Using High-Frequency Evaluations to Estimate Discrimination:
Evidence from Mortgage Loan Officers*

Marco Giacoletti † — USC

Rawley Z. Heimer‡ — Boston College

Edison G. Yu§ — FRB-Philadelphia

Current Draft: July 20, 2021

First Draft: August 24, 2020

Abstract

We develop empirical tests for discrimination that use high-frequency evaluations to limit the scope of unobserved heterogeneity in a conventional benchmarking test. Our approach to identifying discrimination requires two conditions: (1) the subject pool is time-invariant in a short time horizon and (2) there is high-frequency variation to the extent of which evaluators can rely on their subjective assessments. We bring our approach to the residential mortgage market, using data on the near-universe of U.S. mortgage applications from 1994 to 2018. Monthly volume quotas reduce how much subjectivity loan officers apply to loans they process at the end of the month. As a result, the volume of new originations increases by 150% at the end of the month, while application volume and quality are constant within the month. Owing to within-month variation in loan officers' subjectivity, we estimate that Black mortgage applicants have 3.5% to 5% lower approval rates, which explains at least half of the observed approval gap for Blacks. When we use this approach to evaluate policies, we find that market concentration and FinTech lending have had no effect on lending discrimination, but that shadow banking has reduced discrimination presumably by having a larger presence in under-served communities.

Keywords: Performance Incentives, Loan Officers, Mortgages, FinTech Lending, Lending Discrimination

---

*Previously titled "To the Back of the (Lending) Bus: How Loan Officers' Performance Incentives Reveal Lending Discrimination." We thank Michael Fitzpatrick for excellent research assistance. We thank Pat Akey, Mitchell Berlin, Neil Bhutta, Greg Buchak, Yulyia Demyanak, Ran Duchin, Andreas Fuster, Antonio Gargano, Hajime Hadeishi, Cameron LaPoint, Chris Parsons, Micah Spector, Phil Strahan, Nancy Wallace, and seminar participants at the ASU Finance Winter Conference, Baruch College, Boston College, Federal Reserve Bank of Philadelphia, the Joint Finance Seminar, Midwest Finance Association Conference, NBER Inequality, Discrimination, and the Financial System Virtual Conference, NBER Summer Institute, University of Bonn (MacroHistory Lab), University of California Irvine, University of Miami, University of Oregon, University of Southern California, Western Finance Association Conference, and Yale Junior Finance Conference for helpful feedback. This paper does not reflect the views of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System. Any remaining errors or omissions are the authors' responsibility.

†University of Southern California, Marshall School of Business. Phone: +1 (650) 475-6410, Email: mgiacole@marshall.usc.edu

‡Boston College, Carroll School of Management. Phone: +1 (216) 774-2623, Email: heimer@bc.edu

§Federal Reserve Bank of Philadelphia. Phone: +1 (215) 574-6221 Email: edison.yu@phil.frb.org

1 Introduction

A wide range of fields—such as labor markets, the legal system, and credit markets—have documented racial and gender disparities. Yet whether these disparities are the result of discrimination by economic decision-makers—defined as an evaluator treating otherwise identical subjects from minority groups worse than subjects from the majority group—remains in dispute because of the limitations of empirical tests. There has been a growing trend toward using experiments and correspondence studies to test for discrimination (Bertrand and Duflo, 2017). Nonetheless, tests for discrimination that use observational data have several advantageous features. Such tests are accessible to a wide range of researchers, they are easy to replicate and scale, they can be used to estimate aggregate costs of discrimination in a given market, and policymakers can easily implement them.

However, tests for discrimination based on observational data face a number of econometric challenges that limit their appeal. The most straightforward test for discrimination is an audit or "benchmarking" test. Benchmarking tests claim to find discrimination when minority groups receive unfavorable evaluations relative to the majority group. But, benchmarking tests are vulnerable to criticisms of omitted variable bias—differences in group characteristics, which the researcher does not observe, can cause differences in evaluations across groups.

Alternatively, researchers can use an "outcome test" (e.g., Becker, 1957). Instead of comparing differences in how groups are evaluated, outcome tests compare the ex-post success of these evaluations. The marginal minority will have better ex-post outcomes than the marginal majority subject because minority groups face higher thresholds for inclusion when they are subject to discrimination. Though intuitively appealing, outcome tests are notoriously difficult to implement, most notably because of the "infra-marginality" problem—the average difference in ex-post outcomes can be a poor approximation of the difference in marginal outcomes (Ayres, 2002). Recent research has made significant progress to improve econometric methods (e.g., Arnold et al., 2018), but addressing the infra-marginality problem requires additional modelling and distributional assumptions (Simoiu et al., 2017). Furthermore, ex-post outcomes can be the result of self-fulfilling prophecies (e.g., female students underperform in math because gender stereotypes reduce investment in females' math education; Bordalo et al., 2016) and ex-post outcomes are often not easily measured (e.g., worker productivity can be difficult to measure and discrimination can also distort proxies for productivity, such as wages).

We propose a modification to a conventional benchmarking test that limits the scope for omitted variables to drive observable differences across groups. Our approach is motivated by the observation that evaluators' subjectivity can often vary substantially within short time intervals. For example, employers that have immediate staffing needs can ill afford to turn away job applicants. TSA agents might reduce their screening of travelers when they are at the end of their shifts or there are long queues. Police officers that have monthly quotas would issue tickets to all drivers that exceed the speed limit on the last day of the month. Our approach starts with a benchmarking test, but addresses the problem of omitted variables by exploiting such high-frequency evaluations. The approach requires two simple assumptions: time-varying discrimination and time-invariant unobserved characteristics both in a short time interval. The identification rationale is straightforward. If the evaluations of a group vary within a short time interval, then these differences cannot be driven by unobserved subject characteristics, because the unobserved characteristics are time-invariant.

We apply our approach to high-frequency data on mortgage applications, to test for discrimination in the U.S. residential mortgage market.[1] We obtain the time-stamped version of the Home Mortgage Disclosure Act (HMDA) data, covering the near-universe of mortgage applications from 1994 to 2018 with over 500 million loan applications across more than 28,000 lenders.

Figure 1 demonstrates our key source of high-frequency variation in the mortgage market and the foundation of our empirical approach. The figure shows the volume of new originations and new applications relative to the first day of a given month. The total volume of new mortgage

[1]The literature can be traced back at least as far as the public release of HMDA data and the work of Munnell et al. (1996). Ladd (1998) summarizes much of the older literature and frames longstanding debates. Other foundational papers include Berkovec et al. (1994); Tootell (1996); Berkovec et al. (1998). Recent work studies differences in mortgage rates and fees (e.g., Bhutta and Hizmo, 2020).

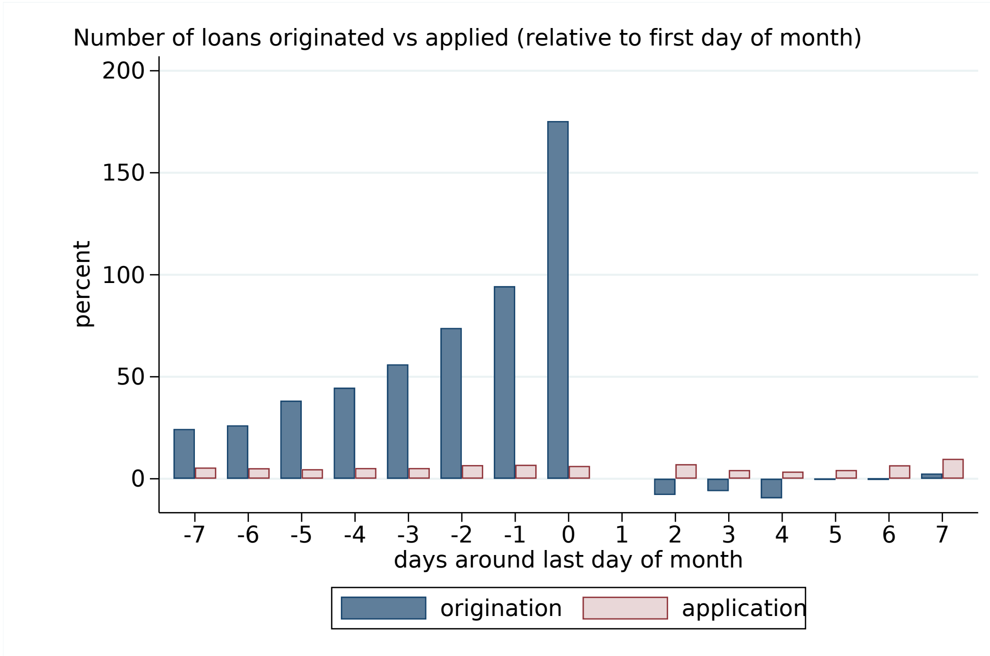


Figure 1: The figure shows average percentage abnormal daily loan origination volume, and loan application volume (measured as number of originations and applications) in the U.S., for the last eight days of the month, and the first seven days of the following month. The figure reports the average across all months over the sample period from January 1994 to December 2018 from the HMDA data. Abnormal volume is computed with respect to loan originations and applications on the first day of the following month.

originations increases by more than 150% on the last day relative to the first day of a given month. At the same time, the number of submitted mortgage applications stays constant over the course of the month. These patterns reveal a crucial feature of the mortgage application process: loans are processed by individual loan officers who have monthly performance targets that determine their compensation.[2] Moreover, this within-month pattern in loan approvals unveils the component of loan officers' decision-making that is orthogonal to observable and unobservable factors affecting loan originations (e.g., credit market conditions, applicant characteristics, and firm-level characteristics). Drawing from Becker (1957), a profit-maximizing agent can give disparate treatment to minority populations until market competition makes discrimination economically untenable.

[2] Though we are unable to obtain the compensation of individual loan officers, the most common compensation scheme includes commissions that are set based on the number of loans and the loan amount originated. For example, the Mortgage Bankers Association describes industry standards for loan officers' compensation: *https://www.mba.org/publications/insights/archive/mba-insights-archive/2019/is-it-time-to-rethink-compensation-x253848*. Loan officers may also face disciplinary action if they fail to meet their quotas several months in a row (Tzioumis and Gee (2013)). Given this non-linear incentive scheme, Tzioumis and Gee (2013) and Cao et al. (2020) document end-of-month bunching in a large U.S. commercial bank and in two Chinese banks, respectively.

Loan officers have an economic incentive to meet end-of-month performance incentives. As such, loan officers' subjective favoritism toward applicants has to attenuate at the end of the month relative to the beginning of the month. Therefore, the within-month pattern, combined with a conventional benchmark test, allows us to estimate the extent to which loan approval decisions can be attributed to loan officers' subjectivity towards applicants.

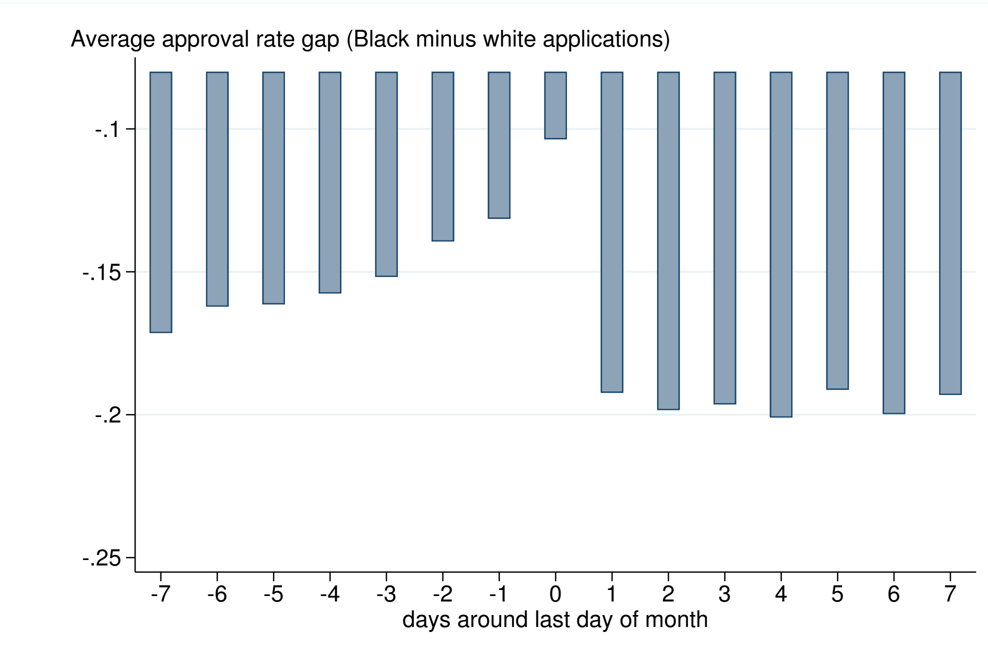


Figure 2: The figure reports the difference between the fraction of approved loans, out of all approved and denied loans in the U.S., for Blacks minus the one for whites, on each of the last eight days of the month and the first seven days of the following month.

Exploiting this within-month variation, our tests for discrimination estimate the difference in approval rates between Black and white applicants at the start of the month relative to the end of the month. Figure 2 summarizes our main finding. It shows the difference in application approval rates between Black and white applicants over the course of any given month. In the first seven days of the month, Black applicants have 20 percentage point lower approval rates than white applicants. The approval gap gets smaller over the course of the month. The approval gap between Blacks and whites is just 10 percentage points on the last day of the month. Of course, these raw estimates do not account for observable differences in loan applications.

Regression analysis confirms the graphical evidence in Figure 2. The regression tests are saturated with a rich set of fixed effects that control for time-varying economic conditions at precise geographic levels, namely county-month, as well as lender-month fixed effects that control for factors, such as regulations, that would affect lending at the institution level. The regressions also include applicant characteristics interacted with day-of-month fixed effects to allow lenders' decision criteria to change flexibly over the course of the month. In our most stringent tests, the difference in the Black-white approval gap between the start and the end of the month is 3 to 5 percentage points. This constitutes a lower bound on the share of the Black-white approval gap that is due to loan officers' subjectivity, relative to the approval gap that can be attributed to unobservable group-differences. The estimates suggest that loan officers' subjective decision-making explains at least half of the overall difference in approval rates between Black and white applicants after controlling for observable characteristics. Furthermore, these estimates are similar across different types of mortgage lending, such as FHA loans and refinances. This robustness across mortgage products helps exclude alternative explanations, such as there being a financial incentive to close on the last day of the month.

We use these estimates to assess the magnitude of discrimination in mortgage lending over the past several decades. Using a back of the envelope calculation for the upper-bound of the costs of discrimination, if the approval rate gap for every day of the month was as small as the last day of the month, about 1.4 million more Black applications would have been approved rather than denied between 1994 and 2018. This difference in loan approvals corresponds to approximately \$213 billion (in 2018 dollars) total loan volume since 1994.

Our approach to estimating discrimination hinges on simple assumptions that we derive and find support for via narrative and in the data. The first assumption is that the loan officer has time-varying costs of being subjective. In our setting, loan officers have nonlinear contract incentives.[3]

[3] Importantly, with volume quotas, the optimal strategy would be to approve all loan applications. However, in practice, there are several constraints on this strategy. Lenders set origination standards that an application has to exceed and loan officers may have a fixed quantity of mortgage credit that they can distribute within a month. Loan officers can use their discretion and work to sidestep the origination standards by either using risk-based pricing or appealing to other "soft" criteria (e.g., the applicant holds other accounts with the bank).

Loan officers that fail to meet their volume quotas will have reduced compensation and risk getting fired.

The second assumption is that the characteristics of the subject pool are time invariant. Indeed, we find that application volume, the relative share of Black applications, and loan application quality (for both Black and white applications) are all constant over the course of the month. The remaining threat to identification is that there are differential trends by race in the quality of applications that get processed over the course of the month. As evidence against this explanation, we find that high-quality and low-quality Black mortgages have similar amounts of bunching toward the end of the month.

To further support the assumption of time-invariant application quality, we use a new sample of HMDA data (post-2018) that includes the use of automated underwriting systems' (AUS) recommendations. The AUS recommendations are generated by computer algorithms—such as those produced by Fannie Mae and Freddie Mac—and they are intended to offer a race-neutral evaluation of applications. We find that there is a racial gap in AUS recommendations—Black applicants are recommended for approval approximately 6 percentage points less frequently. However, unlike *actual* approval rates, the gap in AUS recommendations is roughly constant over the course of the month. Moreover, our regression evidence is robust to including the AUS recommendation as a control variable. These findings suggest that lenders' subjective decision-making causes the residual difference in approval rates over the course of the month.

In contrast to other methodologies, our approach does not require ex-post outcomes to test for discrimination. Nevertheless, we show that a conventional outcome test is potentially misleading about the levels of discrimination in mortgage lending. We find that Black mortgages have significantly higher unconditional average rates of default, which could be interpreted as evidence that mortgage markets actually favor minorities. Instead, this result is likely caused by the inframarginality problem – Black and white mortgage applicants have different risk distributions. We compare the subsequent default rates of applications approved at the start of the month to those approved at the end of the month. We do not observe within-month variation in the default rate gap between Black and white applications. As such, these findings suggest that our approach potentially counteracts the shortcomings of a conventional outcome test.

Furthermore, our approach offers guidance, relative to both benchmarking and outcome tests, as to whether observed discrimination is caused by taste-based versus statistical discrimination. Though we are unable to definitively disentangle competing theories of discrimination, we develop additional assumptions to assess the possible contribution of these theories. Put simply, the case for statistical discrimination requires asymmetric information between evaluators and subjects. Because of the high-frequency nature of our data, statistical discrimination would require loan officers' information sets about applicants to change over the course of the month. However, the applicant pool is time invariant, which leaves some form of taste-based discrimination as the most likely explanation. Related, we consider the role of inaccurate beliefs (see e.g., Bohren et al., 2020) and a similar logic makes this explanation unlikely.

Finally, our approach is advantageous because it can easily be applied to evaluate the effect of market policies and market innovations on the quantity of discrimination. We consider three important features of modern mortgage lending: market concentration in banking, FinTech lending, and shadow banking. We find that the amount of discrimination due to loan officers' subjectivity is unaffected by both market concentration and FinTech lending. This result is largely consistent with the fact that our regressions include lender-by-month fixed effects and that the component of loan officer subjectivity our approach uncovers occurs *within-lender*. Moreover, despite these changes to the banking sector, loan officer compensation incentives have largely remained constant throughout our sample, and even mortgage lending at FinTech lenders involves significant discretion from human loan officers. On the other hand, we find that shadow banks have lower levels of subjective discrimination against Black applicants. This is likely the result of shadow banks—owing to their lower regulatory requirements—having a larger presence in under-served communities.

Related Literature

Our paper is related to advances in the literature on identifying discrimination by economic decision-makers. Our approach is akin to empirical papers that use changes to evaluation settings to identify discrimination against minority groups. For example, Goldin and Rouse (2000) shows that blind auditions reduce employment discrimination against female orchestra musicians. Police officers are less likely at night than during the day to pull over Black motorists because the driver's race is difficult to identify (Pierson et al., 2020). These empirical papers identify discrimination by comparing situations in which evaluators know the subject's gender and race to situations in which they do not. Our approach is different because there is no change in the loan officer's knowledge of applicants' race. We show that discrimination can be identified under certain assumptions about the applicant pool and loan officers' reliance on subjective assessments.

More specifically, our paper joins the literature on discriminatory lending practices in consumer credit markets.[4] Our empirical approach is grounded in evidence that loan officers have significant discretion in loan processing decisions (see e.g., Engelberg et al., 2012; Chen et al., 2016; Cortés et al., 2016; Demiroglu et al., 2021). Guided by these findings, we bring the confidential HMDA data to the question of lending discrimination. Recent papers advance the literature on discrimination in mortgage lending by obtaining richer data sets that allow for more control variables in a classic benchmarking test (Bartlett et al., 2021; Bhutta et al., 2021). Our approach to estimating discrimination is less reliant on cross-sectional control variables, although our approach can only suggest a lower bound on discrimination. Furthermore, most papers are unable to distinguish between taste-based and statistical discrimination (Bohren et al., 2019 and Dobbie et al., 2020 are notable exceptions). Our approach can offer guidance about which type of discrimination is most likely. Also, many papers use evidence from confidential internal data from one

[4]Several papers use different approaches to provide evidence that individual loan officers discriminate against other races and women. In no particular order, Fisman et al. (2017, 2020) use data from Indian banks to show that loan officers are more favorable to culturally proximate applicants. Beck et al. (2018) use evidence from an Albanian bank and Montoya et al. (2019) use a field experiment at a Chilean bank to uncover evidence of gender discrimination in consumer lending. Ambrose et al. (2020) find that minorities pay higher fees when they have a white mortgage broker. Other papers study discrimination in auto lending (e.g., Butler et al., 2019; Lanning, 2021).

or two lending institutions. In contrast, our paper uses the universe of U.S. mortgage applications over a 25-year period to connect racial disparities in lending to the incentives of individual loan officers. This allows us to address crucial questions about external validity, investigate the effects of the market structure, and quantify the scope of racial bias in mortgage markets.

