present similar tests for subprime and prime borrowers respectively. The loan amount quartile assignments for the tests in Panels B and C are based on the full sample, as in Panel A. The *Minority* coefficient is reported in terms of percentage points, and the standard errors are clustered by state-year.

Panel A: All Borrowers

	Loan Amount Q1 (Mean Amount = $9,837)	Loan Amount Q2 (Mean Amount = $16,973)	Loan Amount Q3 (Mean Amount = $23,072)	Loan Amount Q4 (Mean Amount = $35,159)
	Rate	Rate	Rate	Rate
	(1)	(2)	(3)	(4)
Minority	1.186***	0.743***	0.694***	0.377***
	(0.226)	(0.210)	(0.121)	(0.110)
Individual Controls	Yes	Yes	Yes	Yes
Auto Loan Characteristics	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes
Origination Month Indicators	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes
R-Squared	0.451	0.496	0.479	0.422
Observations	6,361	6,363	6,362	6,361

Panel B: Subprime Borrowers

	Loan Amount Q1	Loan Amount Q2	Loan Amount Q3	Loan Amount Q4
	Rate	Rate	Rate	Rate
	(1)	(2)	(3)	(4)
Minority	2.246***	1.216***	1.711***	0.854***
	(0.502)	(0.416)	(0.308)	(0.272)
Individual Controls	Yes	Yes	Yes	Yes
Auto Loan Characteristics	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes
Origination Month Indicators	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes
R-Squared	0.430	0.484	0.515	0.471
Observations	1,701	1,519	1,398	1,243

Panel C: Prime Borrowers

	Loan Amount Q1	Loan Amount Q2	Loan Amount Q3	Loan Amount Q4
	Rate	Rate	Rate	Rate
	(1)	(2)	(3)	(4)
Minority	0.537**	0.472***	0.309***	0.176**
	(0.214)	(0.175)	(0.078)	(0.083)
Individual Controls	Yes	Yes	Yes	Yes
Auto Loan Characteristics	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes
Origination Month Indicators	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes
R-Squared	0.271	0.301	0.337	0.361
Observations	4,596	4,790	4,905	5,062

Table A.10: Race and Auto Loan Interest Rates - Subsample Tests

The regressions in this table examine the effect of borrower race on auto loan interest rates for various subsamples. The sample of auto loans is constructed from the Credit Bureau/HMDA Matched Panel and contains one observation for each new auto loan originated from 2011-2017 (the time period over which interest rates are available). We also require the loan to be the borrower's only outstanding auto loan at origination. Column 1 presents the results using the full sample, as in Table 8. Column 2 restricts the sample to college-educated borrowers (those with a record of student loan debt on their credit report at some point during our sample). Column 3 examines high income borrowers (those with income above the sample median of $58,000). Column 4 examines middle-aged borrowers (ages 40 to 64). The *Minority* coefficient is reported in terms of percentage points, and the standard errors are clustered by state-year.

	Full Sample	College-Educated	High Income	Middle-Aged
	Rate	Rate	Rate	Rate
	(1)	(2)	(3)	(4)
Minority	0.704***	0.713***	0.456***	0.635***
	(0.117)	(0.158)	(0.103)	(0.132)
Individual Controls	Yes	Yes	Yes	Yes
Auto Loan Characteristics	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes
Origination Month Indicators	Yes	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes
R-Squared	0.440	0.467	0.431	0.449
Observations	25,523	10,466	12,683	12,368

