Discrimination in the Auto Loan Market

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

June 25, 2019

Corresponding author email address: emayer@smu.edu

Acknowledgements: For helpful comments we thank seminar participants at Rice University, Southern Methodist University, Texas Christian University, and University at Buffalo. 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.

Discrimination in the Auto Loan Market

Abstract

We provide evidence of discrimination in the auto loan market. Combining credit bureau records with borrower characteristics, we find that Black and Hispanic applicants' loan approval rates are 1.5 percentage points lower than White applicants', even controlling for creditworthiness. In aggregate, this discrimination leads to over 80,000 minorities failing to secure loans each year. Results are stronger in more racially biased states and where banking competition is lower. Minorities who receive loans pay interest rates 70 basis points higher than comparable White borrowers. *Ceteris paribus*, minority borrowers have *lower* ex post default rates, consistent with preference-based racial discrimination. An anti-discrimination enforcement policy initiated in 2013, but halted in 2018, was effective in reducing unexplained racial disparities in interest rates by nearly 60%.

Over 100 million U.S. consumers had automobile debt in 2017, making auto loans the most widely used form of installment credit. Yet, compared to the markets for other consumer credit products like mortgages or student loans, the auto loan market is relatively unstructured, unregulated, and opaque. The lack of transparency makes it harder to monitor the factors lenders consider, potentially including characteristics like race and ethnicity. Indeed, the Consumer Financial Protection Bureau (CFPB) issued specific guidance to auto lenders in 2013 on how the Equal Credit Opportunity Act applies to auto loans.[1]

There are not many academic studies of discrimination in auto lending. Identifying discrimination requires information on applicant/borrower race and outcomes, but auto lenders are not required to report application or loan-level data.[2] Therefore, past studies of auto lending practices are largely suggestive or incomplete. Our study builds an extensive, novel, and rich dataset in order 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. Linking the two databases presents a challenge because they do not share a common identifier. However, information on originated mortgages is reported with sufficient granularity in each dataset that we can uniquely match the majority (69%) of mortgages in the credit bureau data to HMDA based on mortgage characteristics. The credit bureau records provide a panel data structure and information on financial outcomes including auto loans, while the HMDA data provide borrower demographics. Our matched dataset contains roughly 79,000 people per year

[1] The 2013 CFPB Bulletin can be found here: https://www.consumerfinance.gov/policy-compliance/guidance/implementation-guidance/bulletin-indirect-auto-lending-compliance/

[2] 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.

between 2005 and 2017. We use these data to test whether minorities face discrimination in the auto loan market, and find strong evidence that they do.

A number of prior studies interpret lower approval rates and higher interest rates for minorities as evidence of lending discrimination in other markets (e.g. Munnell et al. (1996) and Bayer, Ferreira, and Ross (2018)). However, racial disparities in access to credit could arise from taste-based discrimination (Becker (1957)), omitted variables, or statistical discrimination (Phelps (1972)). To identify taste-based discrimination, Becker (1957, 1993) proposes an "outcome test" that compares the profitability of loans to marginal White and minority borrowers. If lenders discriminate, loans to marginal minority borrowers should be more profitable because the bar is higher. Researchers typically use loan performance as a proxy for profitability, and lower ex post default rates for minorities are considered strong evidence of discrimination (Ferguson and Peters (1995)). We take advantage of the scope of our data and test for discrimination using all three approaches, evaluating differences in loan approvals, interest rates, and subsequent defaults.

Our first tests focus on loan approval rates. We use a broad set of controls including borrower characteristics (e.g., age, sex, income), ZIP code characteristics, and state-by-year fixed effects. Importantly, we control directly for applicants' *financial health* (credit score, debt, debt to income ratio, and debt past due). Few other lending discrimination studies have such a rich set of controls. We estimate that minority applicants have a lower approval rate by 1.5 percentage points, which is comparable to the effect of a 26 point (32% of a standard deviation) reduction in borrower credit score. The difference in approval rates is 60% larger (2.4 percentage points) for minority applicants with subprime credit scores, where qualitative preferences likely have greater influence. A 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. Because we find the strongest evidence of discrimination among lower credit quality applicants, and our sample consists of *homeowners,* who typically have better credit than the average auto loan applicant, our estimates may understate the true magnitude of discrimination.

A concern when testing for discrimination in credit markets is that race correlates with creditworthiness in some way that lenders observe, but researchers do not. If so, we should see racial disparities in credit approval, even absent any racist preferences. We argue that credit card applications provide an ideal setting for such a falsification test. Unlike auto loans, which typically involve personal interaction, most credit card decisions are made using statistical algorithms that provide less opportunity for direct discrimination (e.g. Gross and Souleles (2002), Moore (1996), and Tsosie (2016)). We find that, on average, the *same* minority applicant who faced lower approval rates on auto loans does not face lower approval rates on credit cards, *during the same year*. This finding suggests that the human element of auto lending, rather than actual differences in creditworthiness, leads to the lower approval rates for minorities.