Second, our paper contributes to a growing literature on how market structure and technology affect consumer lending. First, recent papers find support for classic theories arguing that competition reduces discrimination in consumer lending (Buchak and Jørring, 2017; Butler et al., 2019). Such papers suggest that racial discrimination declines because of changes to the composition of lending institutions. We find that discrimination by individual decision-makers can persist within organizations even when there are differences in institution-level competition across markets. Second, there is significant debate over how the growth of FinTech lending affects the allocation of credit, with a particular interest in the effects on disadvantaged borrowers. In theory, FinTech can reduce intermediation costs which can pass through to consumers (e.g., Tang, 2019) or improve screening (Berg et al., 2020a). However, the literature finds that FinTech algorithms have either no effect or negative effects on the supply of credit to disadvantaged consumers (see e.g., Fuster et al., 2021; Bartlett et al., 2021). Our contribution is to show that biases in human decision-making can survive advances in loan processing technology (similar to findings on the introduction of machine learning to judicial outcomes, as shown by Kleinberg et al., 2018).

Finally, we make a unique contribution to the literature on performance-based compensation, with a particular focus on financial intermediation. The effects of performance-based compensation have been studied in a range of settings, such as manufacturing (Oyer, 1998), software sales (Larkin, 2014), government contracts (Liebman and Mahoney, 2017), healthcare (Li et al., 2014; Gravelle et al., 2010), firm managers (Bandiera et al., 2007) and accounting (Murphy, 2000). The literature has also studied how performance incentives within banks affect loan officers' effort and performance (Agarwal and Ben-David, 2012; Cole et al., 2015), and information production (Hertzberg et al., 2010; Qian et al., 2015; Berg et al., 2020b). To the best of our knowledge, our paper is the first to study variation in performance incentives combined with racial biases in human decision-making.

2 Identifying Discrimination

This section presents a formal discussion of our empirical setup. Existing frameworks for identifying discrimination, face important challenges when dealing with differences in unobserved characteristics across subject groups. Our approach is to "filter out" these unobserved differences using high frequency data and fluctuations in decision makers' subjectivity.

More specifically, our approach extends conventional tests for discrimination, called either audit or benchmarking. These tests compare the conditional likelihood that a minority subject group receives disparate treatment relative to the majority group, after controlling for other observable characteristics from the point of view of the researcher. Assume that the decision maker is considering whether to take a potential favorable decision affecting an individual, such as loan approval, exemption of TSA screening, or hiring. The researcher claims to have uncovered discrimination when she rejects the null of no difference in the conditional likelihood of the favorable decison between minority and majority groups, for instance Blacks and whites, and instead finds that the likelihood is significantly smaller for Blacks. Specifically, the researcher claims discrimination when she finds that:

$$P(Y|W,X) > P(Y|B,X) \tag{1}$$

where $P(Y|R,X)$ is the probability of receiving a favorable decision, conditional on race $R \in \{W,B\}$ (white or Black) and a vector of observable characteristics X to the researcher. However, this approach is exposed to the criticism that the difference in the estimated conditional probability between white and Black subject groups might be driven by unobserved characteristics that are relevant for the assessment of by the decision makers, but are not included in the vector of controls X used by the researcher. To see that, assume for simplicity that there is a binary unobserved variable $Z \in \{Z_L, Z_H\}$, such that the following assumptions are satisfied:

Assumptions Set (A)

$$\begin{aligned}
&\text{No discrimination:} && P(Y|W,X,Z_k) = P(Y|B,X,Z_k) \\
& && \text{for } k \in \{H,L\} \\
&\text{Higher quality predicts higher favorable decision probability:} && P(Y|R,X,Z_H) > P(Y|R,X,Z_L) \\
&\text{On average white applicants have better unobservables:} && P(Z_H|W,X) > P(Z_L|B,X)
\end{aligned}$$

The inequality in favorable decisions formalized by equation (1) holds under the set assumptions above when omitting the variable Z, even though *decision-makers do not discriminate when all of the characteristics are accounted for* (see Appendix A.I). The differences in the observed conditional probability for different races simply capture the differences in the unobserved characteristics. In the mortgage-lending setting as an application of the methodology in this paper, Black and white applicants have substantially different observable characteristics (see Table 1). Such differences raise concern that there might be also substantial differences in unobservables.

In this paper, we show that we can refine existing approaches to address the identification problems due to the systematic differences in unobservables across subject groups. Rather than only testing for the differences in the likelihood of a favorable decision between racial groups, we use high-frequency data to test whether those differences vary over a short period of time. Because discrimination is determined by the subjective judgment of the evaluators, under the null of no discrimination, and if applicant characteristics remain constant over time, there shall be no change in the decision probability over time. On the other hand, discrimination would predict a change in the relative favorable decision probability over time.

To formalize this idea, let there be two time periods, $T \in \{Start, End\}$. Assume that evaluators have more scope to be subjective in period $Start$ relative to period End. Then, in the presence of time-varying discrimination we expect to find:

$$P(Y|W,X,End) - P(Y|B,X,End) < P(Y|W,X,Start) - P(Y|B,X,Start) \qquad (2)$$

where $P(Y|., X, .)$ is the probability of receiving a favorable decision, conditional on race (white or Black), a vector of observable characteristics X, and in a specific period ($Start$ or End). Note that the presence of unobservable quality characteristics systematically correlated with race cannot alone explain the effects in equation (2). Consider the following set of assumptions that characterize a situation in which there is no discrimination:

Assumptions Set (B)

No discrimination:	$P(Y\|W, X, Z_k, T) = P(Y\|B, X, Z_k, T)$
	for $k \in \{H, L\}$
Higher quality predicts higher favorable decision probability:	$P(Y\|X, Z_H, T) > P(Y\|X, Z_L, T)$
On average white applicants have better unobservables:	$P(Z_H\|W, T) > P(Z_H\|B, T)$
No time pattern in subject group quality:	$P(Z_H\|R, X, Start) = P(Z_H\|R, X, End)$

The first three assumptions are the same as in **Assumptions Set (A)**, while the last assumption states that the unobserved characteristics of the applicants, for both whites and Blacks, are on average constant over time. Jointly, these assumptions imply (see Appendix A.I):

$$P(Y|W, X, End) - P(Y|B, X, End) = P(Y|W, X, Start) - P(Y|B, X, Start).$$

Thus, the condition in equation (2) indeed amounts to a rejection of the null of no discrimination.

2.1 Distinguishing Taste-Based from Statistical Discrimination

This section explores the extent to which our approach can distinguish between the two broad categories of discrimination. Under "taste-based" discrimination, minorities are subject to disparate treatment because evaluators have animus toward them. Under "statistical" discrimination, evaluators are uncertain about the abilities of any given subject. Evaluators form their beliefs after observing the subject's race. Minorities are subject to disparate treatment when evaluators have developed beliefs that minority subjects have worse abilities. Evaluators do not need to have accurate beliefs about minorities to apply disparate treatment (see e.g., Bohren et al., 2020).

We consider evaluators j who, over a short time-period, for example a month or a week, evaluate subjects i. Each evaluator j has perceived net benefits from making decisions that favor subject i equal to $U^j(X_i, Z_i, R_i, t)$, where X_i and Z_i are vectors of observable and unobservable (from the perspective of the researcher) characteristics, R_i is the subjects' race (e.g., $R_i = W$ for a white applicant and $R_i = B$ for a Black applicant), and t is the point in time in which the evaluation is conducted.

The evaluator's net benefits can be decomposed into two components:

$$U^j(X_i, Z_i, R_i, t) = b^j(X_i, Z_i, R_i, t) + E_j[u_i|X_i, Z_i, R_i, t], \tag{3}$$

where $b^j(X_i, Z_i, r_i, t)$ is the subjective net benefits of evaluator j conditional on all characteristics, and $E_j[u_i|V_i, Z_i, R_i, t]$ is the statistical component. The statistical component can be written as

$$E_j[u_i|X_i, Z_i, R_i, t] = E[u_i|X_i, Z_i, R_i, t] + \tau_j(R_i, t), \tag{4}$$

where $\tau_j(.)$ is the bias of decision maker j when forming expectations conditional only on the information about the race of an applicant.

We can then use our stylized framework to characterize different types of discrimination, for example against Black subjects with respect to white subjects:

- Taste-based discrimination: $b^j(X_i, Z_i, W, t) > b^j(X_i, Z_i, B, t)$
- Statistical discrimination: $\tau_j(W, t) > \tau_j(B, t)$

The decision maker will take a decision favorable to subject i as long as the net benefit is positive:

$$b^j(X_i, Z_i, R_i, t) + E_j[u_i|X_i, Z_i, R_i, t] + v_{i,j,t} > 0,$$

where $v_{i,j,t}$ is a random preference shock, i.i.d. across subjects and evaluators, and independent of information on subject characteristics and evaluators' beliefs. We can then introduce the variable $y_{i,j,t}$, which is equal to one if subject i receives a favorable decision from evaluator j at time t, and has likelihood function:

$$\mathcal{L}(y_{i,j,t}) = Pr(y_{i,j,t} = 1)^{I(y_{i,j,t}=1)}[1 - Pr(y_{i,j,t} = 1)]^{1-I(y_{i,j,t}=1)}$$
$$Pr(y_{i,j,t} = 1) = E[y_{i,j,t}|X_i, Z_i, R_i, j, t] = F(X_i, Z_i, R_i, j, t).$$

If we assume the function $F(X_i, Z_i, R_i, j, t)$ can be approximated with a liner specification, then we can write:

$$y_{i,j,t} = \beta_1 r_i + \beta_2 (r_i \times t) + \eta X_i + \phi Z_i + a_t + \epsilon_{i,j,t} \tag{5}$$

where $r_i = 1$ if $R_i = B$. Equation (5) can be estimated in the data. Within this specific framework, we can state the predictions of our general discussion in the previous section along the following lines:

1. The two types of discrimination listed above (driven by taste or statistical) would cause estimates of $\beta_1 < 0$. However, as the previous section outlines, β_1 will be a biased estimate unless the researcher fully controls for observable (X_i) and unobservable (Z_i) characteristics, or r_i is uncorrelated with any omitted characteristics.

2. Estimates of β_2 will be different from zero if the magnitude of discrimination changes over time, regardless of the type of discrimination. If subject pool characteristics (X_i and Z_i) are not correlated with the evaluation time t, estimates of β_2 will be unbiased even if the researcher does not perfectly control for time-invariant characteristics.

How can this approach distinguish between different theories of discrimination? In principle, any type of discrimination can be subject to high-frequency fluctuations, and thus produce non-zero estimates of β_2. However, if the unobserved variation across subject pools and the evaluator's statistical inference problem are time-invariant, our approach allows the researcher to attribute discrimination to the source of time-variation in the evaluator's decision-making.

Consider the case of statistical discrimination. Statistical discrimination is caused by the evaluators' statistical inference problem. Therefore, the researcher can reasonably assume the findings are caused by statistical discrimination if she can provide evidence of time-variation in the evaluators' information set. Now consider taste-based discrimination. The evaluator's subjective preferences against minorities causes disparate treatment. The researcher can assume taste-based discrimination if she has evidence that evaluators' subjectivity is time-varying.

In the following empirical analysis, we focus on residential mortgage lending in the U.S. Our source of time-variation in evaluations is the fact that loan officers have monthly volume quotas. These monthly volume quotas generate within-month variation in loan officers' subjectivity. At the same time, loan officers observe the same information about applications that they process at the start of the month relative to the end of the month. As such, any finding of discrimination due to within-month differences in evaluations can be attributed to loan officers' subjective preferences.

3 Data

The empirical results in this paper are based on the confidential version of the HMDA data available to researchers in the Federal Reserve System. The dataset contains the largest sample of mortgage applications available in the U.S. The public version of the data includes information on applicant characteristics – race, gender, reported income, and location of the property – and identifiers for the lenders that received the applications. The data cover the entire geography of the U.S. over the period from January 1994 through December 2019. Moreover, the data provide information on mortgage contract characteristics, such as whether the application is for a new home purchase or refinancing, the loan amount, the lien, and whether the property is owner-occupied. The primary distinguishing feature of the confidential version of the HMDA data is that it contains the exact date on which each application was submitted by a potential borrower and the date on which the lender took action on the application, either by originating the loan or denying the application.

This information has been employed in several prior papers (see e.g., Cortés et al., 2016). Because the exact timing of the lender decisions is crucial to our study, Section 5.4 describes the relation between the timing of denial, origination and approval in the US mortgage market.

Annual mortgage applications in the HMDA dataset over the years from 1994 to 2019 are between 10.1 and 37.3 million, and originations between 7.2 and 23.7 million. The number of active lenders by year is between 5,700 and 9,800, and the average number of originations per lender is between 750 and 2,900.

Panel (a) of Table 1 shows summary statistics over the period from 1994 to 2018, and across different applicant groups based on race. Approximately 67% of applicants are white and 7% are Black. Other races are not separately identified in our analysis and are grouped, along with applications that do not specify race, into a single category called "Other race" that includes 26% of all observations. Black applicants apply for smaller loans on average, have the highest fraction of low income applicants (59.8%, compared to 46% for whites and 47.8% for other races), and have the lowest approval rate (63.25%, compared to 80.72% for whites and 69% for other races). When considering approved loans, 73.7% are to whites, 5.7% to Blacks and 20.6% to applicants of other races.

To obtain more detailed information on characteristics and performance of *originated* mortgage loans we also construct two additional datasets. First, we merge HMDA with the Black Knight McDash (McDash) dataset. We construct the merged sample with an algorithm similar to the one used by Rosen (2011). Individual observations in HMDA and McDash are merged using loan origination date, loan amount, zip code, lien type, loan type, loan purpose, and occupancy type (owner occupied, absentee or investment property). The match rate of the merge is about 60%.[5] McDash provides further information on individual loan contracts, such as the mortgage interest rate, rate type (fixed or adjustable rate), the mortgage term, whether the loan is conforming, borrowers' FICO scores, and the quality of the supporting documentation submitted by the borrower. Panel (b) shows summary statistics by race for this matched sample. The composition

[5]Lender and consumer identities were anonymized for the merged dataset used in this analysis. There is no within-month pattern in the match rates.

of loans by race lines up with the one shown in Panel (a), for the entire HMDA sample. Black borrowers again obtain smaller loans and have lower income than other borrowers. Moreover they are also more likely to have FICO below prime (below 660) and loan-to-value (LTV) above 80%.

Second, we obtained the extended version of the HMDA dataset, available for the years 2018 and 2019. This new version of the dataset contains more detailed underwriting for loan applications, such as credit scores, LTVs, and debt-to-income ratios. Moreover, it also contains the approval recommendation generated for each loan by lender's Automated Underwriting System (AUS). Statistics for this last dataset are reported in panel (c) of Table 1, and show that the composition of the sample across race group and other applicant characteristics is similar to the one of for the full HMDA sample displayed in panel (a).

4 Identifying Assumptions

Our identification strategy relies on high-frequency variation in evaluators' subjective decision making. This section provides support for the identification assumptions: (1) the pool of mortgage applicants is time-invariant and (2) there is time-variation in loan officers' reliance on their subjective assessments.

4.1 The Applicant Pool is Time-Invariant

First, our tests identify discrimination under the assumption that the composition of the applicant pool is time-invariant. Figure 3 shows how the composition of new applicants evolves over the course of the month. Panel (a) plots the average share of Black applicants submitted on each day of the month. Black applicants' share of new applications is roughly constant at approximately 7% on each day of the month. This confirms our identifying assumption that the racial composition of applicants is time-invariant.

We also verify that other characteristics of the applicant pool—characteristics that could correlate with race—are constant over the course of the month. The HMDA data before 2018 has limited information on the creditworthiness of applicants. However, the data contain applicants' income, which is an important input into lender's decision-making and is likely correlated with other variables that determine whether an application is approved (e.g., credit scores). Figure 3, Panel (b) reports the fraction of applicants that have levels of personal income that are below the median of applicants within a county during a given year. Panel (c) shows the share of such applicants with the added restriction that the application becomes a new origination. In both panels, we divide the sample into applications submitted by Black and white applicants. As such, these figures explore whether the quality of applications within and across races changes within the month. These figures show that application quality is constant.

Lastly, Panel (d) studies the composition of the applicant pool with outstanding applications (i.e., applications that have been submitted but have yet to receive an approval decision) in the lenders' inventory over the course of the month. We explore this measure of application inventory because it captures what applications the loan officer has the opportunity to work on at any point during the month. Panel (d) also sorts the outstanding applications by income and by race. Again, we find that the applicant pool is constant over the course of the month, both in terms of the racial composition of the applicant pool and the quality of the applications outstanding.

4.2 Time-Variation in Subjective Assessments of Applicants

4.2.1 Monthly Volume Quotas and End-of-Month Bunching in Mortgage Originations

Mortgage loan officers tend to receive commissions calculated as a percentage of the total amount they originate over the month. They can also receive bonuses for meeting monthly origination targets, as well as face disciplinary actions or be fired for failing to meet volume targets.[6] The use of volume-based incentives is acknowledged by U.S. regulations and directives from the Consumer Financial Protection Bureau (CFPB). U.S. law permits the use of volume-based incentives, but it

[6]See Tzioumis and Gee (2013), and evidence from practitioners' research and discussions, such as what is reported in the following articles on industry standards for loan officers' compensation in the U.S., published by the Mortgage Bankers Association (*https://www.mba.org/publications/insights/archive/mba-insights-archive/2019/is-it-time-to-rethink-compensation-x253848*), and by consumer websites (*https://www.investopedia.com/ask/answers/120214/whats-average-salary-loan-officer.asp* and *https://www.thetruthaboutmortgage.com/loan-officer-jobs/#salary*).

restricts the use of commissions based on the terms and performance of individual loans (see, most recently, the dispositions of Regulation Z, implementing the Truth in Lending Act).[7]

We find that monthly volume quotas cause large increases in new mortgage originations at the end of the month. Figure 1, described in the introduction, presents the average volume of new originations per day relative to the first day of any given month. The volume of new mortgage originations grows over the course of the month. The origination volume is more than 150% larger on the last day relative to the first day of a given month. The figure documents clear evidence of "bunching" at the end of any given month.

The end-of-month bunching in mortgage originations is robust across time and to seasonal factors. The end-of-month increase in originations occurs in every year of our sample, which suggests that the finding is not caused by business cycles and is therefore unlikely to be caused by fluctuations in the demand for mortgages (see Appendix Figure A.1). Also, the end-of-month bunching occurs in every month of the calendar year (see Appendix Figure A.2, which plots the average number of new originations on the first and last seven days of each month within a given year).[8] This suggests that the finding is not caused by seasonality in mortgage demand.[9]

Building on our graphical evidence, we use regression analysis to show that the within-month pattern in originations is not caused by confounding factors. We estimate the following regression:

$$\log(N_t) = \beta_{lw} I_{lw} + \beta_{fw} I_{fw} + a_{ym} + a_{dow} + a_{holiday} + e_t \tag{6}$$

where the dependent variable $\log(N_t)$ is the log of the number of originated mortgages by lender i on day t. The regression includes year-month, day-of-week, and bank-holiday fixed effects, which are a_{ym}, a_{dow}, and $a_{holiday}$, respectively. I_{lw} and I_{fw} are dummies equal to one for days in the last

[7] Volume-based incentives are the first form of compensation mentioned in the section on *Permissible Methods of Compensation* in the most recent revision of Regulation Z, available at *https://www.federalregister.gov/documents/2013/02/15/2013-01503/loan-originator-compensation-requirements-under-the-truth-in-lending-act-regulation-z.*

[8] As a testament to the quality of our micro-level data, recurring-day bank-holidays are clearly visible in Figure A.2. Origination volume is abnormally low on the first day of January, on Christmas and on July 4th.

[9] It is beyond the scope of our paper to disentangle competing mechanisms for why the volume targets cause the end-of-month bunching in originations. It may be optimal for loan officers to exert less effort at the start of the month relative to the end. Loan officers may procrastinate or be overconfident in their ability to meet their quotas.

week of the month and the first week of the following month. The coefficient of interest, β_{lw} (β_{fw}), measures the difference between the average origination volume in the last (first) seven days of the month, relative to the middle days of the month.

The regression estimates confirm that loan origination volume increases significantly in the last days of the month. In Table 2, when origination volume is measured as the log number of loans, the point estimate of β_{lw} is 31%, and the estimate of β_{fw} is -15%. When origination volume is measured as the total dollar amount originated per day, the point estimates are 36% and -14%. This gives us estimates of the increase in origination volume between the first and last week of the month of 46% and 50%, which are qualitatively consistent with the evidence shown using the raw data in Figure 1. Our findings are unlikely to be explained by lending seasonality because the estimates are robust to including a rich set of calendar time fixed effects (see e.g., Murfin and Petersen, 2016).

We also show that the end-of-month bunching in new originations is consistent with loan officers managing the inventory of applications over the course of the month. Figure A.3 in the Appendix shows the inventory of applications that await a decision (approval, denial, or withdrawal by the applicant) for each day within the month. There is a sharp drop in inventory over the last week of the month, driven by the spike in originations, and then a steady increase taking place over the first two weeks of the following month.

4.2.2 Linking Origination Volume to Loan Officers' Performance

Next, we connect loan officers' performance incentives to the end-of-month bunching in new originations. To do so, we consider how loan officers' monthly volume targets affect their economic incentive to approve and deny applications. Specifically, we expect that loan officers have to increase the pace of new originations when they are not on track to meet their quotas. Though our data does not contain the origination targets set by each lender, we infer that loan officers' volume targets are a function of mortgage lending seasonality and the lender's internal projections. As such, we expect that each lender will have their own month-by-month benchmarks that are a function of their origination volume in prior years (e.g., origination volume in March 2012 is a reasonable estimate of the volume target in March 2013).