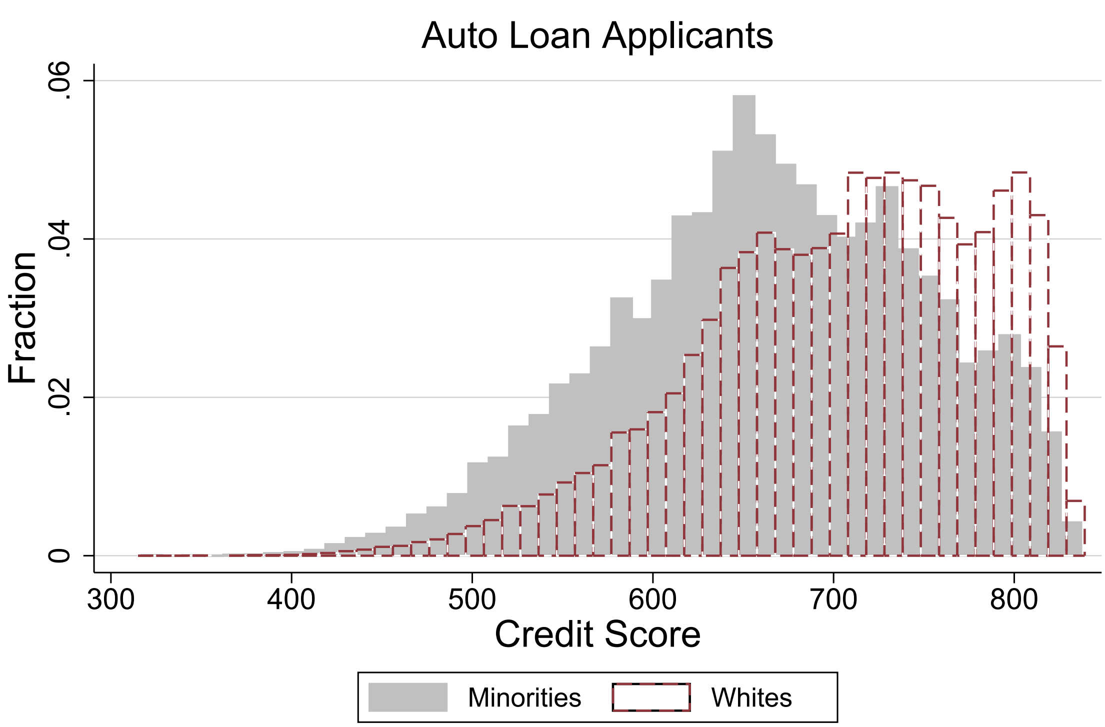

Figure A.1
Credit Score Distributions for Minority and White Applicants
This figure presents the credit score distributions for minority and White auto loan applicants in our Credit Bureau/HMDA Matched Panel. The dataset includes credit bureau records for the years 2005-2017 (see Section 3.3 for details).

Appendix B — Back-of-the-Envelope Calculations

In this Appendix, we use two approaches to estimate the total number of minority auto loan applicants who fail to secure loans each year, that they would have received if they were White (*MinoritiesDeniedPerYear*). In each approach, we estimate this number by multiplying an estimate of the total number of minorities applying for auto loans per year (*YearlyMinorityApps*) by the reduction in their probability of approval due to discrimination. This reduction in credit approval rates is already estimated by the coefficient on *Minority* in Column 2 of Table 4, which we refer to as *MinorityCoefficient*. The two approaches differ only in how they estimate *YearlyMinorityApps*. The first approach is simple and naive, whereas the second approach is data-driven and produces the estimates we reference in the paper. Below we describe the two approaches, and how they may over or underestimate *MinoritiesDeniedPerYear*.

B.1 Naive Estimate of the Number of Applicants Denied Credit Each Year Due to Discrimination

In this approach, we take the average number of borrowers applying for auto loans each year in our 1% sample of credit bureau data, and multiply it by 100 to estimate the number of U.S. residents with a credit history that apply for auto loans each year. We then make the naive assumption that Black and Hispanic borrowers apply for auto loans exactly as often as other borrowers. Using this assumption, we estimate *YearlyMinorityApps* by multiplying the number of auto loan applications per year by the fraction of the U.S. population that is Black and/or Hispanic (approximately 29% according to the 2010 Census). We then obtain an estimate of the number of minority applicants denied auto loans each year due to discrimination, by multiplying *YearlyMinorityApps* by the 1.5 percentage point *MinorityCoefficient* from Table 4.

On average, there are 338,972 borrowers applying for auto loans each year in our credit bureau data. Therefore,

Estimate of $\textit{YearlyMinorityApps} = 338,972 \times 100 \times 0.29 = 9,830,188$

Estimate of *MinoritiesDeniedPerYear* $= 9,830,188 \times 0.015 =$ **147,453**

B.2 Data-Driven Estimate of the Number of Applicants Denied Credit Each Year Due to Discrimination

B.2.1 Estimate the Number of Minority Auto Loan Applicants Per Year

First, note that we only observe auto loan applicants' race in our final dataset, the Credit Bureau/HMDA Matched Panel. Therefore, we need to walk through the filtering process that determines which auto loan applications end up in our final dataset. Understanding the filters allows us to estimate the percentage of all auto loan applications by minorities in the United States that end up in our final dataset (call this fraction F_{Final}).