Next, we examine the cross-sectional variation in discrimination. First, we test whether discrimination is stronger in states where racial biases are more prevalent. Following Stephens-Davidowitz (2014), we measure states' racial bias using Google Search Volume for racial slurs. We find that the effect of race 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 also test whether competition ameliorates discrimination. Whereas statistical discrimination is profitable for lenders and should persist despite competition, taste-based discrimination is costly and should be rooted out by competition (e.g., Buchak and Jørring (2017)). Consistent with taste-based discrimination, we find stronger results in low-competition environments.

Discrimination also affects an intensive margin of credit provision through higher interest rates for minorities. *Ceteris paribus*, minorities pay 70 basis points more on their auto loans (APR) than White borrowers. This magnitude is comparable to the effect of a 37 point drop in credit score. This result is especially notable because we find it in the sample of borrowers who were approved—at lower average approval rates—for the loans. Moreover, the effect of minority status increases to 125 basis points for borrowers in high racial bias states.

Some of these results could reflect an omitted variable bias if minorities are worse credit risks, even net of our extensive controls.[3] If true, it would lead to higher ex post default rates for minorities in our tests. *Ceteris paribus*, we find that minorities have a *lower* default rate in the full sample. In the subprime sample, default rates are a statistically significant 2.3 percentage points lower for minorities, consistent with loans to these marginal minority borrowers being more profitable than loans to marginal White borrowers. These results provide strong evidence that the racial disparities we document in

[3] Such an omitted variable bias would still have difficulties explaining the cross-sectional patterns in discrimination we find, and the results of our falsification test based on credit card applications.

credit approval and interest rates are generated by taste-based discrimination rather than omitted variable bias or statistical discrimination.

In our final set of tests, we evaluate whether increased oversight of auto lenders affects discrimination. We exploit a sharp increase in the CFPB's scrutiny of indirect auto lenders in 2013. Our differences-in-differences tests show that the additional interest (APR) paid by minorities decreased from 84 basis points, to 35 basis points in the post-event period (a 58% decrease). A triple differences test shows that the reduction in discrimination occurred primarily in areas where indirect auto lending is most prevalent, providing evidence that we are indeed capturing the effect of the CFPB's actions. These findings are particularly relevant considering that CFPB oversight is an area of active debate—in fact, in 2018 Congress passed a joint resolution nullifying the 2013 Bulletin the CFPB used to spearhead its anti-discrimination enforcement policies.

Our paper is related to prior work documenting racial disparities in approval rates for mortgages (e.g. Munnell et al. (1996)), credit cards (Cohen-Cole (2011)), and peer-to-peer loans (e.g. Pope and Sydnor (2011)).[4] Studies also show that minorities pay higher interest rates on mortgages (e.g. Bayer, Ferreira, and Ross (2018)). However, prior studies rarely include default rate tests (often due to data constraints), which makes inferences about discrimination precarious. For example, evidence from the mortgage market suggests that Black borrowers default more (e.g. Berkovec et al. (1998)), raising questions about whether racial disparities in approvals and interest rates reflect actual taste-based

[4] Also, see studies on the role of race in high-cost lending (Dobbie et al. (2018)), and small business lending (e.g., Blanchflower, Levine, and Zimmerman (2003) and Fairlie, Robb, and Robinson (2018)).

discrimination. 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 taste-based discrimination.

The primary contribution of our paper is to provide substantial evidence of lending 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 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 cannot 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 the first tests for taste-based discrimination in this market using ex post default rates. Each of our tests provides strong evidence that discrimination is prevalent in the U.S. auto loan market.

2. Background Information on Auto Lending

In this section we provide some general information about the U.S. auto loan market.[5] In 2017, 91% of U.S. Households had automobiles, and roughly 70% of auto purchases were used vehicles.[6] 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 have auto debt as of 2017, with aggregate balances over $1.1 trillion.

[5] Unless otherwise specified, auto lending statistics in this section come from an industry report, which can be found here: https://www.experian.com/assets/automotive/quarterly-webinars/2017-q4-safm.pdf.

[6] Household automobile ownership comes from the National Household Travel Survey. The composition of auto purchases comes from the fact that new vehicle purchases totaled 17.1 million in 2017 according to the Bureau of Economic Analysis, and used vehicle purchases totaled 39.2 million according to Edmunds, a leading automotive information provider.