Based on these observations, we construct a measure of whether or not loan officers at a given lender are likely to be on track to meet their performance targets. The measure relates the current month's origination volume relative to prior year's:

$$RelPerf_{i,ym} = \frac{AvgVol_{i,ym}}{AvgVol_{i,ym'}} \tag{7}$$

where $AvgVol_{i,ym}$ is the average daily volume of mortgage loans that have been issued by lending institution i and in year and month ym, excluding the last 7 days of the month. The denominator is the average daily volume of mortgage loans issued by the same lending institution in month ym', exactly one year before ym. We conjecture that the denominator of equation (7) proxies for the volume target for institution i, which is based on the performance in the same month of the previous year. We expect loan officers to be behind their volume targets when the value of $RelPerf_{i,ym}$ is small. Loan officers that are behind their volume targets would be motivated to increase their lending at the end of the month.

Indeed, origination volume at the end of the month increases by a larger amount when loan officers are more likely to miss their quotas. Figure 4 shows origination volume around the end of the month. The figure splits the sample into lenders that have values of $RelPerf_{i,ym}$ in the top quartile of lenders in a given month and lenders with values of $RelPerf_{i,ym}$ in the bottom quartile. The month-end increase in originations is substantially larger when $RelPerf_{i,ym}$ is in the bottom quartile. This provides evidence that loan officers increase the pace of new originations at the end of the month in order to meet their performance targets, and suggests that the end-of-month increase in origination volume is caused by loan officers' monthly volume quotas.

Though we attribute the increase in new originations at the end of the month to loan officers' monthly volume quotas, lenders may also have incentive to "window-dress" at month end in order to meet the criteria of regulatory exams. Specifically, lenders may increase originations to disadvantaged neighborhoods in an effort to meet the requirements of upcoming Community Reinvestment Act (CRA) examinations.[10]

We find that the increase in new originations at the end of the month is unlikely to be caused by upcoming CRA examinations. Table 2, columns (5) and (6), sort lenders by whether or not they have a CRA exam scheduled in the following month. Then, separately for the two samples, we estimate the specification in equation (6) where the dependent variable is the logarithm of daily origination volume. The month-end increase in originations occurs for both lenders that are and are not subject to CRA examinations. Surprisingly, the end-of-month increase in lending is smaller for institutions that are subject to CRA exams. Lenders that have (do not have) CRA exams increase their origination volume by 29% (54%) at the end of the month.

5 Testing for Lending Discrimination Using High-Frequency Evaluations

This section uses the modified benchmarking test developed in Section 2 to test for discrimination in the mortgage lending data.

5.1 High-frequency Benchmarking Test

We test for discrimination in mortgage lending by estimating how differences in approval rates across races change within any given month. Our tests use the following linear probability regression specification:

$$\begin{aligned} Appr_j = \delta_{lw,Black}\left(I_{lw} \times I_{Black}\right) + \delta_{fw,Black}\left(I_{fw} \times I_{Black}\right) + \delta_{lw} I_{lw} + \delta_{fw} I_{fw} + \\ + \delta_{Black} I_{Black} + BX_j + a_{ym,c} + a_{ym,i} + a_{dow} + a_{holiday} + u_j \end{aligned} \tag{8}$$

[10] CRA exams are administered by four different federal regulators—the FRB, the FDIC, the OCC, and the OTS—and they are conducted every two years for large banks and every five years for small- and medium-size lenders. Lenders know in advance the exam dates. CRA exams consist of a review of the lender's fair lending practices, designed to ensure that lenders meet the credit needs of disadvantaged communities in markets that they serve.

where the unit of observation is the individual loan application.[11] The dependent variable $Appr_j$ equals one if the loan is approved. Independent variables I_{fw} and I_{lw} equal one when the application decision is made in the first or the last week of the month, respectively. I_{Black} is equal to one for Black applicants. X_j is a vector that contains characteristics for mortgage application j and the corresponding applicant: loan amount, conforming loan status, loan type (conventional, or government guaranteed or insured, such as FHA, VA, and USDA loans), occupancy type (owner occupied or absentee), loan purpose (new purchase or refinancing), and applicant income. Year-month-county, year-month-lender, day of the week, and holiday fixed effects are $a_{ym,c}$, $a_{ym,i}$, a_{dow}, and $a_{holiday}$, respectively. The coefficients of interest, $\delta_{lw,Black}$ and $\delta_{fw,Black}$, capture the average approval rate for Black applicants in the last and first week of the month both relative to the other days of the month.

We begin the regression analysis by reporting split-sample tests of Black and white applicants that compare approval rates at the start of the month to those at the end of the month (Table 3). We find that approval rates for Black applicants are 12 percentage points larger in the last week of the month relative to the first week (column 1). Approval rates for white (and other) applicants are 8 percentage points larger in the last week (column 2). Comparing these results, Black applicants gain an additional 4 percentage point increase in approval rates over the course of the month relative to white applicants.

Next, we use the pooled sample of HMDA data to test the regression model in equation (8). The specification contains the complete set of interaction terms between Black applicants and applicants of other races (Table 3, columns 3 through 6). Because we find that the estimates are robust across specifications, we describe the most restrictive specification: column (6). The point estimate of $\delta_{lw,Black}$, the abnormal approval rate for Black applicants in the last week of the month, is equal to 2.7 ppt. The estimate of $\delta_{fw,Black}$, the abnormal approval rate in the first week of the month, is equal to -0.7 ppt. This implies that the relative likelihood of approval for

[11] Appendix Table A.3 presents estimates of logistic regression models that are analogous to the linear probability model in equation 8. We reach similar conclusions when we use logit regressions even though we are unable to include the full set of fixed effects that we include in our main OLS regressions.

Black applications increases by 3.4 ppt if the application is processed in the last seven days of the month.[12]

The estimates of the within-month difference in approval rates for Black applicants are large. For context, we estimate a baseline 6.8 ppt difference in approval rates between Black applicants and applicants of other races (the coefficient estimate on I_{Black}). This estimate is equivalent to what a conventional benchmarking test would estimate as the amount of discrimination against Black applicants. However, a conventional benchmarking test is unable to determine whether the 6.8 ppt difference is caused by racial biases or whether it reflects the unobserved heterogeneity across races. On the other hand, because our empirical design suppresses the cross-sectional variation across applicants' races, we can attribute the within-month approval gap of 3.4 ppt to loan officers' subjectivity. As such, the ratio of the within-month difference to the unconditional difference—3.4 divided by 6.8, or 50%—approximates the share of the observed racial gap in approval rates that can be attributed to subjective decision-making. In other words, we attribute at least half of the racial gap in approval rates to racial bias.

These regression tests confirm the graphical evidence that the approval gap between Black and other applicants converges over the course of the month (presented in Figure 2 and described in the Introduction). We augment this aggregate evidence by plotting how the approval gap changes over the course of the month estimated from the saturated regression model in Table 3, column (6). Figure 5(b) plots the average day-by-day residual difference in approval rates after controlling for application characteristics. The approval gap in the first seven days of the month is approximately equal to 7 ppt. The approval gap during the last seven days of the month shrinks to approximately 1 ppt on the last day of the month. Therefore, after controlling for loan characteristics, there is almost no difference in application approval rates across races on the last day of any given month.

The regressions in Table 3 also convey insight into how differences across lending institutions affect mortgage credit for Black applicants. Notably, the literature attributes much of the difference in approval rates between Black and white applicants to different lending institutions catering to

[12]This number is reported in the ”last-first (black-other)” row in the table.

different types of borrowers and that different applicants choose to apply for mortgages at certain types of institutions. We gain insight into the role of selection across institutions by examining how including lender fixed effects affects the regression estimates. Including lender fixed effects reduces the magnitude of the un-interacted coefficient on I_{Black} from -0.10 in column (2) to -0.07 in column (3). This result implies that lender fixed effects are a crucial source of unobservable variation driving the Black-white approval gap. On the other hand, lender fixed effects have a negligible effect on the within-month approval gap. The difference in approval gaps between the start and end of the month is 0.040 without and 0.035 with lender fixed effects. These results suggest that we capture a component of loan officer decision-making that exists within lenders and is invariant to institutional differences. These results are reassuring for our empirical design and the interpretation of the findings, because incentive compensation schemes are widely used across lending institutions.

Our finding that the approval gap for Black applicants is reduced at the end of the month is highly robust (see Appendix Table A.1 for the following robustness tests). The estimates are not much changed across different types of mortgage applications – new home purchases, conforming mortgages, and refinances. Controlling for the applicant's gender and including Black-year fixed effects also do not affect the estimates. Furthermore, the results are robust to replacing calendar month fixed effects with fixed effects that span the end and start of successive calendar months (e.g., January 15 to February 14). Lastly, the estimates are robust to interacting application characteristics—the loan-level control variables—with indicators for first and last week of the month. This allows the effect of application characteristics on approval decisions to flexibly change over the course of the month. For example, this regression specification would allow applicants' incomes to have a stronger effect on loan decisions at different times within a month.

Finally, Table 4 estimates our modified benchmarking test on different subsamples of the HMDA data that help exclude alternative interpretations of the results. Columns (1) and (2) show that the Black approval gap for new purchases and refinances decline by approximately the same amount, 0.035, at the end of the month. This result suggests that our findings are unlikely to be explained by borrowers strategically closing at the end of the month in order to minimize the costs of housing transitions (e.g., not having to pay rent and a mortgage at the same time). Also, the Black approval gap at the end of the month declines by a similar amount for FHA conforming and conventional loans (columns 3 and 4). This suggests that our approval gap findings cannot be explained by a financial incentives from the FHA to close at the end of the month. Specifically, for FHA mortgages originated before January 2015, the borrower is required, at the time of paying off the loan, to pay the entire interest for the month, including the part that has not yet accrued. Column (5) shows that the approval gap decline is similarly large for the sample of loans restricted to conventional new purchases, which further suggests that our findings are not caused by financial incentives to close at the end of the month. Lastly, these alternative explanations relate to the possibility that some borrowers are financially constrained and need to close at the end of the month to minimize costs. Yet we find that higher-income borrowers—that are presumably less financially constrained—have a larger decline in the end-of-month approval gap (columns 6 and 7).

5.2 Challenges to Identification of Time-Varying Discrimination

The evidence that the approval gap for Black applicants declines at the end of the month is consistent with loan officers having less scope for subjective decision-making when they have monthly volume quotas, as shown by the framework outlined in Section 2. Yet, we consider plausible challenges to the interpretation that the change in approval rates over the course of the month is evidence of discrimination.

Instead of considering specific alternative explanations for why the Black approval gap nearly disappears at the end of the month, we describe how the empirical design limits the scope for alternative theories. First, it is unlikely that the within-month variation in approval rates can be explained by variation across lenders, because lender fixed effects do not have much effect on the estimates.

Second, our empirical strategy rules out the possibility of time-invariant unobserved differences across applicant groups. Therefore, any candidate alternative explanation has to have within-month variation and also has to have differential effects on Black applicants relative to other applicants. Not only does this confine alternative explanations to factors that vary within the month, it gives us an avenue to test alternative theories. For example, suppose that the indicator variable for Black applicants reflects other unobserved characteristics, such as the riskiness of the loan, and that loan officers delay processing high-risk applications. If application risk explains the convergence in approval rates across race over the course of the month, then the observed riskiness of the loan application would explain within-month changes in approval rates. Put simply, we would expect to find that originations of observably high-risk applications submitted by Black applicants would bunch more than low-risk applications at the end of the month.

Guided by these bounds on alternative theories, we take a holistic approach to confronting alternative interpretations by examining the within-month quantity of loan originations sorted by credit scores (and applicant incomes). We study credit scores because they are possibly the most important ingredient in loan approval decisions and mortgage pricing. Credit scores would also correlate with other application attributes that could form the basis for alternative explanations. For example, credit scores directly measure the ex-ante risk of the application, and low credit score applicants would be more likely to file low-documentation applications. As section 3 describes, the data only contains credit scores for applications that are approved. As such, we study the quantity of new originations over the course of the month instead of approval rates. However, such tests are similar to testing for differences in approval rates because we have shown that mortgage demand does not vary within the month.

We find that alternative explanations related to application quality are unlikely to explain the within-month approval gap. Figures 6(a) and 7(a) plot the quantity of new originations sorted by credit scores and incomes for applications submitted by Blacks and whites, respectively. Strikingly, the volume of originations for prime-credit-score (FICO ≥ 660) and subprime (FICO < 660) Black applicants are nearly identical over the course of the month (Figure 6(a)). We would have expected to find relatively more end-of-month bunching for subprime Black applicants if the results simply reflected characteristics—such as risk—that correlate with applicants' credit scores. Also, the end-of-month bunching of originations is larger for Blacks than whites for both prime and subprime applicants (comparing the levels in Figure 6(a) to those in Figure 7(a)). The difference between Black and white originations would have been attenuated for prime applicants if characteristics related to credit scores explained the within-month approval gap.

We find similar evidence when we sort the volume of new originations into quartiles by applicant incomes (Figure 6(b) and Figure 7(b)). Testing for end-of-month bunching across applicant incomes not only fortifies evidence from sorting by credit scores but also allows us to present evidence from the full HMDA sample. We find that there is substantial end-of-month bunching for all four income quartiles. Moreover, in each corresponding quartile, the end-of-month bunching for Black applicants is significantly larger than for white applicants. These findings cast doubt on alternative explanations related to within-month variation in application quality.

5.3 Testing for Omitted Time-varying Characteristics

The empirical design that we use to estimate discrimination limits the scope of omitted variables to factors that change between the start and the end of the month. The regression tests suppress cross-sectional differences across races. However, other characteristics—including characteristics that can correlate with race—might also vary within the month. We address this possibility by assessing the extent to which observable characteristics vary within the month by using the HMDA-McDash sample and the recently available HMDA data that includes loan origination recommendations from Automated Underwriting Systems (AUS).

We start by directly testing whether certain types of loans are more or less likely to be originated over the course of the month using the HMDA-McDash merged sample. Specifically, Table 5 tests the regression specification in equation (8), but replaces the dependent variable with variables that measure loan quality for originated loans: an indicator for subprime loans (column 1), loan-to-value (LTV) ratios (column 2), an indicator for low-documentation loans (column 3), and the interest rate (column 4). We find that, of the four loan characteristics, only low-documentation loans are more likely to be originated in the last week of the month, but the effect is only weakly statistically significant. Furthermore, Black applicants with low-documentation applications are relatively *less* likely than white applicants to be originated in the last week of the month. These results, combined with the above tests, suggest that race has effects that are independent from loan characteristics—characteristics that vary by race in the cross-section—over the course of the month.[13]

Next, we compare the differences between AUS recommendations and lenders' actual loan decisions around the end of the month by using the extended HMDA data available for 2018 and 2019. Figure 8 plots the day-by-day difference between AUS recommendations made for Black and white applicants. The figure also plots the difference between Black and white approval rates. Effectively, the figure demonstrates how the AUS system evaluates Black applicants relative to how Black applicants are evaluated by lenders.

We find that the AUS evaluations do not bunch around the end of the month. Instead, the AUS system recommends that Black applicants should be approved approximately five to eight percentage points less often than white applicants. This result suggests that Black approval rates should be approximately five to eight percentage points lower when using the objective, race-neutral criteria contained in the loan application. At the same time, loan officers' actual decisions in the new version of the HMDA data exhibit substantial end-of-month bunching. The baseline Black-white approval gap in the new HMDA data is only slightly less than in the older data. The Black-white approval gap falls substantially towards the end of the month. The approval gap nearly converges with the AUS recommendations on the last day of the month.

We further show that our regression evidence on the within-month approval gap is robust to controlling for AUS recommendations to account for differences across applicants (Table 6). The

[13]To also provide an out-of-sample test of our methodology, we test for changes in end-of-month approvals for other underrepresented groups. The HMDA data contains limited demographic information but we can observe applicants' gender. Table A.2 presents regression estimates of equation (8) that tests whether the approval rate for Female applicants converges toward the end of the month. Column (6) reports the regression results with the full controls and shows that the Female approval gap shrinks from 7.4 ppt in the first week of the month to 6 ppt in the last week of the month. This suggests that there is discrimination against Female applicants.

first two columns replicate our main regressions on the new HMDA data while not controlling for the AUS recommendations. Columns (3) through (6) control for the AUS recommendations. Column (6) includes indicators for the type of ASU system used to evaluate the loan. The regressions confirm that the Black-white approval gap is 3 to 4.5 percentage points smaller at the end relative to the start of the month. Using the same sample of loans, the difference between Black and white AUS recommendations changes significantly less over the course of the month. The difference between Black and white AUS recommendations is between 1.2 and 1.8 percentage points smaller at the end relative to the start of the month.

Our findings on the AUS recommendations of Black and white applications further suggest that we identify time-varying discrimination against Black applicants. Our identification strategy is vulnerable to the possibility that there are time-varying differences across Black and white applicants that emerge around the end of the month. For example, high-quality Black applicants may be more likely than high-quality white applicants to strategically close on the last day of the month. Because the AUS recommendations do not exhibit the same level of end-of-month bunching in approval rates, and because the regression results are robust to controlling for the AUS recommendations, it is unlikely that there are unobservable differences between Black applications at the start and the end of the month.

5.4 The Timing of Originations

Though we show that application volumes are constant within any given month, a lingering challenge to our interpretation of the results is that they might reflect differences in how long it takes to finalize originations. In this section, we provide evidence that time-to-origination is unlikely to explain why the Black approval gap shrinks at the end of the month.

We start by showing that the within-month approval gap is robust to accounting for the time to action—origination or denial—on the loan (Table 7). We find that estimating equation (8) while controlling for time-to-action does not significantly affect the coefficients on the within-month approval gap (column 1). We also estimate the approval gap using sub-samples of the data sorted on time-to-action: 1 to 30 days, 31 to 60 days, 61 to 90 days and more than 90 days (columns 2 through 5, respectively). We find that the within-month approval gap for Black applicants holds in all four sub-samples. Furthermore, the estimates across sub-samples do not exhibit patterns that would suggest that the approval gap is explained by how long loan officers hold applications in inventory.

Next, we directly examine how processing times change over the course of the month and by race. Table 8 contains regressions that set time-to-origination as the dependent variable. Columns (1) and (2) regress time-to-origination in number of days on indicator variables for the first and last week of the month (and the full set of fixed effects and loan level controls). There is no difference in the time to origination for loans issued at the start versus the end of the month. Originations to Black applicants take three days longer on average. However, the within-month difference across races in time-to-origination is small. Black applications take just a half-day longer to originate when they close in the last week relative to the beginning of the month.

These results counter at least two relevant alternative interpretations of our finding that the Black approval gap reduces significantly at month-end. First, Black borrowers might prefer to close at the end of the month, potentially because they are more likely to be liquidity constrained. Indeed, some media (i.e., web searches) suggest that borrowers should close at the end of the month because it will minimize advance interest payments at the time of closing. However, closing at the end of the month might not be financially optimal because borrowers will also have to make their first mortgage payment up to a month earlier (the first payment is due on the first day of the month following the closing date).[14] Second, our findings counter the alternative explanation that loan officers might delay processing Black applications because, for example, they are more difficult to process.

These findings also address concern over institutional features of the data. In particular, the origination date is not necessarily the same as when loan approvals are decided. Our analy-

[14] Additionally, prior tests also suggest that our results are not caused by financially constrained borrowers having loan officers set the origination date at the end of the month. Our estimates of the within-month approval gap are robust to borrowers' income, LTV ratios, or other measures of creditworthiness that would correlate with financial constraints.

sis uses the origination date as a proxy for lenders' approval decisions because the data does not contain distinct records of "approval" dates. For this reason, our tests account for the delay between approval-decision and origination dates by having a wide period—the last seven days of the month—under which loan officers might be under pressure to meet origination quotas. In addition, by showing that time-to-origination is not a relevant confound, we gain additional comfort that the delay between loan decisions and origination dates does not explain our findings.

Finally, we describe the institutional details of the loan origination process that help explain why the delay between approval decisions and originations is unlikely to confound our interpretation. In the loan origination process, lending institution send borrowers a closing disclosure document that says the borrower is "cleared to close." The document confirms the mortgage conditions and closing costs. This information is provided only after confirmation that the applicant's documentation is satisfactory and the application meets the underwriting standards.[15] At the lending institution, loan "processors" prepare the mortgage documentation in collaboration with the loan officer. They set a closing date, on which the loan documentation will be signed, and the mortgage will be originated. However, loans can be denied between the "cleared to close" and the close date because institutions can recheck borrowers' financial information such as employment status during this period.

The processor tends to set the closing date within a few days of the cleared to close communication, especially for home-purchases.[16] First, sellers might pressure buyers to close quickly. Borrowers might even be bound by signed clauses that state the buying agreement is valid only if funds are provided by a certain date. Second, buyers and lenders will try to close before the

[15]When the mortgage application is submitted, the borrower provides information on her income, credit history, and the collateral for the loan. The lender then processes this information, verifying the details concerning the characteristics of the borrower, the appraisal of the property, and the Title, to make sure that the asset does not already have pending claims. These features of the underwriting process are overseen by the underwriter and tend to take several weeks. If the information provided is incomplete, or the underwriter establishes that the loan does not meet underwriting standards, the borrower may either require further information or guarantees from the prospective borrower, or deny the loan. Loan officers have some leeway in determining the outcome of these proceedings since they coordinate the interactions between the bank and the borrower.