Let us consider the filtering process for a randomly selected minority borrower-year from 2005-2017 during which the borrower applied for auto credit (call this borrower-year *TargetApp*). To make it into our final dataset, *TargetApp* must make it through three sequential filters: making it into our 1% credit bureau sample, belonging to a borrower who is a candidate to be matched to the HMDA data, and being successfully matched to the HMDA data. We refer to the probabilities that *TargetApp* makes it through these three filters as $F_{CreditBureau}$, $F_{MatchCandidate}$, and $F_{Matched}$, respectively. Therefore, the probability that *TargetApp* makes it into our matched dataset is:

$F_{Final} = F_{CreditBureau} \times F_{MatchCandidate} \times F_{Matched}$.

Filter 1: Credit Bureau Sample

The probability that *TargetApp* appears in our credit bureau sample ($F_{CreditBureau}$) should be 1%, because these data are a 1% sample of all U.S. Residents with a credit history and Social Security number.

Filter 2: Must Belong to a Candidate for the Match to HMDA

In order to be a candidate for the match to HMDA, the borrower from *TargetApp* must take out a mortgage between 2010 and 2016, and the mortgage must fit the following requirements:

1) Must be borrower's only first lien mortgage at the time of origination.

2) Person must live in an MSA directly following the mortgage origination.

3) Person must be the only applicant on the mortgage loan.

Fortunately, because we have the 1% sample of credit bureau data, we can calculate the probability that a randomly selected borrower-year during which the borrower applies for auto credit, belongs to a borrower who takes out this type of mortgage between 2010 and 2016. Using the credit bureau data, we calculate this probability (based on all auto loan applicants) to be 8.77%, which we use as our estimate of $F_{MatchCandidate}$ (the probability for minority applicants).

It is important to note that this approach assumes that minority auto loan applicants are just as likely as White applicants to take out a home purchase or refinance loan on their own (no co-applicant), for their primary residence located in an MSA. Based on our summary statistics showing that, even within the matched sample of homeowners, minorities have lower credit scores on average, we would expect minority auto loan applicants to be less likely to become this type of homeowner than White applicants. Therefore, $F_{MatchCandidate}$ likely overstates the probability that the minority borrower from *TargetApp* is a candidate for the match to HMDA. This overstatement of $F_{MatchCandidate}$ would bias our estimate of F_{Final} upwards, which would in turn bias our final estimate of the total number of minority applicants denied credit downwards (making it conservative).

Filter 3: Candidate Must be Successfully Matched to HMDA

For the borrower from *TargetApp* to be in the final matched dataset, a mortgage they take out fitting the match criteria must actually be successfully matched to HMDA. The probability of a credit bureau mortgage that fits the match criteria being successfully matched to HMDA is calculated in the summary statistics describing the match in Table 1, and is 68.82%. This approach assumes that minorities' mortgages are just as likely to be matched as White borrowers', and this assumption is supported by the results in Table 2 showing that race does not affect the likelihood of being matched. Therefore, 68.82% should be an accurate estimate of $F_{Matched}$.

Estimate *YearlyMinorityApps*

Based on the filters described above, the probability that *TargetApp* makes it into our final matched dataset is:

$$\begin{aligned} F_{Final} &= F_{CreditBureau} \cdot F_{MatchCandidate} \cdot F_{Matched} \\ &= 0.01 \times 0.0877 \times 0.6882 \\ &= 0.0006036 \end{aligned}$$

Therefore, we can estimate the total number of minority auto loan applications per year as the number of them in our sample per year, multiplied by $1/F_{Final}$. Based on the summary statistics in Table 3, our sample contains 42,565 minority auto loan applicant-years from 2005-2017, i.e. 3,274 applications per year. Therefore,

Estimate of *YearlyMinorityApps* $= \frac{3{,}274}{0.0006036} = 5,424,122$

B.2.2 Calculate the Final Estimate

We use the data-driven estimate of the number of minorities applying for auto credit each year, and the reduction in loan approval rates that minorities face, to estimate the number of minority applicants denied auto credit each year due to discrimination.

Estimate of *MinoritiesDeniedPerYear* $= 5,424,122 \times 0.015 =$ **81,362**

It is important to note that we are assuming that *MinorityCoefficient* is based on a representative sample of minority auto loan applicants. However, the sample of applicants from our matched dataset are homeowners (or soon-to-be homeowners), and are likely of higher credit quality than the average minority auto loan applicant. Because we find evidence that lower credit quality borrowers face stronger discrimination, this suggests that our estimate of *MinorityCoefficient* likely understates the true effect for the population of minority auto loan applicants. Therefore, our estimate of the total number of minorities denied credit due to discrimination each year is likely conservative.