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 ranged 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.), indirect auto lenders, and "buy here pay here" lenders. While banks usually interact directly with consumers (direct lending), indirect auto lenders partner with car dealerships to originate loans, and do not interact with the consumer.[8] Indirect auto lenders are either the "captive" financing arm of a major auto manufacturer (e.g. Ford Motor Credit Company) or an independent auto finance company. "Buy here pay here" lenders are typically used car dealerships that originate loans on-site. Based on 2017 originations, the market shares were banks (53.3%), indirect auto lenders (40.3%, mostly from captives), and "buy here pay here" lenders (6.4%).[9]

For auto loans financed indirectly, the consumer and lender never meet. Instead, the dealership helps the consumer obtain financing. For example, a typical scenario would

[7] The credit score mentioned is the Vantage Score, which we use throughout the paper. 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.
[9] "Buy here pay here" dealerships typically do not report their loans to the major consumer credit bureaus, which will prevent these loans from being included in our analyses.

begin with the customer choosing a car and then completing an application for credit. The dealer then submits this credit application to an indirect lender. The lender evaluates the application, *with no information on applicant race*. The lender decides on credit approval, and gives the dealer a minimum interest rate, not seen by the customer. 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.

At first, indirect lending may appear to limit concerns about discrimination, because lenders do not observe race in the credit application, or in person as a bank's loan officer would. However, to the contrary, regulators have expressed serious concerns about discrimination in indirect auto lending. The key point in the indirect lending process at which discrimination is most likely to occur is when the car dealership representative sets the dealer markup. At this point, the dealership representative has observed the consumer's race in person, and typically has an incentive to charge a high markup since indirect lenders share the profits with dealerships.

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 responsible under 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 7 we test the effect of the CFPB's increased scrutiny of auto lenders on our measures of discrimination.

3. Data and Methodology

We merge two data sources to connect individuals' 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.[11] Second, we use information on bank branch locations and deposits from the

[10] For a list of CFPB enforcement actions, see: https://www.consumerfinance.gov/policy-compliance/enforcement/actions/.
[11] See, https://trends.google.com/trends/?geo=US.

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 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.[12]

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

[12] We have also run our tests including borrowers matched based on joint mortgages, and the results are similar. However, we focus on single applicants because we know the HMDA information applies directly to the borrower. In contrast, for joint mortgages, we are unable to determine whether the borrower in the credit bureau data is the first applicant or the co-applicant in the HMDA data.

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 is matched to should be 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 credit bureau data. Since HMDA excludes some small mortgage lenders, and 89% of HMDA mortgages were unique based on the matching characteristics, we view the match rate of 69% as reasonable.

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.[13] 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

[13] It is more difficult to match people moving from one property to another, than people purchasing their first property. This fact is likely explained by seasoned homebuyers being more likely to change location, which could cause their location in the credit bureau data to be less likely to match the property location in the HMDA data.

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 in these regressions, indicating no evidence of selection bias either directly or through the combination of race and income.

[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, and are more likely to already have a mortgage in 2010.

[Insert Table 3]

4. Applicant Race and Auto Credit Approval

In this section, we test whether race affects access to auto credit. We start by selecting all borrower-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.[14] 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.[15] Panel B of Table 3 summarizes the characteristics of the auto loan applicants from 2005-2017.[16] 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

[14] 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 will be understated.
[15] See for example Bhutta and Keys (2016), Akey et al. (2018), Akey, Heimer, and Lewellen (2018), Brown, Cookson, and Heimer (2018), and Mayer (2018).
[16] Although borrowers are matched to HMDA based on mortgages originated from 2010-2016 (due to the data availability of matching variables), we can observe their auto loan applications in prior years as well. The sample starts in 2005 rather than 2004, because we need one prior year to construct lagged controls.

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)). 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 last result is particularly noteworthy because approval for subprime borrowers typically involves more loan officer discretion, lowering the marginal cost of discriminatory decision making.

[Insert Table 4]

Next, we use credit card application data from the credit bureau to conduct a simple falsification test (Column 6 of Table 4). Credit card approval decisions are generally made using quantitative algorithms. This automation reduces the human element, and mitigates the opportunity for taste-based discrimination. We examine credit card approval rates in the same borrower-years as the auto loan applications in our sample. If the lower minority approval rates on auto loan applications are justified by hard information available to lenders, but not to us as econometricians, then we should also observe lower minority approval rates on credit card applications. On the other hand, if there is no difference in credit card approval rates, it would suggest that the racial disparities in auto lending stem from the human component of the lending process. It appears that they do. We find that the same minority applicants with lower auto loan approval rates, do not face lower credit card approval rates, during the *same borrower-years*.

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 states where racial biases are more prevalent. 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.1.[17] In Column 1 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).[18]

[Insert Table 5]

[17] 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).
[18] We find similar results using the racial bias index from Levine, Levkov, and Rubinstein (2014), based on interracial marriage rates.

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.1). 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 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 2). This result is consistent with the gap in approval rates stemming from costly taste-based discrimination, which competition should limit (e.g., 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. In Column 3 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.[19] This finding suggests that the racial disparities in credit approval that we document are not driven by a particular type of lending institution. In a related vein, 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 very similar results (see Table A.2).