[16]In October 2015, the Consumer Financial Protection Bureau introduced the "Know Before You Owe" rule, imposing a minimum 3-day waiting period between the receipt of the closing disclosure document and the closing date, to make sure that borrowers were granted enough time to review the documentation and contact their lawyer before committing to the mortgage.

expiration of interest rate locks. Because the majority of loans are fixed rate loans, lenders issue interest rate locks at the time of application to guarantee that loan conditions do not change during the underwriting process.

How does the lag time between clear-to-close and origination affect our empirical design and the interpretation of our results? It is important to note that loan officers have incentive to finalize originations, not just cleared-to-close, before the end of the month to meet origination volume targets. Thus, loan officers might work with the loan processor to finalize the origination more quickly but rarely ever would they want to stall the process of closing. Indeed, we find that time-to-origination is not a relevant confound for the approval gap. Therefore, so long as our empirical tests allow a sufficient time frame at the end of the month to accommodate the lag between decision and closing, there is little reason to suspect that our empirical design is biased by this feature of the data.

6 Alternative Ways to Measure Discrimination

The standard benchmarking test characterizes the null of no discrimination as there being no difference between the approval rates of minority and majority applicants. Thus, following the literature, our tests in Section 5 build on the benchmarking approach by studying changes in approval rates—we characterize the no-discrimination null as the gap in approval rates between Black and other applicants being constant over the month. This section explores two alternative ways to measure discrimination: (1) changes to the ratio of mortgage approvals across race and (2) outcome tests.

6.1 Testing for discrimination using ratios of originations

Using changes in approval rates to define discrimination has the drawback that Black approval rates are evaluated against unconditional approval rates which may also change within-month, and the regression specification used to estimate the benchmarking test does not fully control for changes to unconditional approval rates. To illustrate this challenge with the approximate numbers from the HMDA data, suppose that out of 1,000 applicants, 800 are approved (100 Black, 700 non-Black) and 200 are denied (60 Black, 140 non-Black). The approval rate for Blacks would be 62.5% and the approval rate for non-Blacks would be 83.3%, which would give an approval gap of 20.8%. Suppose that the lender decides that they need to find and approve 30% more applicants (holding constant the total number of denials) and they do so allocating the new originations proportionally across races, according to the historical shares of originations, and thus without explicitly considering race or any characteristic correlated with race. The lender would then approve 130 Black applicants and 910 non-Black applicants, and the approval gap would shrink to 18.2%.

As such, an alternative way to characterize the no-discrimination null is to test for changes in the *share* of loans that are originated to Black applicants. In the preceding example, notice that Blacks constitute one-eighth of all originations even after the lender decides to approve more loans. We evaluate this alternative null first by testing for changes in approval ratios over the course of the month and second by evaluating how the approval gap would be expected to change following a proportional change in unconditional approvals.

6.1.1 Changes in the ratio of originations

We find that the share of loans originated to Black applicants increases substantially at the end of the month. Figure 9(a) plots the share of all approved applications submitted by Black applicants on a given day. On the first day of the month, Black applicants account for approximately 4.4% of approved loans. In the last week of the month, the share increases steadily, reaching just over 6% on the last day of the month. Similarly, Appendix Figure A.4 shows the ratio of the number of originations to Black applicants over the number of originations to white applicants, on each day of the first and last week of the month. The ratio increases sharply in the last week of the month.[17]

[17] Furthermore, we study whether the within-month convergence in approval gap is sensitive to the share of Black applicants that a lender processes. Figure 9(b) shows that the within-month change in approvals occurs across the full range of lenders. The figure reports the median share of approved loans from Black applicants, along with the 25th and 75th percentile, across lenders that issued at least 10 loans per day on average over the year. The median share is close to 5.5% in the first two weeks of the month. However, it steadily increases, across the entire distribution, in the last week of the month. On the last day of the month, the median share is above 6%, the 25th percentile is approximately 4% and the 75th percentile is close to 10%.

We also test for changes to the ratio of new originations using regression analysis at the loan-level (Appendix Table A.3). We use the full sample of originations and denials to estimate:

$$I_{Black,j} = \delta_{lw,Appr}\left(I_{lw} \times Appr\right) + \delta_{fw,Appr}\left(I_{fw} \times Appr\right) + \delta_{lw} I_{lw} + \delta_{fw} I_{fw} + \\ + \delta_{Appr} Appr + BX_j + a_{ym,c} + a_{ym,i} + a_{dow} + a_{holiday} + v_j \tag{9}$$

where $I_{Black,j}$ equals one for applications submitted by Black borrowers and $Appr$ equals one for approved (originated) loans. The other variables are the same as in equation 8. The coefficient on the interaction term between the last and first week of the month, and the indicator for approvals—$\delta_{lw,Appr}$ and $\delta_{fw,Appr}$—should not be statistically different from zero under the null of no discrimination.

The regression estimates provide evidence of discrimination against Black applicants. The coefficient estimates of $\delta_{lw,Appr}$ are statistically significant and positive (columns 1 through 3 of Table A.3). This implies that the share of Black originations increases at the end of the month, even after controlling for application characteristics. The coefficient estimates are robust to different fixed effects and loan level controls, and the magnitude is large. The relative increase in the share of Black originations (out of Black and white originations) is in between 3.4% and 6.6%, a magnitude which is roughly consistent with the increase in approval rates measured in Table 3.

6.1.2 Benchmarking test under the null of a proportional change in approvals

Furthermore, we derive an alternative counterfactual for our approval gap test. In our prior tests, we evaluate the within-month change in the Black approval gap against the null of zero difference. Instead, we can evaluate the within-month change against the null that allows for a proportional change in loan originations over the course of the month.

We find that the within-month change in the Black approval gap is substantially larger than would be predicted under the null that there is a proportional change in origination volume. Figure A.5(a) plots the Black approval gap on each day at the start and end of month. We then include a line that shows how the approval gap would change if origination volume increases as much as it does in the data and the share of loans originated to Black applicants is held constant. Though the approval gap decreases at the end of the month, the size of the hypothetical decrease is significantly smaller than the approval gap decrease in the data. Indeed, the counterfactual reduction in Black approval gap is equal to just half of the reduction observed in the data 50% of what we observe in the data.

We can draw similar conclusions—that the within-month decrease in Black approval gap is too large to be explained by the increase in origination volume—from our baseline regression analysis (see equation 5 and regression estimates in Table 3). We calculate conditional estimates of the approval gap under the assumption that the approval gap changes when the approval rate changes proportionally. We estimate that the approval gap would shrink by just 1-1.5% under the assumption that origination volume increases proportionally, which accounts for just 30%-40% of the magnitudes we estimate in Table 3.

An alternative way to assess the counterfactual effect caused by the increase in all originations is to estimate how large the increase should be in order to match the reduction in Black approval gap in the data. Figure A.5(b) shows daily origination volume relative to the first of the month. The figure also includes a line for how large the increase in originations would need to be in order to match the magnitude of the reduction in the Black approval gap. Under the assumption that shares across races stay constant, the increase in origination volume at the end of the month would need to be 22% to 45% larger than in the data.

6.2 Outcome tests

Lastly, we study the ex-post performance of originated loans in order to connect our setting and empirical design to a classic outcome test. Table 9 estimates the regression specification in equation 8 but sets the dependent variable equal to one for mortgages that face a 90-day delinquency within 5 years after origination and zero otherwise.[18] Column (1) includes the full sample of originations.

[18]Appendix Table A.4 repeats the same analysis but sets the dependent variable equal to one for loans that defaulted within 5 years after origination, while in Table A.5 the dependent variable is equal to one for loans that were

The remaining columns restrict the sample to loans that might be considered more risky or more difficult to evaluate at the time of origination. Column (2) restricts the sample to subprime loans ($FICO < 660$), column (3) to high loan-to-value loans ($LTV > 80\%$), and column (4) to low documentation loans. All of the regressions include the full set of loan-level control variables available in the merged HMDA-McDash data.

The regression results show that loans originated to non-Black borrowers become delinquent at the same rate whether they are issued at the start or the end of the month. Also, loans to Black borrowers are approximately five percentage points more likely to become delinquent in each of the four samples. Interestingly, we find that delinquencies for loans to Black borrowers originated in the last week of the month tend to not be statistically different from the baseline difference for Black borrowers. Loans originated in the first week of the month are more likely to become delinquent but the difference is small in magnitude relative to the baseline.

These findings on ex-post performance help interpret the within-month approval gap that we observe in our modified benchmarking tests. First, the regressions that estimate the approval gap might not control for unobservable differences across applicants and these unobservable differences might be orthogonal to *ex-ante* observable characteristics. However, these unobservable ex-ante differences are likely to predict ex-post performance, which we can observe. Therefore, we can infer from the results of the outcome test that the within-month empirical design is not biased by unobservable differences that vary within the month. Second, by showing that there are no within-month differences in ex-post delinquency rates, these findings explicitly counter the explanation that loan officers approve riskier loans at the end of the month.

7 Lenders, Market Structure and Discrimination

Our empirical approach to estimating discrimination gives us the ability to study the effect of policies on the quantity of discrimination. In this section, we explore how two important features of

terminated (due to default or refinancing) within 5 years. We find similar results across all of these different measures of ex-post non-performance.

the mortgage lending market—financial innovation and competition across lenders—affect the estimates of the approval gap for Black applicants that can be attributed to loan officers' subjectivity.

The rise of FinTech and shadow banks has been a major trend in mortgage lending over the last decade in the United States.[19] First, we compare FinTech lenders to non-FinTech lenders. FinTech lenders rely on complex models rather than human decisions in loan approvals and have been shown to improve the speed of the origination process (Fuster et al., 2019). Second, we compare shadow banks to mortgage originators from "traditional" depository institutions. Shadow banks are less regulated than traditional banks. Buchak et al. (2018) show that having fewer regulations allows shadow banks to have a larger operating presence in underserved communities. However, the overall effect of FinTech and shadow banking on minorities' access to credit is still an open question (see e.g., Fuster et al., 2021; Bartlett et al., 2021). We use our novel empirical strategy to shed new light on this important topic.

Our analysis uses the classifications of FinTech lenders and shadow banks provided by Buchak et al. (2018).[20] Their hand-collected classification defines FinTech lenders as those that have a large online presence and that process the majority of mortgage applications online. However, the authors note that human interaction is not completely absent for FinTech lenders. Mortgage applicants have to engage with a loan officer during the closing process even if applications are submitted online. On the other hand, shadow banks have a straightforward classification. The authors define non-shadow banks as mortgage originators that also take deposits, and classify all other mortgage lenders as shadow banks. In this section, we restrict the sample of mortgage applications from HMDA to the period between 2014 and 2018, since the rise in FinTech and shadow banking is a recent phenomenon.

[19]Buchak et al. (2018) show that the market share of shadow banks has nearly doubled, from 30% in 2007 to more than 50% in 2017.

[20]Fuster et al. (2019) use loan processing times to classify FinTech lenders. Though we find this measure appealing, it would not be well-suited to our analysis. Our analysis focuses on the differences in loan processing across races. To the extent that there could be differences in the processing times across races, such an analysis would confound the classification of FinTech lenders.

To provide a test for the differences between FinTech and non-FinTech lenders, and shadow and non-shadow banks, we estimate a new regression equation:

$$
\begin{aligned}
Appr_j = {} & \delta_{lw,Black,Z}\left(I_{lw} \times I_{Black} \times Z_i\right) + \delta_{fw,Black,Z}\left(I_{fw} \times I_{Black} \times Z_i\right) + \\
& + \delta_{lw,Black}\left(I_{lw} \times I_{Black}\right) + \delta_{fw,Black}\left(I_{fw} \times I_{Black}\right) + \delta_{lw} I_{lw} + \delta_{fw} I_{fw} + \\
& + \delta_{lw,Z}\left(I_{lw} \times Z_i\right) + \delta_{fw,Z}\left(I_{lw} \times Z_i\right) + \delta_{Black} I_{Black} + \delta_{Black,Z}\left(I_{Black} \times Z_i\right) \\
& + BX_j + a_{ym,c} + a_{ym,i} + a_{dow} + a_{holiday} + e_j.
\end{aligned}
\tag{10}
$$

This specification is constructed by augmenting equation (8) with interaction terms, capturing the effect of lender characteristic Z_i on approval rates (coefficients $\delta_{lw,Z}$ and $\delta_{fw,Z}$), and on the approval rates for Black applicants (coefficients $\delta_{lw,Black,Z}$ and $\delta_{fw,Black,Z}$). The variable Z_i will consist of either a dummy equal to one for FinTech lenders, or a dummy equal to one for shadow banks.

Our results for FinTech lenders are reported in column (1) of Table 10. Interestingly, the coefficient $\delta_{lw,Z}$ is positive and statistically significant, and close to 3%. Thus, there is a higher increase in approval rates for white applicants in the last week of the month for FinTech lenders. However, the incremental effects on Black applicants are not significant. The coefficients on the individual dummies $\delta_{lw,Black,Z}$ and $\delta_{fw,Black,Z}$ are statistically indistinguishable from 0, and their difference is also not statistically significant at conventional confidence levels.

In column (2) we turn to shadow banks. The coefficient $\delta_{lw,Z}$ is negative, while $\delta_{fw,Z}$ is positive, thus taking the opposite signs as the baseline month-end fluctuation in approval rates. Both are statistically significant. For white applicants, the increase in approval rates is smaller for shadow, rather than non-shadow, banks by roughly 2.8%. Moreover, the marginal increase in approval rates for Black applicants ($\delta_{lw,Black,Z} - \delta_{fw,Black,Z}$) is smaller by 1.8% (significant at the 95% confidence level) for shadow banks. Thus, shadow banks seem to experience smaller month-end effects in general, and in particular for Black applicants. Consistent with this finding, the coefficient on the interaction between the Black applicant and the shadow bank dummy reveals that in normal days (outside the last and first week of the month) approval rates for Black applicants are 5% higher for shadow banks than for traditional depository institutions. Thus, shadow banks appear to have smaller month-end effects, and a more equal treatment of white and Black applicants. This is consistent with our conjecture that more efficient institutions would leave less room for taste-based discrimination. However, quite surprisingly, this does not seem to be the case for FinTech lenders, which experience larger month-end spikes in approvals, and are subject to a similar degree of taste-based discrimination as non-FinTech lenders.

There are two potential explanations for this result. First, the criteria currently used in the literature to identify FinTech companies might be flawed, or noisy. Second, while up-to-date statistical and machine learning models may provide loan officers with more accurate insights on credit risk, FinTech lenders may still leave loan officers free to blend model insights with their own subjective judgment. Both explanations have potentially important implications for our current assessment and understanding of the role of FinTech lenders in the mortgage market, and deserve further investigation in future research.

We now turn to the relationship between lending market structure and month-end fluctuations in approval rates, and ultimately the relationship between market structure and discrimination.

We first focus on local (county-level) market concentration. To measure concentration, in each county and year in our sample we construct two measures based on the number of mortgage originations by lender using the HMDA data: the share of total mortgages originated by the 4 institutions with the largest number of originations, and the Herfindahl-Hirschman Index (HHI) based on within county shares of mortgage originations in the previous year.

Our results are reported in Table 10. In columns (3) and (4) we report estimates from the specification in equation (10) (with Z now a dummy equal to one for above median top 4 share or above median HHI counties), which we use to test the triple interaction effects between last and first week of the month, Black applicant and market concentration. We find that these effects are not statistically significant. Differences between high and low concentration counties in the month-end approval rates increase for whites are significant, but quantitatively negligible. If we interpret concentrations as a proxy for local competition among lenders (Cetorelli and Strahan, 2006), then this evidence suggests that local competition among lenders does not translate into higher competition and smaller room for subjective decision making and biases at the level of individual loan officers.

An alternative channel through which we may capture the effects of local competition is by comparing large and small lenders. Small lenders are naturally more concentrated in a small number of markets, and might compete more fiercely in those markets. Thus, we construct a proxy for lender size, equal to the average number of mortgage originations per year. We compute the overall average annual mortgage origination volume for each lender and split lenders into two groups, depending on whether their size is above or below the median across lenders in the United States. In column (6) of Table 10 we find that the magnitude of the change in the approval gap for Black applicants around month-end is indistinguishable between large and small size lenders. Overall, our results suggest that, even when market forces increase competition across institutions, discriminatory practices persist *within* institutions.

8 Conclusions

Tests for discrimination are often unconvincing because subject groups tend to have different unobserved characteristics. We show that high-frequency evaluations can limit the scope of the omitted variable problem when there is variation in the degree to which decision-makers rely on subjective evaluations. Under the null that decision-makers do not engage in discrimination, and assuming that the applicant pool is constant, a decrease in the degree of subjectivity should have no impact on the likelihood of favorable decisions for minority subjects relative to majority subjects. A reduction in disparate treatment for minority subjects would instead reveal the presence of discrimination.

We use our approach to provide new evidence of discrimination in mortgage lending in the U.S. First, we document an "end-of-month effect" in which the volume of new mortgage originations increases by over 150% relative to the start of the month. This increase is caused by the performance incentives of loan officers – the fact that loan officers have monthly performance targets. Next, we show that the within-month pattern of loan approvals varies by the mortgage applicants' race. The gap in approval rates between white and Black applicants attenuates by half at the end of the month, when loan officers need to approve more applications to meet their performance targets. There are no observable within-month racial patterns in application volume and no within-month patterns of application quality that could explain the results.

Our findings have important policy implications for the distribution of credit in consumer credit markets. Legislation such as the Community Reinvestment Act and the Equal Credit Opportunity Act has been implemented over the past several decades to counteract historical inequities in credit access (e.g., red-lining; Appel and Nickerson, 2016; Aaronson et al., 2017). A crucial aspect of such legislation is that it intends to modify the behavior of lending institutions. We show that patterns of discriminatory behavior by loan officers exist *within-institution* and such behavior is not mitigated by important features of the market structure of lending markets—namely FinTech, institution size, and competition across lenders. This suggests that policies targeted toward institutions will have limited effects so long as individuals use their discretion to allocate credit.

Seeing as institution-level policies do not eliminate biases held by individuals, it calls into question what policies would be effective. In accordance with classic economic theories of discrimination (Becker, 1957), competition reduces taste-based discrimination. Such competition occurs in the labor market for loan officers. Loan officers have to meet monthly performance targets otherwise they would have less compensation and risk being fired. However, discrimination can persist if there are barriers to entry in the labor market. Indeed, loan officers need at least a bachelor's degree in a field related to finance or business, they have to obtain and maintain a license, and they often obtain clients through referrals by real estate agents.

Two recommendations emerge from our study. First, the collection of high-frequency data on evaluations, combined with our approach, can be used to estimate the amount of discrimination across a variety of contexts and markets. Second, enhancing the data collection to go beyond the institution and down to individual decision-makers can provide further insight into the factors that determine discrimination. Such data can be used by researchers and policy-makers, as well as by consumers, for instance when shopping for credit in the mortgage market.

References

Aaronson, D., Hartley, D., Mazumder, B., 2017. The effects of the 1930s HOLC “redlining” maps. Working Paper .

Agarwal, S., Ben-David, I., 2012. Do loan officers’ incentives lead to lax lending standards? Working Paper .

Ambrose, B. W., Conklin, J. N., Lopez, L. A., 2020. Does borrower and broker race affect the cost of mortgage credit? The Review of Financial Studies Forthcoming.

Appel, I., Nickerson, J., 2016. Pockets of poverty: The long-term effects of redlining. Working Paper .

Arnold, D., Dobbie, W., Yang, C. S., 2018. Racial bias in bail decisions. The Quarterly Journal of Economics 133, 1885–1932.

Ayres, I., 2002. Outcome tests of racial disparities in police practices. Justice Research and Policy 4, 131–142.

Bandiera, O., Barankay, I., Rasul, I., 2007. Incentives for managers and inequality among workers: Evidence from a firm-level experiment. The Quarterly Journal of Economics 122, 729–773.

Bartlett, R., Morse, A., Stanton, R., Wallace, N., 2021. Consumer-lending discrimination in the fintech era. Journal of Financial Economics Forthcoming.

Beck, T., Behr, P., Madestam, A., 2018. Sex and credit: Is there a gender bias in lending? Journal of Banking and Finance 87.

Becker, G. S., 1957. The Economics of Discrimination. The University of Chicago Press, Chicago, IL, second ed.

Berg, T., Burg, V., Gombović, A., Puri, M., 2020a. On the rise of fintechs: Credit scoring using digital footprints. The Review of Financial Studies 33, 2845–2897.

Berg, T., Puri, M., Rocholl, J., 2020b. Loan officer incentives, internal rating models, and default rates. Review of Finance 24, 529–578.

Berkovec, J. A., Canner, G. B., Gabriel, S. A., Hannan, T. H., 1994. Race, redlining, and residential mortgage loan performance. The Journal of Real Estate Finance and Economics 9, 263–294.

Berkovec, J. A., Canner, G. B., Gabriel, S. A., Hannan, T. H., 1998. Discrimination, competition, and loan performance in FHA mortgage lending. Review of Economics and Statistics 80, 241–250.