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 4 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).[20]

5. Borrower Race and Auto Loan Interest Rates

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

[19] 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.
[20] For matching purposes, we require borrowers to live in an MSA at the time of the credit bureau/HMDA match, but we make no requirements on where a borrower lives in other years.

loan to borrowers in the matched panel. We require the loan to be the borrower's only auto loan at origination so that the loan's characteristics can be accurately measured in the credit bureau data. We also require information on the borrower's scheduled monthly auto payment in order to compute the interest rate.[21] This information is not available from the credit bureau until 2011, and it is missing for 33% 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 6 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 6 Here]

We test whether minority auto borrowers pay higher interest rates by regressing each auto loan's annual percentage rate (APR) 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 7 show that minority borrowers pay interest rates 70 basis points higher than can be explained by observable characteristics.

[21] We compute an annual percentage rate (APR) based on the loan term, loan amount, and scheduled monthly payments, assuming a fixed rate loan.

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.

Two prior studies also examine racial disparities in auto loan interest rates. 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.[22] In an analysis closer to ours, 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.3). These findings suggest that even analyses that control for a set of credit history variables, but not credit scores, likely significantly overstate the effect of race.[23] Omitting credit history variables altogether (even controlling for age, sex,

[22] These confidential data were accessed as a plaintiff's expert and cannot be used for research purposes.
[23] 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

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 7 or see Table A.3.

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 7 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. 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 7, 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 where the borrower lives, however, the analysis in Section 7 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 7]

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,

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.

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.4 for these results). 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).

6. Race and Auto Loan Default Rates

In this section we implement a version of the outcome test proposed in Becker (1957, 1993), to test for discrimination in the auto loan market. Becker (1957, 1993) proposes that, in order to identify taste-based discrimination, researchers should test whether loans to marginal minority borrowers are more profitable than those to marginal White borrowers. The underlying intuition is that this test evaluates whether lenders (or more likely, individual employees of lenders) set the bar higher for minorities due to racial biases/preferences. To implement the outcome test, researchers examine loan performance conditional on loan and borrower characteristics, and lower ex post default rates for minorities are considered strong evidence of discrimination (Ferguson and Peters (1995)). Therefore, we test whether minority auto borrowers are more or less likely to default than comparable White borrowers.

For our default rate 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 the following two calendar years.[24] 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 8, 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. Subprime borrowers may be a more appropriate sample for the outcome test, as they are more likely to be the marginal borrowers. In the subprime sample (Column 2), minority status has a negative and statistically significant

[24] The credit bureau data only allow us to see detailed information on delinquency status as of when the credit file is extracted (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.

effect (2.3 percentage points) on the probability of default. This magnitude is 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. The results from these default rate tests provide evidence that minorities, especially those with subprime credit scores, face taste-based racial discrimination in the auto loan market.

[Insert Table 8]

The tests in this section are an important check on the evidence of discrimination presented in Sections 4 and 5. A serious concern in any study of lending discrimination is that minority status is correlated with an unobserved component of credit risk, creating an omitted variable bias that could skew credit approval and interest rate results in favor of discrimination. However, if this bias exists, it should also skew the default rate results against showing discrimination. Therefore, the default rate approach may be the most conservative. As such, these results may offer the clearest evidence of discrimination.

7. 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 9, respectively. The results in Column 1 show that the additional interest (APR) 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 rates 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.[25] Yet, these credit approval results show that pressure to avoid charging minorities disproportionately high dealer markups/rates did not reduce these borrowers' access to credit, as it might have if the rates were necessary to make these loans profitable. Overall, our tests suggest that the CFPB was effective in mitigating discrimination.

[Insert Table 9]

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

[25] The Consent Order against Ally Financial can be found here: https://files.consumerfinance.gov/f/201312_cfpb_consent-order_ally.pdf

most prevalent. Column 2 of Table 9 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.

In Column 5 of Table 9, 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 9, 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 sample of minorities living in areas with a high share of non-bank auto lending. 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. Panel B of Figure 3 shows no such drop in the additional interest paid by minorities in areas that were less affected by the CFPB's actions due to having a low share of non-bank auto lending. These results provide strong evidence that the CFPB's actions led to a reduction in discrimination by non-bank auto lenders.

[Insert Figure 3]

8. Conclusion

Our paper provides evidence of lending discrimination in the U.S. auto loan market. We find that Black and Hispanic applicants face lower credit approval rates than White applicants after controlling for credit score, income, and a broad set of personal, demographic, and geographic characteristics. The point estimates and a back-of-the-envelope calculation suggest that approximately 80,000 minority applicants fail to obtain auto loans each year due to discrimination.

The effect of minority status on credit approval is larger for applicants living in areas where racial biases are more prevalent, and where banking competition (which should limit discrimination) is less intense. Moreover, minority auto loan applicants, in the same borrower-years, do not face lower approval rates on the applications they submit to credit card lenders, who take the human element out of the lending process. These findings provide evidence that the racial disparities we document in auto loan approval rates stem from discrimination, rather than an omitted variables problem.

When they receive auto loans, minority borrowers (especially those in high racial bias states) pay higher interest rates than White borrowers, controlling for a broad set of borrower, loan, demographic, and geographic characteristics. Furthermore, controlling for these characteristics, we find that minority borrowers default *less*. The results from this outcome test (Becker (1957, 1993)) provide evidence that minorities face taste-based discrimination in the auto loan market.

We show that the CFPB's increased scrutiny of auto lenders starting in 2013 led to almost a 60% decrease in the additional interest that minorities pay on auto loans, with no concomitant decrease in 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

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*.

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.

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, 1998, Discrimination, Competition, and Loan Performance in FHA Mortgage Lending, *Review of Economics and Statistics* 80, 241–250.

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.

Brown, James R., J. Anthony Cookson, and Rawley Z. Heimer, 2018, Growing Up Without Finance, *Journal of Financial Economics* Forthcoming.

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

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.

Dobbie, Will, Andres Liberman, Daniel Paravisini, and Vikram Pathania, 2018, Measuring Bias in Consumer Lending, *Available at SSRN: https://ssrn.com/abstract=3258699*.

Fairlie, Robert, Alicia Robb, and David T. Robinson, 2018, Black and White: Access to Capital among Minority-Owned Startups, *Available at https://siepr.stanford.edu/research/publications/black-and-white-access-capital-among-minority-owned-startups*.

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

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.

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., 2018, Big Banks, Household Credit Access, and Economic Mobility, *Available at https://sites.google.com/site/erikjmayer/research*.

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.

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*.

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

This table presents regressions of measures of access to auto loans and credit cards on applicant race, individual characteristics, and ZIP code characteristics. The outcome variables are indicators for the borrower successfully opening a new auto loan (Columns 1-5) or a new credit card (Column 6). The sample in Columns 1-3 includes all person-years where the individual applies for an auto loan during the year. Columns 4 and 5 restrict the sample to applicants with subprime, and prime credit scores, respectively. The sample in Column 6 includes person-years where the individual applies for both auto credit and a credit card during the year. 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	Falsification Test:
	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Credit Card)
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Minority	−4.465***	−1.480***	−1.661***	−2.375***	−0.840***	0.147
	(0.289)	(0.259)	(0.332)	(0.399)	(0.271)	(0.368)
Minority X Hispanic			0.328			
			(0.410)			
Female	1.598***	1.115***	1.126***	1.492***	1.133***	3.069***
	(0.173)	(0.169)	(0.169)	(0.352)	(0.180)	(0.300)
Age	0.042***	−0.067***	−0.066***	0.017	−0.072***	0.055***
	(0.008)	(0.008)	(0.008)	(0.015)	(0.009)	(0.012)
Log(Income)	3.886***	1.704***	1.711***	4.586***	0.736***	−0.482
	(0.184)	(0.180)	(0.181)	(0.407)	(0.199)	(0.329)
Credit Characteristics						
Credit Score $_{t-1}$		0.057***	0.057***	0.161***	0.013***	0.055***
		(0.002)	(0.002)	(0.004)	(0.003)	(0.003)
Log(Total Debt $_{t-1}$)		0.866***	0.866***	0.403***	0.868***	0.131*
		(0.053)	(0.053)	(0.070)	(0.077)	(0.069)
Debt to Income $_{t-1}$		−0.032	−0.032	0.040	−0.220***	−0.299***
		(0.062)	(0.062)	(0.119)	(0.079)	(0.098)
Log(Past Due Debt $_{t-1}$)		−1.179***	−1.178***	−0.745***	−0.413***	−1.304***
		(0.051)	(0.051)	(0.061)	(0.066)	(0.062)
ZIP Code Characteristics						
Log(Personal Income Per Capita)	1.087*	−0.095	−0.076	0.573	−0.350	−0.301
	(0.629)	(0.611)	(0.611)	(1.088)	(0.701)	(1.060)
Log(Population Density)	−0.014	0.009	0.010	0.067	0.037	0.976***
	(0.065)	(0.065)	(0.065)	(0.142)	(0.072)	(0.120)
Bachelors Degree	5.108***	1.406	1.373	3.907	1.765	3.407
	(1.254)	(1.236)	(1.238)	(2.374)	(1.372)	(2.279)
Commute Using Car	12.020***	10.569***	10.533***	12.663***	8.640***	6.047***
	(1.194)	(1.149)	(1.146)	(2.317)	(1.276)	(2.213)
State-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.047	0.085	0.085	0.105	0.047	0.074
Observations	218,300	214,534	214,534	68,494	146,036	124,739

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)
Minority	-0.906^{***}	-1.268^{***}	-1.259^{***}	-1.509^{***}
	(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
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	0.085	0.085	0.085
Observations	214,534	214,534	214,534	214,534

Table 6: 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 APR	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 7: 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.