Bertrand, M., Duflo, E., 2017. Field experiments on discrimination. In: *Handbook of economic field experiments*, Elsevier, vol. 1, pp. 309–393.

Bhutta, N., Hizmo, A., 2020. Do minorities pay more for mortgages? The Review of Financial Studies Forthcoming.

Bhutta, N., Hizmo, A., Ringo, D., 2021. How much does racial bias affect mortgage lending? evidence from human and algorithmic credit decisions. Working Paper .

Bohren, J. A., Haggag, K., Imas, A., Pope, D. G., 2020. Inaccurate statistical discrimination, Working Paper.

Bohren, J. A., Imas, A., Rosenberg, M., 2019. The dynamics of discrimination: Theory and evidence. American economic review 109, 3395–3436.

Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A., 2016. Stereotypes. The Quarterly Journal of Economics 131, 1753–1794.

Buchak, G., Jørring, A., 2017. Does competition reduce racial discrimination in lending? Working Paper .

Buchak, G., Matvos, G., Piskorski, T., Seru, A., 2018. Fintech, regulatory arbitrage, and the rise of shadow banks. Journal of Financial Economics 130, 453–483.

Butler, A. W., Mayer, E. J., Weston, J., 2019. Discrimination in the auto loan market. Working Paper .

Cao, Y., Fisman, R., Lin, H., Wang, Y., 2020. Target setting and allocative inefficiency in lending: Evidence from two Chinese banks. Working Paper .

Cetorelli, N., Strahan, P. E., 2006. Finance as a barrier to entry: Bank competition and industry structure in local U.S. markets. Journal of Finance 61, 437–461.

Chen, D. L., Moskowitz, T. J., Shue, K., 2016. Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. The Quarterly Journal of Economics 131, 1181–1242.

Cole, S., Kanz, M., Klapper, L., et al., 2015. Incentivizing calculated risk-taking: Evidence from an experiment with commercial bank loan officers. Journal of Finance 70, 537–575.

Cortés, K., Duchin, R., Sosyura, D., 2016. Clouded judgment: The role of sentiment in credit origination. Journal of Financial Economics 121, 392–413.

Demiroglu, C., Ozbas, O., Silva, R., Ulu, M. F., 2021. Do physiological and spiritual factors affect economic decisions? Journal of Finance Forthcoming.

Dobbie, W., Liberman, A., Paravisini, D., Pathania, V. S., 2020. Measuring bias in consumer lending. Review of Economic Studies Forthcoming.

Engelberg, J., Gao, P., Parsons, C. A., 2012. Friends with money. Journal of Financial Economics 103, 169–188.

Fisman, R., Paravisini, D., Vig, V., 2017. Cultural proximity and loan outcomes. American Economic Review 107, 457–92.

Fisman, R., Sarkar, A., Skrastins, J., Vig, V., 2020. Experience of communal conflicts and inter-group lending. Journal of Political Economy 128, 3346–3375.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., Walther, A., 2021. Predictably unequal? The effects of machine learning on credit markets. Journal of Finance Forthcoming.

Fuster, A., Plosser, M., Schnabl, P., Vickery, J., 2019. The role of technology in mortgage lending. The Review of Financial Studies 32, 1854–1899.

Goldin, C., Rouse, C., 2000. Orchestrating impartiality: The impact of "blind" auditions on female musicians. American Economic Review 90, 715–741.

Gravelle, H., Sutton, M., Ma, A., 2010. Doctor behaviour under a pay for performance contract: Treating, cheating and case finding? The Economic Journal 120, F129–F156.

Hertzberg, A., Liberti, J. M., Paravisini, D., 2010. Information and incentives inside the firm: Evidence from loan officer rotation. The Journal of Finance 65, 795–828.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., Mullainathan, S., 2018. Human decisions and machine predictions. The Quarterly Journal of Economics 133, 237–293.

Ladd, H. F., 1998. Evidence on discrimination in mortgage lending. Journal of Economic Perspectives 12, 41–62.

Lanning, J., 2021. Testing models of economic discrimination using the discretionary markup of indirect auto loans. Working Paper .

Larkin, I., 2014. The cost of high-powered incentives: Employee gaming in enterprise software sales. Journal of Labor Economics 32, 199–227.

Li, J., Hurley, J., DeCicca, P., Buckley, G., 2014. Physician response to pay-for-performance: Evidence from a natural experiment. Health Economics 23, 962–978.

Liebman, J. B., Mahoney, N., 2017. Do expiring budgets lead to wasteful year-end spending? Evidence from federal procurement. American Economic Review 107, 3510–49.

Montoya, A. M., Parrado, E., Solıs, A., Undurraga, R., 2019. Gender discrimination in the consumer credit market: Experimental evidence. Working Paper .

Munnell, A. H., Tootell, G. M., Browne, L. E., McEneaney, J., 1996. Mortgage lending in Boston: Interpreting HMDA data. The American Economic Review pp. 25–53.

Murfin, J., Petersen, M., 2016. Loans on sale: Credit market seasonality, borrower need, and lender rents. Journal of Financial Economics 121, 300–326.

Murphy, K. J., 2000. Performance standards in incentive contracts. Journal of Accounting and Economics 30, 245–278.

Oyer, P., 1998. Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality. The Quarterly Journal of Economics 113, 149–185.

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., Ramachandran, V., Barghouty, P., Phillips, C., Shroff, R., et al., 2020. A large-scale analysis of racial disparities in police stops across the United States. Nature Human Behaviour pp. 1–10.

Qian, J., Strahan, P. E., Yang, Z., 2015. The impact of incentives and communication costs on information production and use: Evidence from bank lending. The Journal of Finance 70, 1457–1493.

Rosen, R. J., 2011. Competition in mortgage markets: The effect of lender type on loan characteristics. Economic Perspectives 35, 2–21.

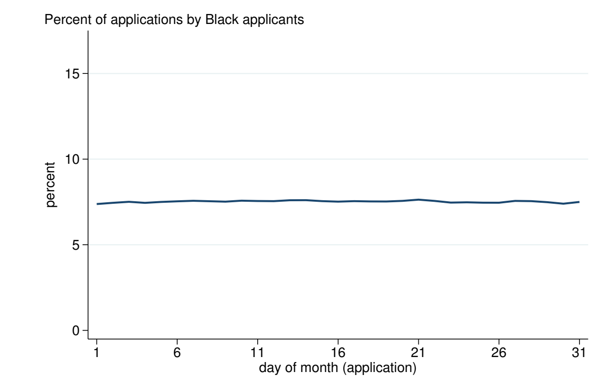
Simoiu, C., Corbett-Davies, S., Goel, S., et al., 2017. The problem of infra-marginality in outcome tests for discrimination. The Annals of Applied Statistics 11, 1193–1216.

Tang, H., 2019. Peer-to-peer lenders versus banks: Substitutes or complements? The Review of Financial Studies 32, 1900–1938.

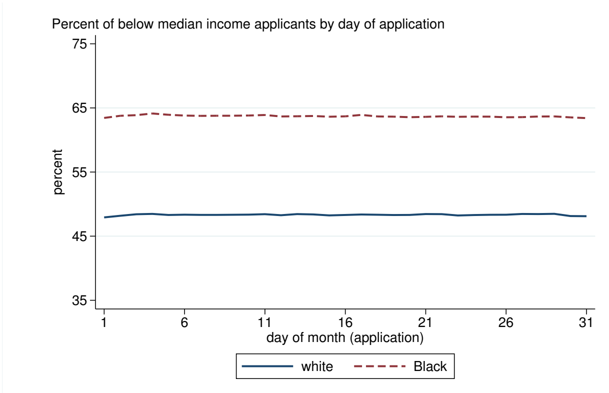
Tootell, G. M., 1996. Redlining in Boston: Do mortgage lenders discriminate against neighborhoods? The Quarterly Journal of Economics 111, 1049–1079.

Tzioumis, K., Gee, M., 2013. Nonlinear incentives and mortgage officers' decisions. Journal of Financial Economics 107, 436–453.

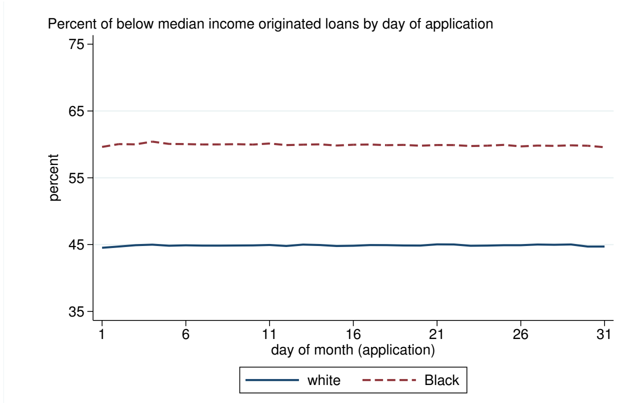
Figures



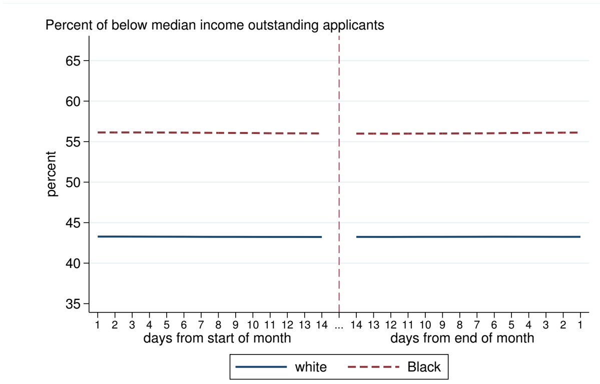
(a) Share of Black Applicants



(b) Originated Loans



(c) All Applications



(d) Low Income, Outstanding Applications

Figure 3: Panel (a) of the figure shows the average fraction of applications by Blacks (out of all applications) based on the day of the month in which the application was filed. Panel (b) shows the fraction of applications by white and Black applicants with income below the county median in the year, based on the day of the month in which the applications were filed. Panel (c) shows the fraction of loans originated to white and Black applicants with income below the county median in the year, based on the day of the month in which the original applications were filed. Panel (d) shows the fraction of applications with income below the county median in the year, out of all applications outstanding on each day of the month, for white and Black applicants. The results in all panels of the figure are based on the HMDA data from January 1994 to December 2018.

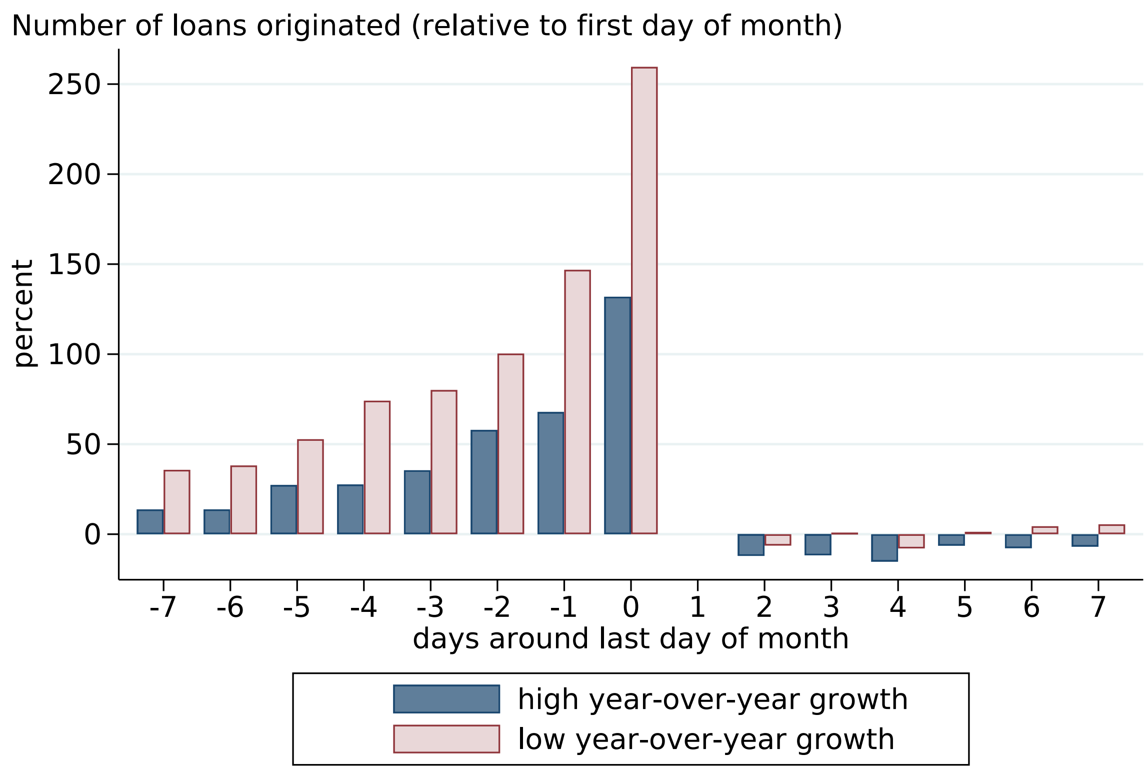
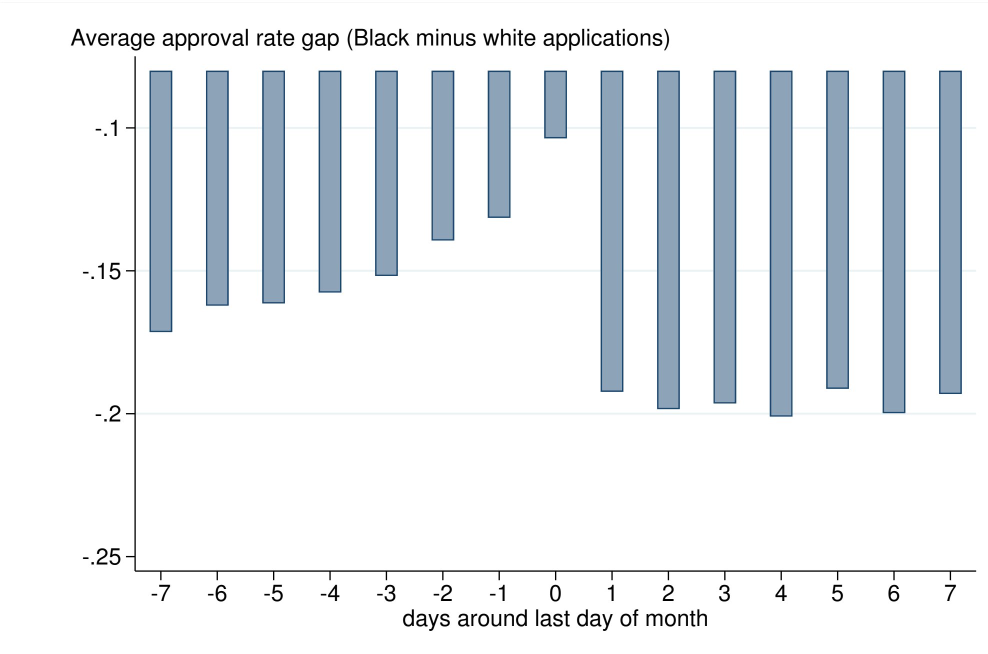
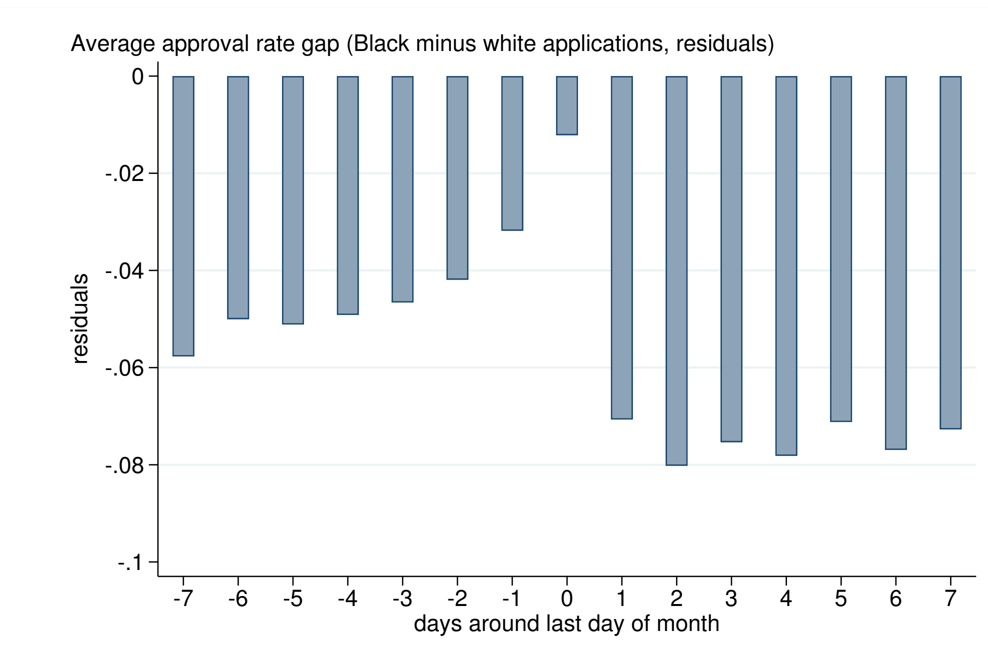


Figure 4: The figure shows average percentage abnormal daily loan origination volume in the U.S., separately for lenders that, on each day, have high (top quartile across all lenders in the year) and low (bottom quartile across all lenders in the year) performance growth relative to the same month in the previous year (defined according to equation 7). The darker bars are for lenders with high performance growth, while the lighter bars are for lenders with low performance growth. Abnormal origination volume is reported for the last eight days of the month, and the first seven days of the following month, and is computed with respect to volume on the first day of the month. The results are based on the HMDA data from January 1994 to December 2018.

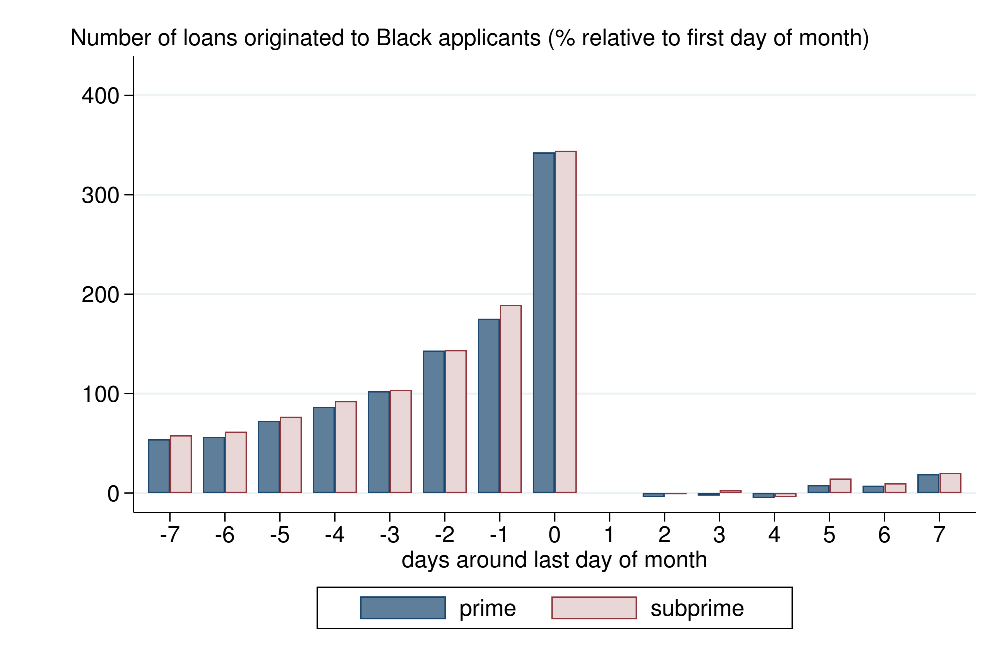


(a) Approval Gap (Raw Data)

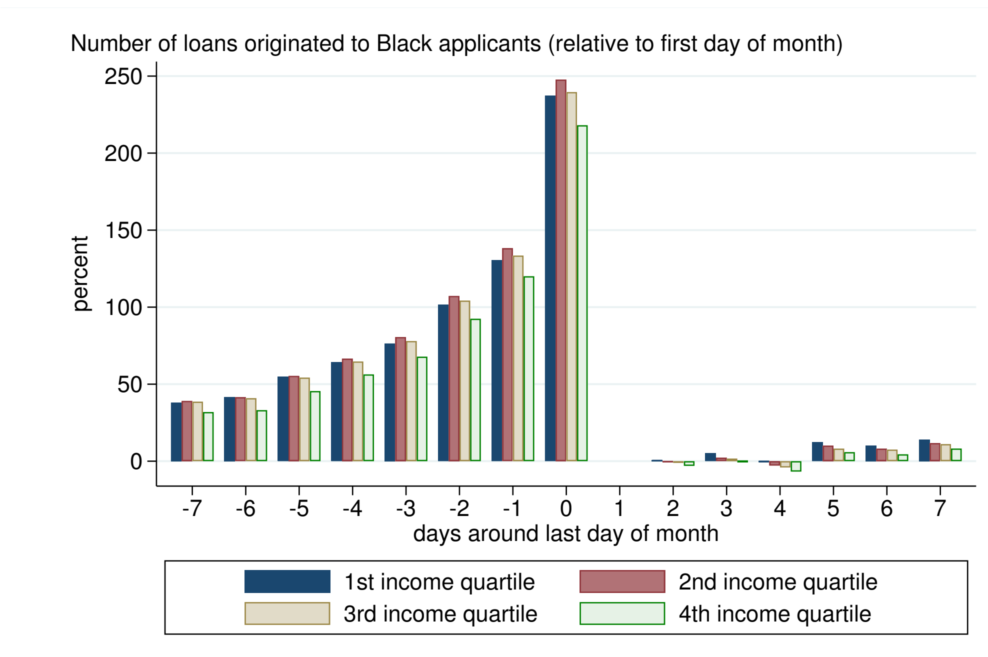


(b) Approval Gap (Residuals)

Figure 5: The figure shows the approval gap for Blacks. Panel (a) reports the difference between the fraction of approved loans, out of all approved and denied loans in the U.S., for Blacks minus the one for whites, on each of the last eight days of the month and the first seven days of the following month. Panel (b) reports residual differences in approval rates after controlling for loan applicant characteristics. The day-by-day difference in approval rates is attained by first estimating the complete specification of equation 8 (see column 6 of Table 3), but omitting the dummies for actions taken in the first and last week of the month. We then average regression residuals on each day, separately for Black and white applicants, and compute the difference between the daily averages to estimate the (controlled) approval gap on each day. Estimates are based on a 5% sample of the HMDA data from January 1994 to December 2018.

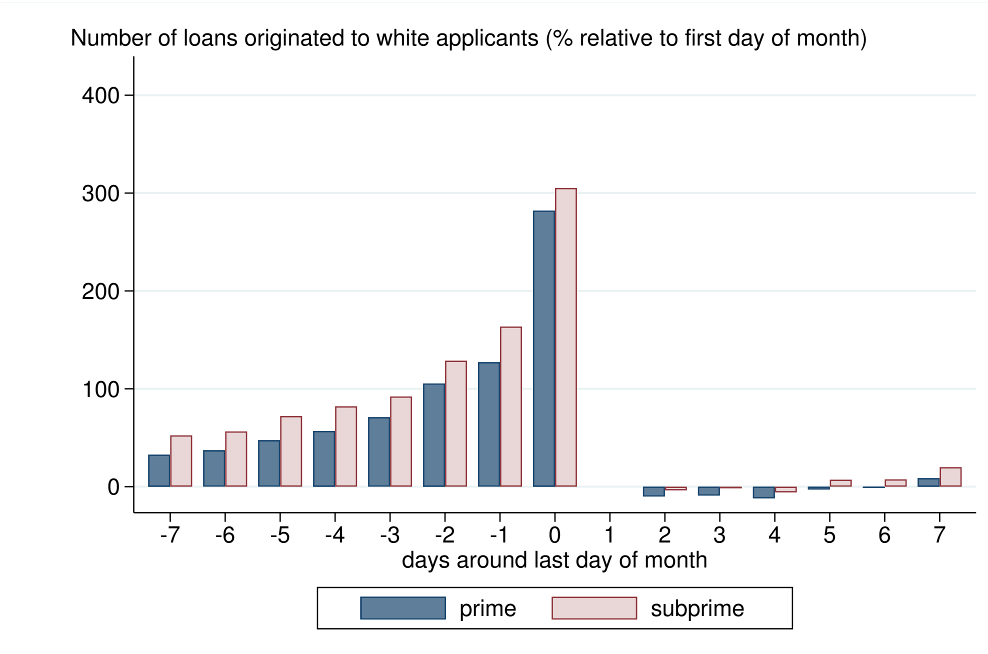


(a) Effects for Prime/Subprime Black Applicants

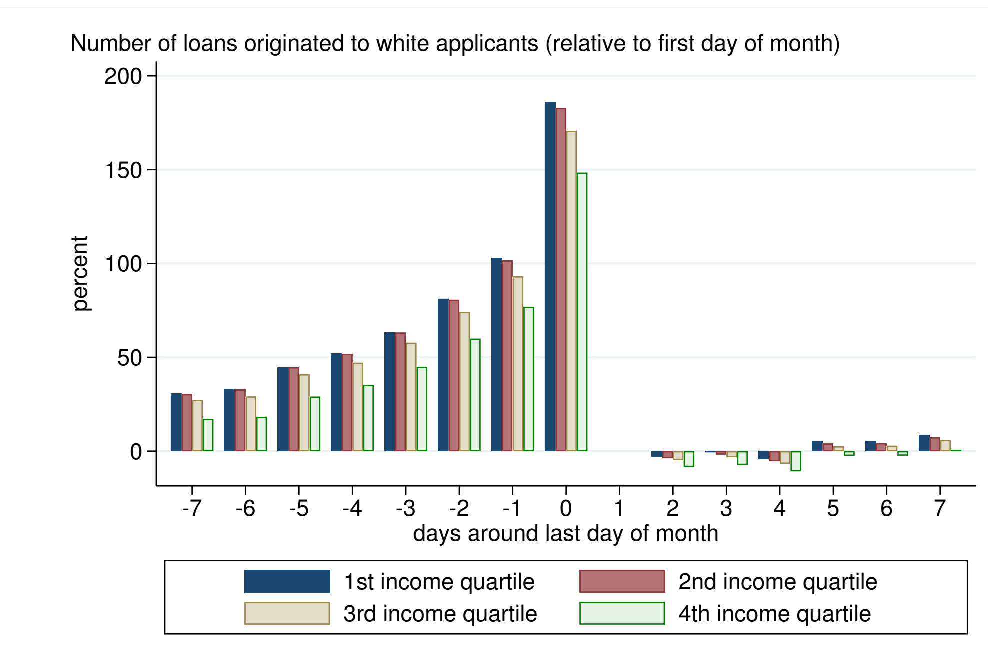


(b) Effects by Income Quartile Black Applicants

Figure 6: Panel (a) of the figure shows average percentage abnormal daily loan origination volume in the U.S., for Blacks with prime (660 or higher) and subprime FICO. Abnormal volume is reported for the last eight days of the month, and the first seven days of the following month, and is computed with respect to loan origination volume on the first day of the following month for each applicant group. Estimates are based on the merged sample of HMDA and Black Knight McDash data from 1994 to 2018. Panel (b) of the figure shows average percentage abnormal daily loan origination volume separately for Blacks belonging to different income quartiles within county. The results in this panel are based on the HMDA sample from January 1994 to December 2018.



(a) Effects for Prime/Subprime White Applicants



(b) Effects by Income Quartile White Applicants

Figure 7: Panel (a) of the figure shows average percentage abnormal daily loan origination volume in the U.S., for whites with prime (660 or higher) and subprime FICO. Abnormal volume is reported for the last eight days of the month, and the first seven days of the following month, and is computed with respect to loan origination volume on the first day of the following month for each applicant group. Estimates are based on the merged sample of HMDA and Black Knight McDash data from 1994 to 2018. Panel (b) of the figure shows average percentage abnormal daily loan origination volume separately for Blacks belonging to different income quartiles within county. The results in this panel are based on the HMDA sample from January 1994 to December 2018.

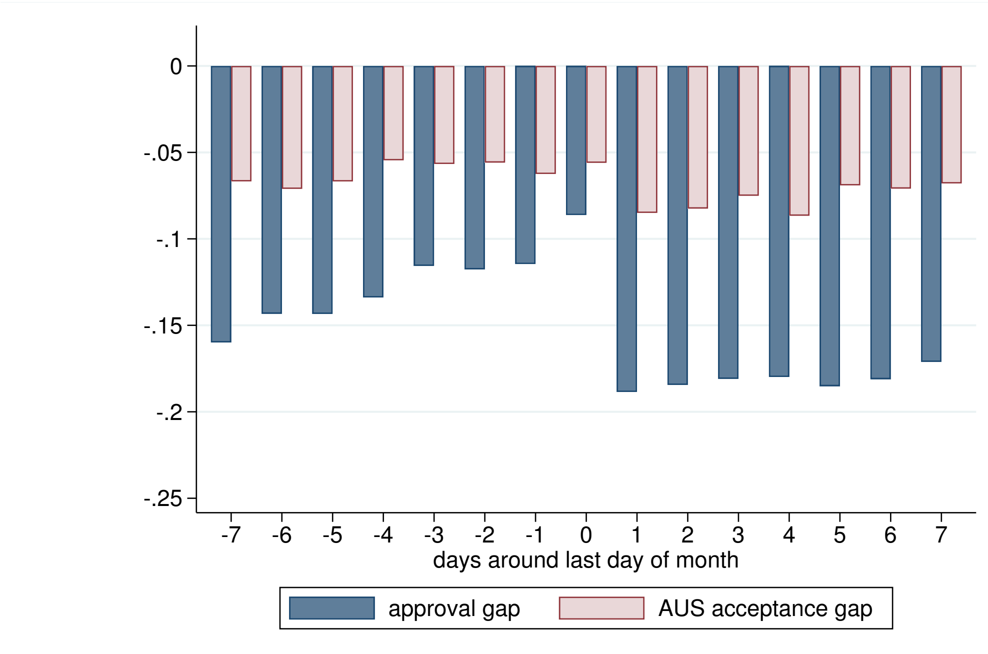
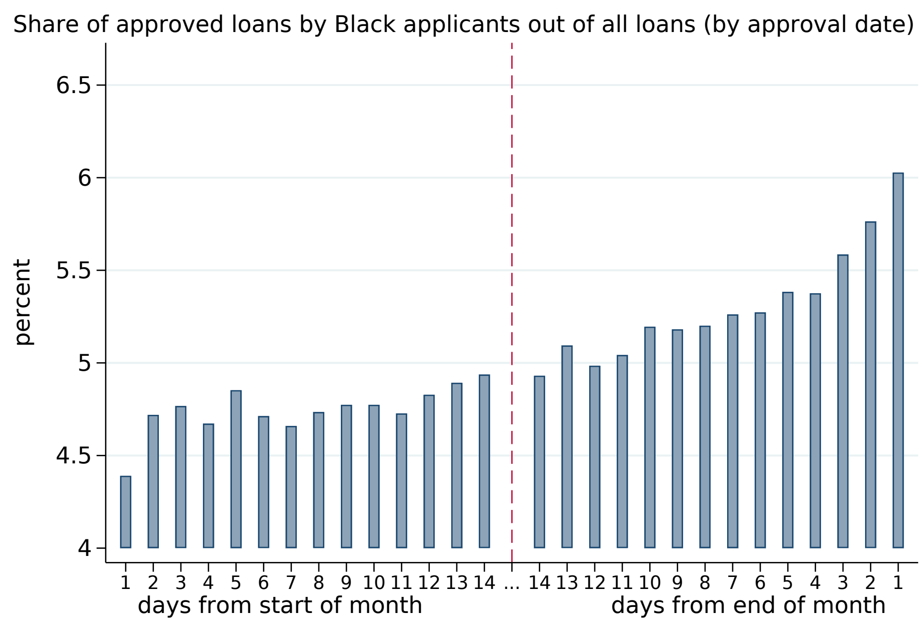
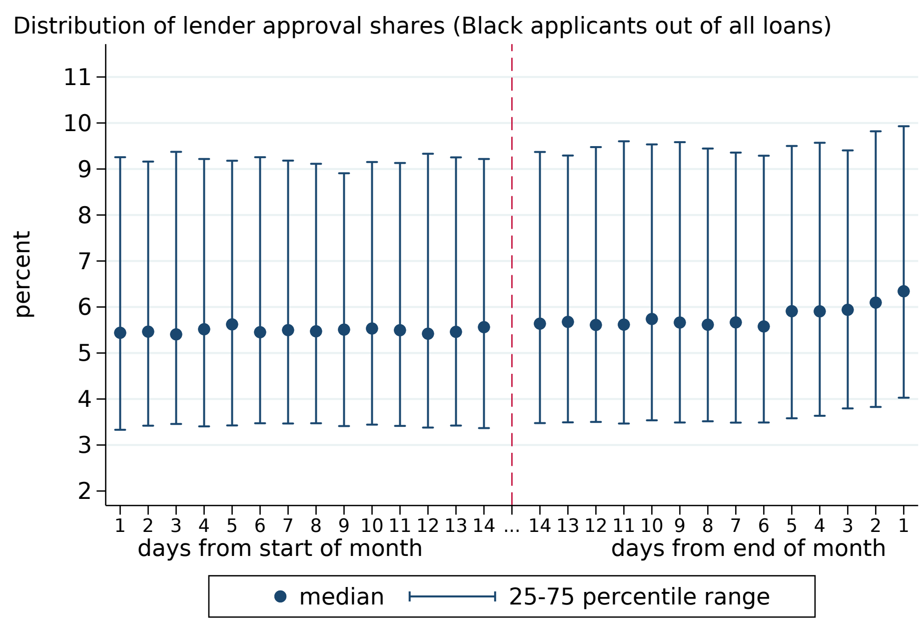


Figure 8: The figure shows the difference between the fraction of approved loans, out of all approved and denied loans in the U.S., for Blacks minus the one for whites, on each of the last eight days of the month and the first seven days of the following month (approval gap). The figure also shows the day-by-day difference in approval recommendations from the Automated Underwriting Systems (AUS). The data comes from the "new" HMDA that covers loan applications in 2018 - 2019.



(a) United States



(b) Individual Lenders

Figure 9: Panel (a) of the figure shows the share of approved applications from Black applicants, out of all approved applications on each day in the two weeks before and after the end of the month in the United States. The results in this panel are based on a 5% random sample of the HMDA data from January 1994 to December 2018. Panel (b) shows the distribution (median, 25th percentile and 75th percentile) of the share of approved applications from Blacks at the lender level on each day, for all lenders that originate on average at least 10 loans per-day. The results in this panel are based on the whole HMDA sample from January 1994 to December 2018.

Tables

	HMDA 1994-2018			HMDA-McDash 1994-2018			New HMDA 2018-2019		
	White	Black	Other Race	White	Black	Other Race	White	Black	Other Race
Observations	16,882,565	1,850,734	6,483,189	29,909,913	2,556,790	7,128,756	2,481,456	260,872	756,638
Shares of Applications	66.95%	7.34%	25.71%	-	-	-	70.92%	7.46%	21.62%
Approval Rate	80.72%	63.25%	69.00%	-	-	-	86.58%	76.35%	82.00%
Average Loan Amount ($ 1,000)	156.21	125.12	176.69	198.34	162.88	231.67	245.24	218.36	294.30
Share Low Income Apps	46.02%	59.75%	47.80%	49.46%	64.73%	50.27%	49.83%	62.12%	48.58%
Share Conforming	87.24%	89.23%	87.09%	94.58%	97.15%	91.65%	-	-	-
Share Primary Residence	91.08%	91.40%	90.25%	91.77%	92.49%	90.54%	92.29%	94.32%	89.79%
Share New Purchases	41.64%	42.73%	32.95%	48.36%	54.34%	44.87%	56.78%	59.67%	50.95%
Share of Originated Loans	73.72%	5.73%	20.55%	75.54%	6.46%	18.00%	73.51%	6.35%	20.14%
Share Below Prime	-	-	-	11.98%	31.05%	12.02%	11.33%	26.32%	10.24%
Share High-LTV	-	-	-	37.30%	58.36%	31.97%	56.42%	75.66%	52.20%
AUS Approval Rate	-	-	-	-	-	-	84.54%	77.65%	83.44%

Table 1: US-level summary statistics for the historical HMDA loan applications data over the years from 1994 to 2018 (5% sample of the data), the merged sample of originated loans from HMDA and McDash over the years from 1994 to 2018, and for the new HMDA data layout for years 2018 and 2019 (20% sample of the data). *Share of Applications* is the share of applications belonging to each group out of the total, *Share Low Income Apps* is the fraction of applicants with income below the median in the county and year of the application, within each group, *Share Conforming* is the fraction of conforming loans, within each group, *Share Primary Residence* is the fraction of loans for which the collateral is the primary residence of the applicant, within each group, *Share New Purchases* is the fraction of loans for new house purchase, within each group, *Approval Rate* is the fraction of approved loans, within each group, and *Share of Originations* is the fraction of originated loans belonging to each group of applicants, out of the total. For the merged HMDA-McDash sample, *Share Below Prime* is the fraction of loans, within each race group, issued to applicants with FICO below 660, while *Share High-LTV* is the fraction of loans issued to applicants with origination LTV higher than 80%. In the new HMDA sample covering 2018 and 2019, the *AUS Approval Rate* is the fraction of loans that, for each race group, was recommended for approval by the Automated Underwriting Systems (AUS) used by the lender (for example, the Fannie Mae's Automated Underwriting System).

	U.S.-Level					
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Num Loans)	log(Num Loans)	log(Num Loans)	log($ Amount)	log(Num Loans) CRA	log(Num Loans) non-CRA
$lastweek$	0.25***	0.31***	0.31***	0.36***	0.18***	0.36***
	(0.052)	(0.016)	(0.012)	(0.013)	(0.013)	(0.013)
$firstweek$	-0.22***	-0.15***	-0.15***	-0.14***	-0.11***	-0.18***
	(0.052)	(0.016)	(0.012)	(0.013)	(0.013)	(0.013)
Holiday FE	NO	YES	YES	YES	YES	YES
Day-of-Week FE	NO	YES	YES	YES	YES	YES
Month FE	NO	YES	NO	NO	NO	NO
Month-Year FE	NO	NO	YES	YES	YES	YES
$last - first$	0.47	0.46	0.46	0.50	0.29	0.54
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	9131	9131	9131	9131	9131	9131
r2	0.0064	0.91	0.95	0.95	0.95	0.94

Table 2: The table reports regression estimates of the abnormal loan originations volume in the last and first week of the month (see equation 6). In columns (1) to (3) the dependent variable is the log of the number of originations per day in the United States. In column (4), the dependent variable is the total dollar amount of loan originations per day in the United States. In columns (5) and (6) the dependent variable is the log number of originations, respectively, for lenders subject to CRA examination and not subject to CRA examination. $lastweek$ and $firstweek$ are dummies equal to one, respectively, in the first and last week of the month. The different columns present estimates based on different choices of lender and seasonality fixed effects. Estimates are based on the sample of HMDA mortgage originations from 1994 to 2018.