	APR (1)	APR (2)	APR (3)	APR (4)	APR (5)	APR (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 8: 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 APR	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 9: 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 APR, or in the probability of credit approval). The standard errors are clustered by state-year.

	Outcome Var = Auto Loan APR			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			
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origination Month Indicators	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	0.398	0.398	0.399	0.057	0.057	0.057
Observations	25,523	25,523	25,523	130,867	130,867	130,867

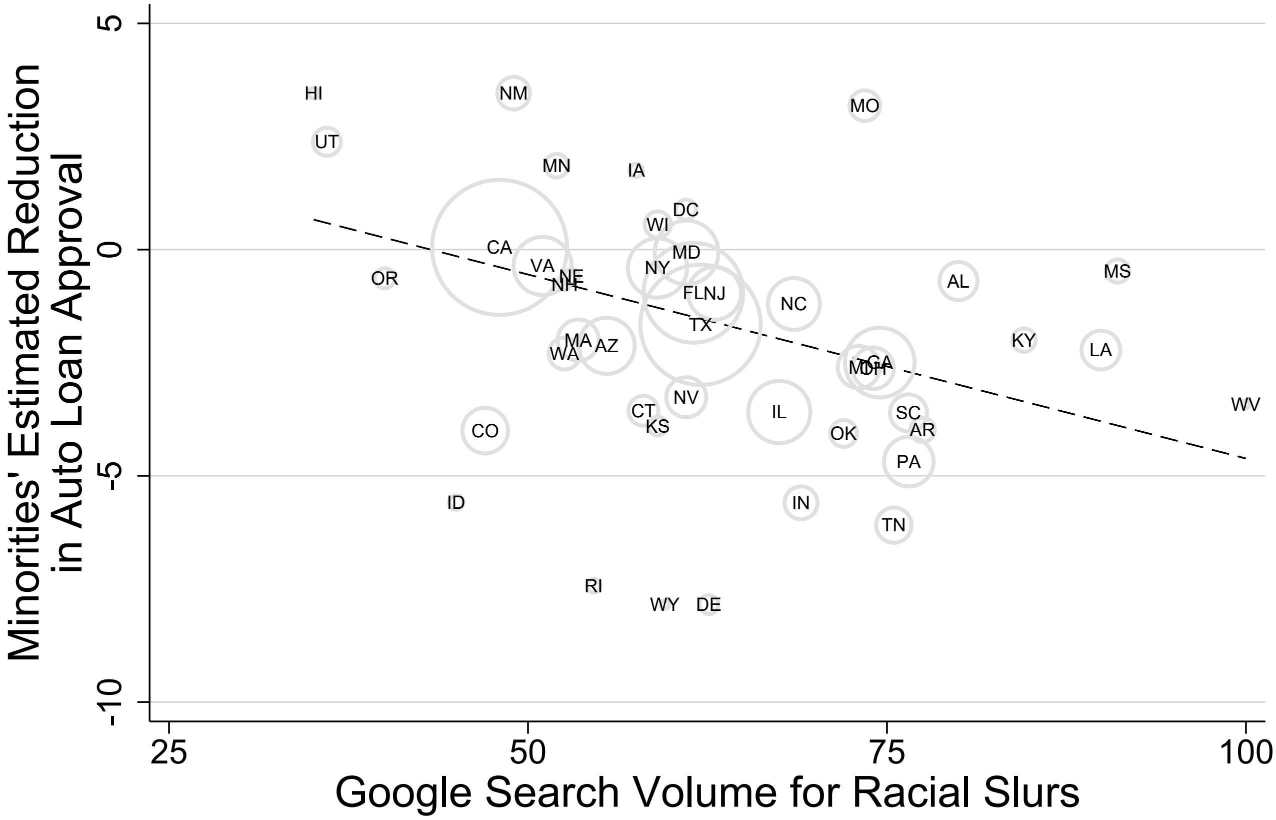

Figure 1
Auto Lending Discrimination and Racial Biases
This figure plots our point estimates of the reduction in auto loan approval rates that minorities face in each U.S. state against the prevalence of racial biases in the state measured using the Google Search Volume for racial slurs (following Stephens-Davidowitz (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).

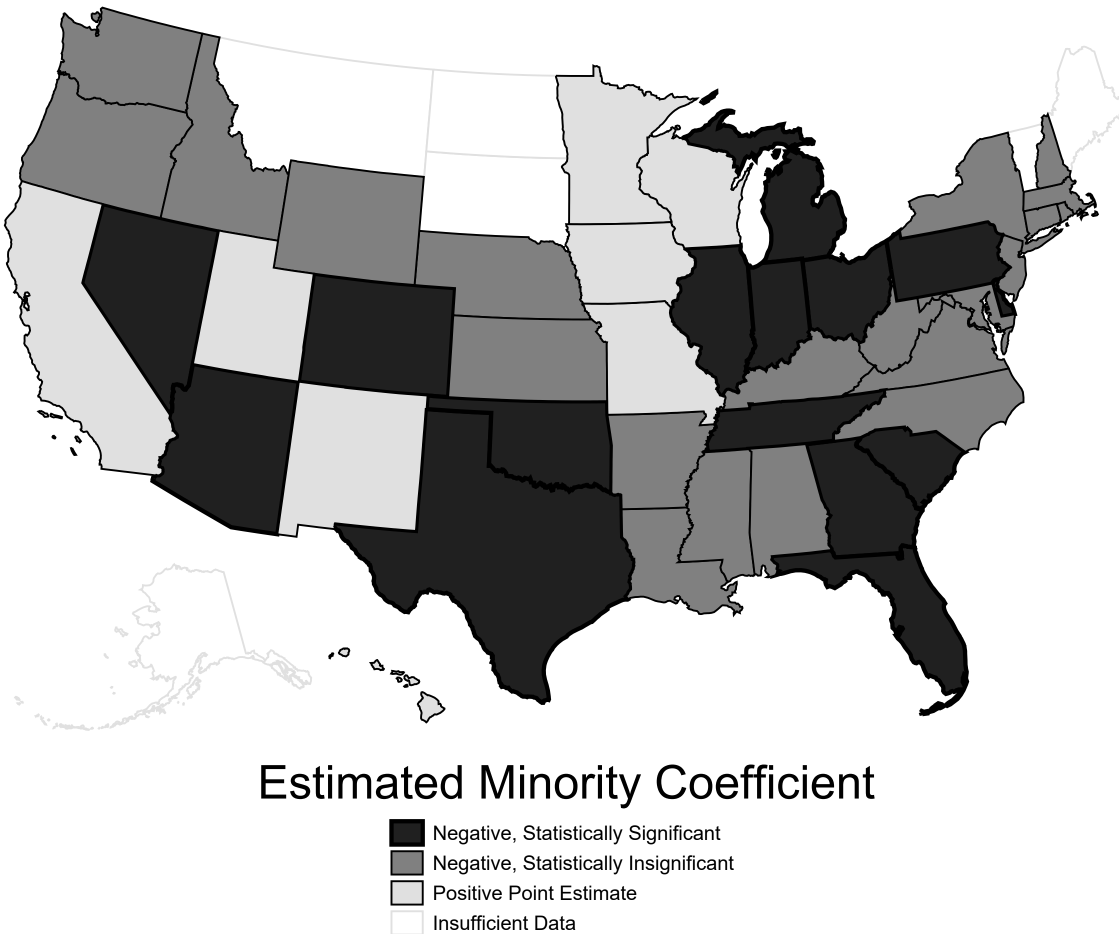

Figure 2
Where is the Evidence of Auto Lending Discrimination Strongest?
This figure presents a map categorizing U.S. states based on whether we find statistically significant evidence that minorities face reduced access to auto credit in the state. Our estimates of whether minorities face reduced access 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.