	(1) approval Black	(2) approval White & Other	(3) approval All	(4) approval All	(5) approval All	(6) approval All
$lastweek$	0.090***	0.057***	0.057***	0.048***	0.044***	0.043***
	(0.0043)	(0.0037)	(0.0037)	(0.0035)	(0.0031)	(0.0031)
$firstweek$	-0.032***	-0.022***	-0.022***	-0.021***	-0.019***	-0.020***
	(0.0028)	(0.0025)	(0.0025)	(0.0021)	(0.0018)	(0.0017)
$black$			-0.12***	-0.10***	-0.070***	-0.068***
			(0.0069)	(0.0062)	(0.0053)	(0.0043)
$black \times lastweek$			0.036***	0.032***	0.028***	0.027***
			(0.0021)	(0.0023)	(0.0022)	(0.0023)
$black \times firstweek$			-0.010***	-0.0081***	-0.0073***	-0.0072***
			(0.0015)	(0.0014)	(0.0024)	(0.0015)
log(income)				0.095***	0.073***	0.071***
				(0.0061)	(0.0039)	(0.0036)
log(loan amount)				0.031***	0.0089***	0.0076***
				(0.0075)	(0.0029)	(0.0023)
is conforming				0.13***	0.092***	0.090***
				(0.0093)	(0.0057)	(0.0057)
Other Loan-Level Controls	NO	NO	NO	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES	YES
Month-Year FE	YES	YES	YES	YES	YES	NO
County FE	YES	YES	YES	YES	YES	NO
Lender FE	NO	NO	NO	NO	YES	NO
Month-Year-County	NO	NO	NO	NO	NO	YES
Month-Year-Lender	NO	NO	NO	NO	NO	YES
$last - first$	0.12	0.079	0.079	0.068	0.063	0.063
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black)$			0.13	0.11	0.099	0.097
$p - value\ (black)$			0.0000	0.0000	0.0000	0.0000
$last - first\ (black - other)$			0.046	0.040	0.035	0.034
$p - value\ (black - other)$			0.0000	0.0000	0.0000	0.0000
N	1440405	18200483	19641147	18464497	18463245	17898939
r2	0.050	0.034	0.041	0.092	0.23	0.32

Table 3: The table reports individual loan-level regression estimates of abnormal approval rates in the last and first week of the month (see equation 8). The dependent variable is a dummy that takes value 1 if a loan application is approved and 0 if it is denied. $lastweek$ and $firstweek$ are dummies equal to one, respectively, if the decision on the application is taken in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. In columns (1) and (2), the sample is restricted to, respectively, Blacks and whites or other race applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) approval New Purchases	(2) approval Refinancing	(3) approval Conforming	(4) approval Conventional	(5) approval Conventional New Purchases	(6) approval Above County Median Income	(7) approval Below County Median Income
$black$	-0.070***	-0.056***	-0.067***	-0.069***	-0.073***	-0.074***	-0.059***
	(0.0036)	(0.0059)	(0.0043)	(0.0049)	(0.0047)	(0.0046)	(0.0044)
$lastweek$	0.035***	0.051***	0.043***	0.042***	0.034***	0.037***	0.049***
	(0.0021)	(0.0049)	(0.0032)	(0.0033)	(0.0025)	(0.0029)	(0.0033)
$firstweek$	-0.014***	-0.025***	-0.020***	-0.019***	-0.012***	-0.017***	-0.022***
	(0.0011)	(0.0026)	(0.0018)	(0.0017)	(0.0012)	(0.0015)	(0.0019)
$black \times lastweek$	0.028***	0.025***	0.026***	0.025***	0.025***	0.029***	0.022***
	(0.0018)	(0.0041)	(0.0022)	(0.0026)	(0.0021)	(0.0017)	(0.0025)
$black \times firstweek$	-0.0063***	-0.0096***	-0.0067***	-0.0065***	-0.0045**	-0.0096***	-0.0050**
	(0.0015)	(0.0025)	(0.0015)	(0.0017)	(0.0020)	(0.0024)	(0.0018)
Loan-Level Controls	YES	YES	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES	YES	YES	YES
$last - first$	0.049	0.076	0.063	0.061	0.046	0.054	0.072
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black)$	0.084	0.11	0.096	0.092	0.075	0.092	0.099
$p - value\ (black)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black - other)$	0.035	0.034	0.033	0.031	0.029	0.038	0.027
$p - value\ (black - other)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	7046904	8705784	16392963	15969777	5665143	8592046	8759206
r2	0.33	0.35	0.32	0.32	0.35	0.31	0.36

Table 4: The table re-estimates the specification in column (6) (including holiday, day-of-week, month-year-county and month-year-lender fixed effects) of Table 3 on subsamples of the HMDA dataset based on loan characteristics. All regressions are based on individual loan-level regression estimates of abnormal origination rates in the last and first week of the month (see equation 8). The dependent variable is a dummy that takes value 1 if a loan application is originated and 0 if it is denied. For all columns the regression specification includes all the controls used in column 6 of Table 3. In columns (1) and (2), the sample is restricted, respectively, to new purchase loans and refinancings. In columns (3), (4) and (5), the sample is restricted, respectively, to conforming and conventional loans, and conventional loans for new house purchases. Finally, in columns (6) and (7), we restrict the sample, respectively, to applicants with income above and below the median in their county. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$ ($lastday$ and $firstday$), and for the difference of the interaction coefficients for Black applicants, along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) Subprime (FICO $<$ 660)	(2) Origination LTV (%)	(3) Low Documentation	(4) Origination Interest Rate
$lastweek$	0.00029	0.0062	0.042*	0.00096
	(0.00079)	(0.070)	(0.023)	(0.0011)
$firstweek$	0.00059	-0.15***	-0.0012	0.00016
	(0.00042)	(0.028)	(0.0012)	(0.00094)
$black$	0.12***	3.09***	0.0014	0.0013
	(0.0066)	(0.24)	(0.0034)	(0.0019)
$black \times lastweek$	0.0020	-0.32***	-0.018**	-0.00076
	(0.0014)	(0.081)	(0.0068)	(0.00068)
$black \times firstweek$	0.0031*	0.098*	-0.00042	0.00022
	(0.0016)	(0.054)	(0.0011)	(0.00071)
LTV	0.0016***		0.00032	0.00043
	(0.00035)		(0.00055)	(0.00044)
log(income)	-0.011***	-9.02***	0.013**	0.000068
	(0.0039)	(0.84)	(0.0051)	(0.0081)
log(loan amount)	-0.041***	19.8***	-0.028	-0.049***
	(0.0061)	(1.60)	(0.022)	(0.012)
is conforming	-0.0061**	7.98***	0.053***	-0.036***
	(0.0030)	(0.67)	(0.012)	(0.0062)
FICO 620:659		-1.59***	0.019	-0.00015
		(0.46)	(0.025)	(0.010)
FICO 660:719		-2.83***	0.043	-0.020*
		(0.56)	(0.040)	(0.011)
FICO 720:759		-3.99***	0.051	-0.036***
		(0.71)	(0.051)	(0.011)
FICO 760:799		-7.34***	0.052	-0.051***
		(0.74)	(0.054)	(0.0099)
FICO $\geq$ 800		-10.6***	0.060	-0.065***
		(0.67)	(0.052)	(0.0081)
Other Loan-Level Controls	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES
$last - first$	-0.0003	0.1500	0.0430	0.0008
$p - value$	0.7600	0.0830	0.0690	0.6800
$last - first\ (black)$	-0.0014	-0.2600	0.0250	-0.0002
$p - value\ (black)$	0.6100	0.0077	0.1500	0.9100
$last - first\ (black - other)$	-0.0011	-0.4200	-0.0180	-0.0010
$p - value\ (black - other)$	0.6600	0.0005	0.0130	0.3700
N	27701585	27701585	18125835	24187005
r2	0.26	0.51	0.51	0.78

Table 5: The table reports regression estimates of the difference in characteristics between mortgages originated in the last and first week of the month. The dependent variables are a dummy equal to one for subprime loans (with FICO $<$ 660, see column (1)), the mortgage LTV at origination (column (2)), a dummy equal to one for mortgages for which the applicant provided low documentation (column (3)), and the mortgage interest rate at origination (column (4)). $lastweek$ and $firstweek$ are dummies equal to one in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with their *p-value*s. Standard errors are clustered by lender and year. Estimates are based on the merged sample of HMDA and Black Knight McDash data from 1994 to 2018.

	(1) approval	(2) approval	(3) approval	(4) approval	(5) approval
$black$	-0.080***	-0.022***	-0.042***	-0.017***	-0.016***
	(0.0049)	(0.0021)	(0.0054)	(0.0018)	(0.0017)
$lastweek$	0.028***	0.023***	0.023***	0.021***	0.021***
	(0.0019)	(0.0016)	(0.0017)	(0.0016)	(0.0016)
$firstweek$	-0.018***	-0.015***	-0.016***	-0.014***	-0.014***
	(0.0018)	(0.0016)	(0.0021)	(0.0021)	(0.0021)
$black \times lastweek$	0.027***	0.021***	0.020***	0.019***	0.018***
	(0.0027)	(0.0024)	(0.0024)	(0.0023)	(0.0022)
$black \times firstweek$	-0.018***	-0.013***	-0.015***	-0.013***	-0.012***
	(0.0028)	(0.0024)	(0.0025)	(0.0026)	(0.0025)
$AUS\ approved$			0.41***	0.32***	0.36***
			(0.022)	(0.025)	(0.019)
Standard Loan-Level Controls	YES	YES	YES	YES	YES
New Loan-Level Controls (2018-2019 HMDA)	NO	YES	NO	YES	YES
AUS Type Controls	NO	NO	NO	NO	YES
Holiday FE	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES	YES
$last - first$	0.045	0.038	0.039	0.035	0.035
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ \ (black)$	0.090	0.073	0.074	0.066	0.066
$p - value\ \ (black)$	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ \ (black - other)$	0.045	0.035	0.035	0.031	0.031
$p - value\ \ (black - other)$	0.0000	0.0000	0.0000	0.0000	0.0000
N	3544630	3544630	2819531	2537705	2537705
r2	0.25	0.38	0.30	0.36	0.36

Table 6: The table reports individual loan-level regression estimates of abnormal approval rates in the last and first week of the month (see equation 8), based on the new HMDA data (which contain more detailed information on applicants and loan characteristics) available for years 2018 and 2019. *Standard Loan-Level Controls* stands for the set of controls used in column 6 of Table 3, for the HMDA sample covering the years from 1994 to 2018. *New Loan-Level Controls* stands for the new controls available for the latter sample. In particular, we include in the regressions dummies for quintiles of debt-to-income ratio, applicants' FICO and loan-to-value ratio. In columns (3), (4) and (5) we include a dummy equal to one when we observe that the Automated Undersriting System (AUS) used by the lender recommended approval of the loan. *AUS Type Controls* is a set of dummies selecting the different types of AUS models used by lenders. The dependent variable is a dummy that takes value 1 if a loan application is approved and 0 if it is denied. $lastweek$ and $firstweek$ are dummies equal to one, respectively, if the decision on the application is taken in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 20% random sample of the new HMDA data for 2018 and 2019.

	(1) approval	(2) approval 1-30 Days TTA	(3) approval 31-60 Days TTA	(4) approval 61-90 Days TTA	(5) approval $>$ 90 Days TTA
$lastweek$	0.038***	0.053***	0.027***	0.027***	0.023***
	(0.0026)	(0.0033)	(0.0031)	(0.0030)	(0.0034)
$firstweek$	-0.019***	-0.016***	-0.017***	-0.019***	-0.017***
	(0.0018)	(0.0017)	(0.0017)	(0.0027)	(0.0027)
$black$	-0.056***	-0.072***	-0.043***	-0.037***	-0.030***
	(0.0039)	(0.0050)	(0.0032)	(0.0035)	(0.0032)
$black \times lastweek$	0.023***	0.023***	0.021***	0.015***	0.015***
	(0.0020)	(0.0025)	(0.0020)	(0.0023)	(0.0037)
$black \times firstweek$	-0.0094***	-0.0043**	-0.015***	-0.0053	-0.0056
	(0.0017)	(0.0019)	(0.0027)	(0.0033)	(0.0036)
log(TTA)	0.11***				
	(0.0069)				
Loan-Level Controls	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES	YES
$last - first$	0.057	0.069	0.044	0.046	0.040
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black)$	0.089	0.096	0.080	0.065	0.061
$p - value\ (black)$	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black - other)$	0.032	0.027	0.036	0.020	0.021
$p - value\ (black - other)$	0.0000	0.0000	0.0000	0.0000	0.0000
N	16503563	9038519	4907587	1420393	1332868
r2	0.35	0.40	0.29	0.37	0.42

Table 7: The table reports individual loan-level regression estimates of abnormal approval rates in the last and first week of the month (see equation 8), controlling for time to action (TTA), defined as the number of days between the application date and the date in which action (approval or denial) of the loan is taken. In column (1), the log of TTA is included as a control. In columns (2) to (5), the sample is restricted to loans with TTA, respectively, between 1 and 30 days, between 31 and 60 days, between 61 and 90 days, and longer than 91 days. The dependent variable is a dummy that takes value 1 if a loan application is approved and 0 if it is denied. $lastweek$ and $firstweek$ are dummies equal to one, respectively, if the decision on the application is taken in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) Time-To-Orig.	(2) Time-To-Orig.	(3) Time-To-Den.	(4) Time-To-Act. Black	(5) Time-To-Act. White & Other
$lastweek$	-0.49**	-0.51***	0.73***	-0.51***	-0.55***
	(0.15)	(0.15)	(0.12)	(0.15)	(0.15)
$firstweek$	-0.55***	-0.54***	0.73***	-0.91***	-0.49***
	(0.13)	(0.13)	(0.16)	(0.19)	(0.14)
$black$		3.08***	-0.85***		
		(0.39)	(0.16)		
$black \times lastweek$		0.26*	0.13		
		(0.13)	(0.14)		
$black \times firstweek$		-0.26**	-0.065		
		(0.11)	(0.16)		
$denial$				-13.1***	-12.3***
				(1.72)	(1.40)
$denial \times lastweek$				1.61***	1.51***
				(0.22)	(0.22)
$denial \times firstweek$				1.73***	1.40***
				(0.28)	(0.24)
Loan Level Controls	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES	YES
$last - first$	0.059	0.024	-0.0050		
$p - value$	0.79	0.91	0.98		
$last - first\ (black)$		0.54	0.19		
$p - value\ (black)$		0.028	0.28		
$last - first\ (black - other)$		0.52	0.19		
$p - value\ (black - other)$		0.0031	0.24		
$last - first\ (den. - orig.)$				-0.12	0.11
$p - value\ (den. - orig.)$				0.69	0.74
N	11924971	11924971	3915189	1011794	11116857
r2	0.29	0.29	0.40	0.45	0.31

Table 8: The table reports individual loan-level regression estimates of abnormal mortgage processing time in the last and first week of the month. Our estimates are based on the regression specification introduced in equation 8. In columns (1) and (2), the sample is restricted to originated loans, and the dependent variable is the time to origination, defined as the number of days between the application date and the origination date. In column (3), the sample is restricted to denied applications, and the dependent variable is the time to denial. In columns (4) and (5), the sample contains both originations and denials, respectively, for Black applicants only and for white or other race applicants only. $lastweek$ and $firstweek$ are dummies equal to one, respectively, if the decision on the application is taken in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. $denial$ is a dummy equal to one for denied applications. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants (and the difference of the interaction coefficients for denials in column (4)), along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) 5-Year Delinquency	(2) 5-Year Delinquency FICO < 660	(3) 5-Year Delinquency LTV $> 80\%$	(4) 5-Year Delinquency Low Docs
$black$	0.045***	0.055***	0.051***	0.048***
	(0.0044)	(0.0038)	(0.0029)	(0.0067)
$lastweek$	0.00012	0.0022*	-0.000075	-0.00034
	(0.00074)	(0.0013)	(0.00086)	(0.0012)
$firstweek$	0.00098	0.0022	0.0022*	0.0014
	(0.00074)	(0.0015)	(0.0010)	(0.00099)
$black \times lastweek$	0.0015	0.0039*	0.00037	-0.0072
	(0.0019)	(0.0021)	(0.0020)	(0.0052)
$black \times firstweek$	0.0052**	0.0047	0.0047*	0.0078**
	(0.0018)	(0.0033)	(0.0023)	(0.0030)
Loan-Level Controls	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES
$last - first$	-0.00086	0.00001	-0.0023	-0.0018
$p - value$	0.53	1.00	0.22	0.30
$last - first\ (black)$	-0.0046	-0.00077	-0.0066	-0.017
$p - value\ (black)$	0.25	0.89	0.068	0.029
$last - first\ (black - other)$	-0.0037	-0.00078	-0.0044	-0.015
$p - value\ (black - other)$	0.22	0.84	0.15	0.038
N	18797106	3065611	5870489	5278008
r2	0.26	0.25	0.26	0.34

Table 9: The table reports regression estimates of the difference in performance between mortgages originated in the last and first week of the month. The dependent variable is a dummy equal to one for mortgages for which we observe a 90-days delinquency within 5 years after origination. In column (2), the sample is restricted to subprime loans (FICO < 660). In column (3) the sample is restricted to high loan-to-value loans (LTV$> 80\%$), and in column (4) to loans with low documentation. $lastweek$ and $firstweek$ are dummies equal to one in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with their *p-value*s. Standard errors are clustered by lender and year. Estimates are based on the merged sample of HMDA and Black Knight McDash data from 1994 to 2018.

	(1) approval $Z = I_{Fintech}$	(2) approval $Z = I_{ShadowBank}$	(3) approval $Z = I_{HighTop4}$	(4) approval $Z = I_{HighHHI}$	(5) approval $Z = I_{LargeBank}$
$black$	-0.093***	-0.11***	-0.063***	-0.063***	-0.063***
	(0.0043)	(0.0038)	(0.0046)	(0.0046)	(0.0046)
$lastweek$	0.035***	0.054***	0.045***	0.045***	0.045***
	(0.0031)	(0.0081)	(0.0030)	(0.0030)	(0.0030)
$firstweek$	-0.020***	-0.024***	-0.021***	-0.021***	-0.021***
	(0.0034)	(0.0037)	(0.0019)	(0.0019)	(0.0019)
$black \times lastweek$	0.027***	0.031***	0.026***	0.026***	0.026***
	(0.0034)	(0.0035)	(0.0022)	(0.0023)	(0.0022)
$black \times firstweek$	-0.0074***	-0.0087***	-0.0079***	-0.0074***	-0.0079***
	(0.0024)	(0.0024)	(0.0015)	(0.0014)	(0.0015)
$black \times Z$	-0.013	0.050***	-0.011***	-0.012***	-0.011***
	(0.0089)	(0.0076)	(0.0033)	(0.0033)	(0.0033)
$lastweek \times Z$	0.029***	-0.020**	-0.0035***	-0.0040***	-0.0035***
	(0.010)	(0.0086)	(0.0010)	(0.0010)	(0.0010)
$firstweek \times Z$	-0.0069	0.0075*	0.0034***	0.0036***	0.0034***
	(0.0046)	(0.0042)	(0.00088)	(0.00084)	(0.00088)
$black \times lastweek \times Z$	0.0072	-0.0095**	0.0014	0.00063	0.0014
	(0.0049)	(0.0046)	(0.0020)	(0.0019)	(0.0020)
$black \times firstweek \times Z$	0.00063	0.0083*	0.0027	0.0016	0.0027
	(0.0033)	(0.0049)	(0.0020)	(0.0022)	(0.0020)
Loan-Level Controls	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES	YES
$last - first\ (black)$	0.089	0.12	0.100	0.10	0.100
$p - value\ (black)$	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black, Z)$	0.096	0.10	0.099	0.099	0.099
$p - value\ (black, Z)$	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black, Z - noZ)$	0.0066	-0.018	-0.0013	-0.0010	-0.0013
$p - value\ (black, Z - noZ)$	0.15	0.022	0.63	0.71	0.63
N	4159023	4354784	17898971	17898971	17898971
r2	0.26	0.26	0.32	0.32	0.32

Table 10: The table reports individual loan-level regression estimates of abnormal approval rates for Black applicants in the last and first week of the month, interacted with lender and local market characteristics (see equation 10). $lastweek$ and $firstweek$ are dummies equal to one, respectively, in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. $I_{Fintech}$ is a dummy equal to one for loan applications submitted to Fintech lenders. $I_{ShadowBank}$ is a dummy equal to one for loan applications submitted to shadow banks. $I_{HighTop4}$ is a dummy equal to one in counties where the share of the top 4 originators is above and below median (across counties in the United States in the same year). $I_{HighHHI}$ is a dummy equal to one in counties where the HHI index based on lenders origination shares is above and below median (across counties in the United States in the same year). $I_{LargeBank}$ is a dummy equal to one for lenders with size above median. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants across lender groups, along with their *p-value*s. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 2014 to 2018 in columns (1) and (2), and a 5% random sample of the HMDA data from 1994 to 2018 in columns (3), (4) and (5).

Appendix to:

Using High-Frequency Evaluations to Estimate Discrimination: Evidence from Mortgage Loan Officers

(intended for online publication)

A.I Identifying Time-Varying Discrimination

We show how under the assumptions in **Assumption Set (A)**, favorable decision probabilities for whites and Blacks are different. The probability, conditional on race and other observable characteristics, is equal to:

$$P(Y|R,X) = \frac{P(Y,R|X)}{P(R|X)} = \frac{P(Y,R,Z_H|X)+P(Y,Z_L|X)}{P(R|X)} = \frac{P(Y|R,Z_H,X)P(Z_H|R,X)P(R|X)+P(Y|R,Z_L,X)P(Z_L|R,X)P(R|X)}{P(R|X)}$$

$$= P(Y|Z_H,R|X)P(Z_H|R,X) + P(Y|Z_L,R,X)P(Z_L|R,X)$$

where $Z \in \{Z_L, Z_H\}$ is a binary unobservable characteristic, X is a vector of observable characteristics and $R \in \{W, B\}$ is the applicants' race (white or Black). Then, the difference in favorable decision probabilities for whites and Blacks is equal to:

$$P(Y|W,X) - P(Y|B,X) =$$

$$= [P(Y|W,Z_H,X)P(Z_H|W,X) + P(Y|W,Z_L,X)P(Z_L|W,X)] - [P(Y|Z_H,B,X)P(Z_H|B,X) + P(Y|Z_L,B,X)P(Z_L|B,X)]$$

$$= P(Y|Z_H,X)[P(Z_H|W,X) - P(Z_H|B,X)] + P(Y|Z_L,X)[P(Z_L|W,X) - P(Z_L|B,X)] > 0$$

where $P(Y|Z_H,X) = P(Y|W,Z_H,X) = P(Y|B,Z_H,X)$ and $P(Y|Z_L,X) = P(Y|W,Z_L,X) = P(Y|B,Z_L,X)$ from the assumption of no discrimination, and $P(Z_H|W,X) - P(Z_H|B,X) > 0$ and $P(Z_L|W,X) - P(Z_L|B,X) < 0$ due to the assumption of higher unobservable quality characteristics for whites.