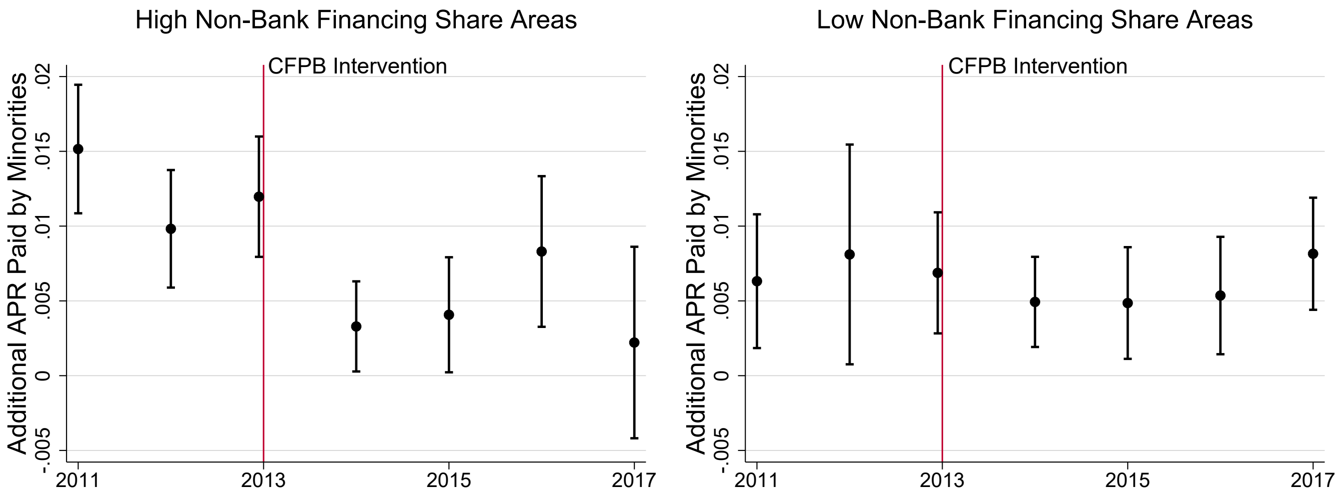

Figure 3

The 2013 CFPB Intervention and Racial Disparities in Auto Loan Interest Rates

This figure shows estimates of the additional interest (APR) minorities pay on auto loans each year from 2011-2017. The 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 7, 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.

Appendix A — Supplementary Tables and Figures

Table A.1: 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.2: Applicant 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 measures of access to auto loans and credit cards on applicant race, individual characteristics, and ZIP code characteristics. The outcome variables are indicators for the borrower successfully opening a new auto loan (Columns 1-5) or a new credit card (Column 6). The sample in Columns 1-3 includes all person-years where the individual applies for an auto loan during the year. Columns 4 and 5 restrict the sample to applicants with subprime, and prime credit scores, respectively. The sample in Column 6 includes person-years where the individual applies for both auto credit and a credit card during the year. 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	Falsification Test:
	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Auto)	Credit Approval (Credit Card)
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Minority	-3.868^{***}	-1.445^{***}	-1.857^{***}	-2.364^{***}	-0.852^{***}	0.733
	(0.338)	(0.319)	(0.409)	(0.553)	(0.316)	(0.461)
Minority X Hispanic			0.740			
			(0.508)			
Female	1.074^{***}	0.784^{***}	0.807^{***}	0.889^{*}	0.932^{***}	3.563^{***}
	(0.225)	(0.217)	(0.217)	(0.462)	(0.220)	(0.413)
Age	0.006	-0.062^{***}	-0.061^{***}	-0.007	-0.063^{***}	0.079^{***}
	(0.010)	(0.009)	(0.009)	(0.020)	(0.010)	(0.015)
Log(Income)	3.400^{***}	1.634^{***}	1.649^{***}	4.704^{***}	0.629^{**}	-0.379
	(0.211)	(0.215)	(0.216)	(0.508)	(0.253)	(0.425)
Credit Characteristics						
Credit Score $_{t-1}$		0.049^{***}	0.049^{***}	0.156^{***}	0.009^{***}	0.046^{***}
		(0.002)	(0.002)	(0.006)	(0.003)	(0.004)
Log(Total Debt $_{t-1}$)		0.841^{***}	0.842^{***}	0.302^{***}	0.951^{***}	-0.210^{**}
		(0.071)	(0.071)	(0.095)	(0.102)	(0.097)
Debt to Income $_{t-1}$		0.020	0.019	0.242	-0.213^{**}	-0.248^{*}
		(0.081)	(0.081)	(0.175)	(0.097)	(0.133)
Log(Past Due Debt $_{t-1}$)		-1.072^{***}	-1.070^{***}	-0.681^{***}	-0.408^{***}	-1.056^{***}
		(0.060)	(0.060)	(0.074)	(0.080)	(0.075)
ZIP Code Characteristics						
Log(Personal Income Per Capita)	0.124	-0.780	-0.734	0.022	-1.105	-0.410
	(0.758)	(0.744)	(0.743)	(1.432)	(0.816)	(1.472)
Log(Population Density)	-0.017	-0.041	-0.039	-0.141	0.057	1.113^{***}
	(0.082)	(0.080)	(0.081)	(0.190)	(0.091)	(0.159)
Bachelors Degree	5.267^{***}	2.015	1.933	4.309	2.551	4.177
	(1.485)	(1.470)	(1.471)	(3.086)	(1.584)	(3.076)
Commute Using Car	10.382^{***}	10.100^{***}	10.019^{***}	11.637^{***}	8.654^{***}	5.243^{*}
	(1.483)	(1.480)	(1.482)	(3.412)	(1.645)	(3.121)
State-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Time Relative to Match Indicators	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.029	0.057	0.057	0.079	0.030	0.054
Observations	132,113	130,867	130,867	38,068	92,796	74,229

Table A.3: 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).

	APR 75th Percentile	APR 75th Percentile	APR 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.4: 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)
	APR	APR	APR	APR
	(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
	APR	APR	APR	APR
	(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
	APR	APR	APR	APR
	(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

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 *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:

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

<u>Filter 1: Credit Bureau Sample</u>

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 borrowers 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.