We now turn to the comparison of the favorable decision probabilities for whites and Blacks at the beginning and the end of the month. The difference in the probability between whites and Blacks is equal to:

$$P(Y|W,X,T) - P(Y|B,X,T) =$$

$$= P(Y|Z_H,W,X,T)P(Z_H|W,X,T) + P(Y|Z_L,W,X,T)P(Z_L|W,X,T) - P(Y|Z_H,B,X,T)P(Z_H|B,X,T) - P(Y|Z_L,B,X,T)P(Z_L|B,X,T)$$

$$= P(Y|Z_H,W,X,T)[P(Z_H|W,X,T) - P(Z_H|B,X,T)] + P(Y|Z_L,W,X,T)[P(Z_L|W,X,T) - P(Z_L|B,X,T)]$$

$$= P(Y|Z_H,X,T)[P(Z_H|W,X,T) - P(Z_H|B,X,T)] + P(Y|Z_L,X,T)[P(Z_L|W,X,T) - P(Z_L|B,X,T)]$$

where $T \in \{Start, End\}$. Exploiting the calculations above, we can then derive the properties of the change in difference between favorable decision probabilities at the beginning and the end of the month:

$$[P(Y|W,X,End) - P(Y|B,X,End)] - [P(Y|W,X,Start) - P(Y|B,X,Start)] =$$

$$= P(Y|Z_H,X,End)[P(Z_H|W,X,End) - P(Z_H|B,X,End)] + P(Y|Z_L,X,End)[P(Z_L|W,X,End) - P(Z_L|B,X,End)]$$

$$-P(Y|Z_H, X, Start)[P(Z_H|W, X, Start) - P(X_H|B, X, Start)] - P(Y|Z_L, X, Start)[P(Z_L|W, X, Start) - P(Z_L|B, X, Start)]$$

$$= [P(Y|Z_H, X, Start) - P(Y|Z_H, X, Start)][P(Z_H|W, X) - P(Z_H|B, X)] + [P(Y|Z_L, X, Start) - P(Y|Z_L, X, Start)][P(Z_L|W, X) - P(Z_H|B, X)] = 0$$

where we set $P(Z_H|W, X, Start) = P(Z_H|W, X, End)$ based on the assumption that applications quality does not change over the month, while $P(Y|Z_H, X, T) = P(Y|W, Z_H, X, T) = P(Y|B, Z_H, X, T)$ and $P(Y|Z_L, X, T) = P(Y|W, Z_L, X, T) = P(Y|B, Z_L, X, T)$, based on the no discrimination assumption. Thus, the rejection of the null that the counterpart of the equation above in the data is equal to zero, leads to a rejection of the no discrimination assumption, conditionally on not having changes in application quality between the beginning and the end of the month.

A.II Additional Figures and Tables

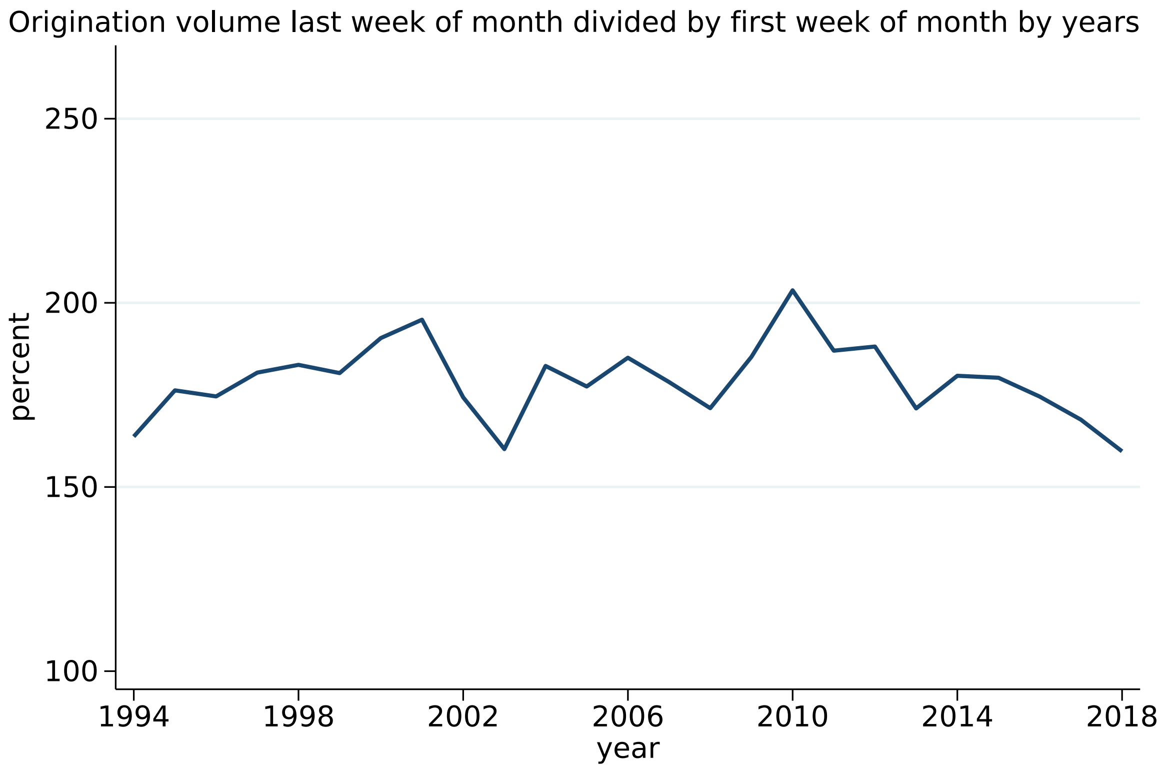


Figure A.1: The figure shows the ratio of average mortgage origination volume in the last week of the month over average mortgage origination volume in the first week of the month, for each year over the period from 1994 to 2018. The evidence is based on the HMDA data from January 1994 to December 2018.

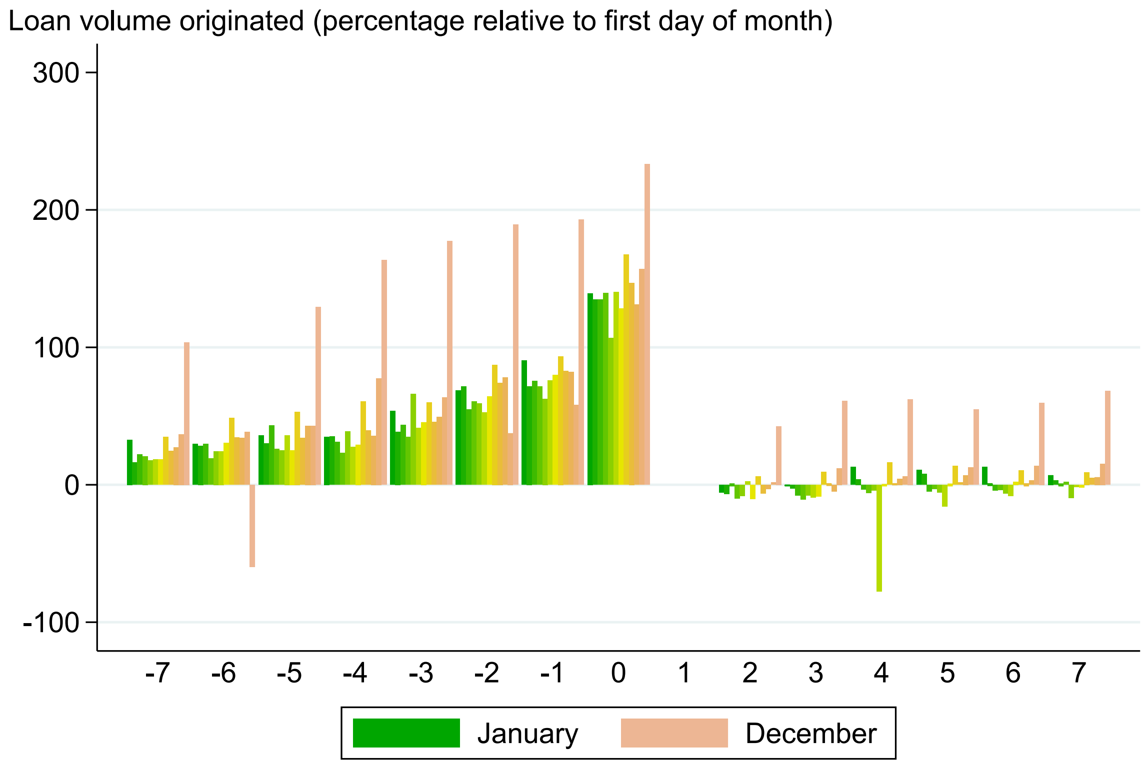


Figure A.2: The figure shows average percentage abnormal daily loan origination volume (measured as number of originations) in the U.S., for the last eight days of the month, and the first seven days of the following month, separately for each calendar month (January to December) over the sample period from January 1994 to December 2018. Abnormal volume is computed with respect to loan origination volume on the first day of the following month. The evidence is based on the HMDA data from January 1994 to December 2018.

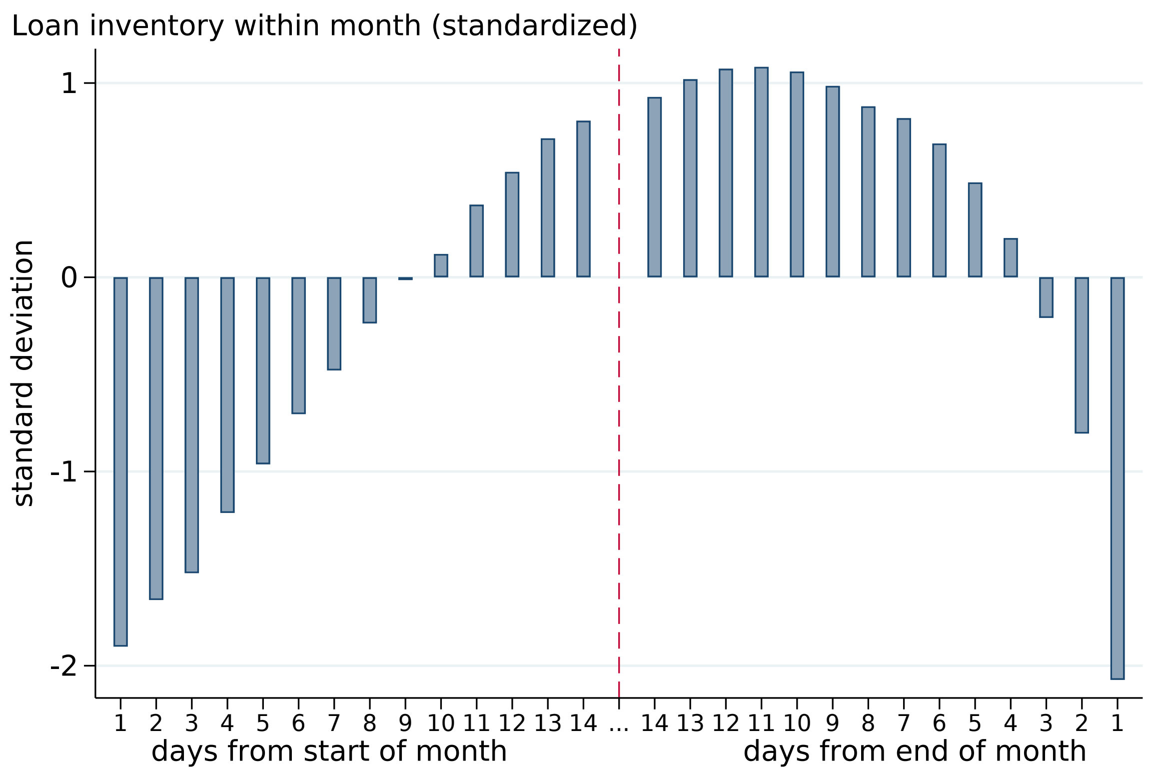


Figure A.3: This figure shows the within-month fluctuation, at the level of the entire United States, in the average number of loan applications in inventory (awaiting a decision) by day of the month. Inventory size is standardized to have mean of zero and standard deviation of one. The evidence is based on the HMDA data from January 1994 to December 2018.

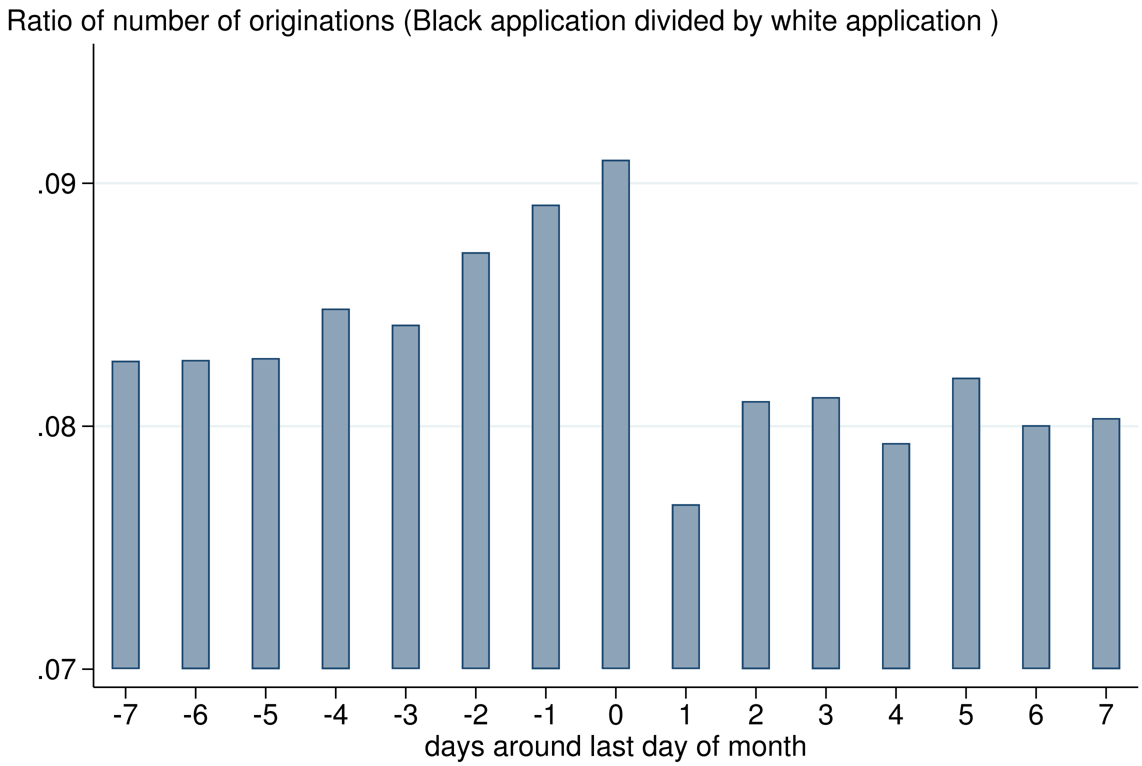
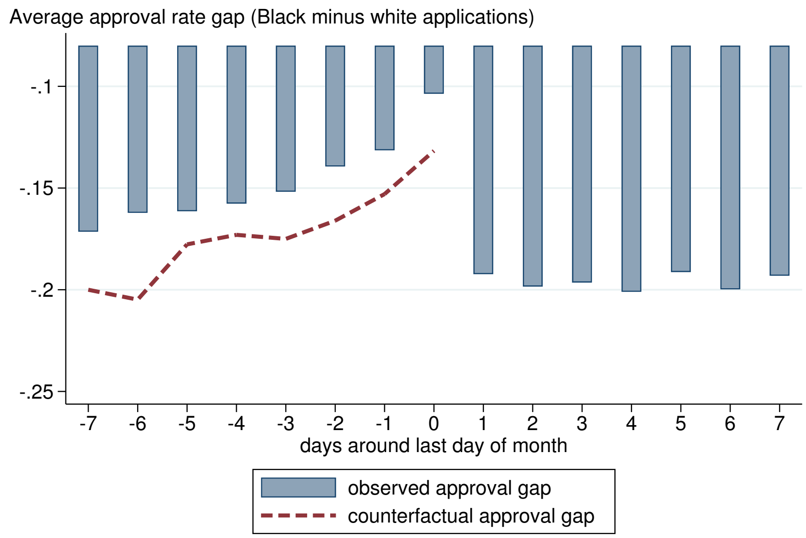
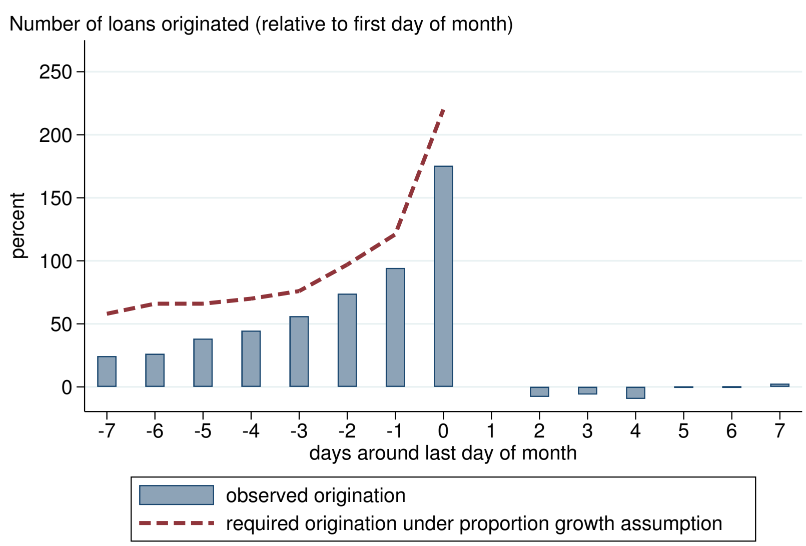


Figure A.4: This figure shows the ratio of the number of originations to Black applicants, over originations to white applicants, for each day in the last and first week of the month, at the level of the entire United States. The evidence is based on the HMDA data from January 1994 to December 2018.



(a) Approval Gap



(b) Origination Volume

Figure A.5: Panel (a) shows the difference between the fraction of approved loans, out of all approved and denied loans in the U.S., for Blacks minus the one for whites, on each of the last eight days of the month and the first seven days of the following month. In the last week of the month, we overlay to the graph a dotted line, showing the change in the approval gap generated by an increase in origination volume matching the one we observe in the data, but with a constant share of loans originated to Black applicants on each day in the first and last week of the month. Panel (b) shows average percentage abnormal daily loan origination volume in the U.S., for the last eight days of the month, and the first seven days of the following month. Along similar lines as in Panel (a), we include for the last week of the month a dotted line, showing the increase in origination volume that would be needed to match the corresponding approval gap on each day in the data, but under the restriction of a constant share of loans originated to Black applicants. The evidence is based on the HMDA data from January 1994 to December 2018.

	(1) approval	(2) approval	(3) approval	(4) approval	(5) approval	(6) approval
$black$		-0.049***	-0.067***	-0.068***	-0.064***	-0.067***
		(0.0036)	(0.0043)	(0.0044)	(0.0043)	(0.0042)
$lastweek$	0.043***	0.043***	0.043***	0.043***		0.071***
	(0.0030)	(0.0027)	(0.0030)	(0.0031)		(0.012)
$firstweek$	-0.020***	-0.019***	-0.020***	-0.019***		-0.013**
	(0.0017)	(0.0017)	(0.0017)	(0.0017)		(0.0050)
$black \times lastweek$	0.027***	0.027***	0.027***	0.028***		0.024***
	(0.0021)	(0.0018)	(0.0021)	(0.0021)		(0.0020)
$black \times firstweek$	-0.0074***	-0.0086***	-0.0073***	-0.0072***		-0.0067***
	(0.0015)	(0.0013)	(0.0015)	(0.0015)		(0.0015)
$female$			-0.017***			
			(0.0025)			
$lastday$					0.081***	
					(0.0068)	
$firstday$					-0.019***	
					(0.0032)	
$black \times lastday$					0.046***	
					(0.0039)	
$black \times firstday$					-0.0069**	
					(0.0033)	
$black$-Year	YES	NO	NO	NO	NO	NO
Loan-Level Controls	YES	YES	YES	YES	YES	NO
Loan-Level Controls-Inter	NO	NO	NO	NO	NO	YES
Holiday FE	YES	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES	YES
Lender FE	YES	YES	YES	YES	YES	YES
Month-Year-County	YES	NO	YES	YES	YES	YES
Month-Year-Census Tract	NO	YES	NO	NO	NO	NO
Month-Year-Lender	YES	YES	YES	YES	YES	YES
$last - first$	0.063	0.062	0.063	0.062	0.10	0.084
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black)$	0.097	0.098	0.097	0.097	0.15	0.12
$p - value\ (black)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (black - other)$	0.034	0.036	0.034	0.035	0.053	0.031
$p - value\ (black - other)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	17898971	12725901	17898971	17893870	17893870	17898971
r2	0.32	0.54	0.32	0.32	0.32	0.32

Table A.1: The table reports several robustness checks for the main results in Table 3, based on individual loan-level regression estimates of abnormal origination rates in the last and first week of the month (see equation 8). The dependent variable is a dummy that takes value 1 if a loan application is originated and 0 if it is denied. For all columns the regression specification includes all the controls used in column 6 of Table 3. In column (1), we interact the $black$ dummy with dummies for each year in our sample. In column (2), we replace the county dummies with census tract dummies and interact the with time controls (year-month). In column (3), we also control for applicant gender. In column (4), we change the definition of month in the time fixed effect, so that they begin and end on the 15th calendar day, rather than on the 1st calendar day. In column (5), we maintain the modified definition of months for the fixed effects, and we focus on changes in the black-applicants gap between the last and first day of the month, and thus $lastday$ and $firstday$ are dummies equal to one, respectively, if the decision on the application is taken in the first and last day of the month. Finally, in column (6), we interact loan-level controls with the first and last week of the month dummies. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$ ($lastday$ and $firstday$), and for the difference of the interaction coefficients for Black applicants, along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) approval Female	(2) approval Male	(3) approval All	(4) approval All	(5) approval All	(6) approval All
$lastweek$	0.070***	0.053***	0.053***	0.044***	0.042***	0.041***
	(0.0043)	(0.0036)	(0.0036)	(0.0034)	(0.0030)	(0.0030)
$firstweek$	-0.025***	-0.021***	-0.021***	-0.020***	-0.019***	-0.019***
	(0.0026)	(0.0025)	(0.0025)	(0.0020)	(0.0017)	(0.0016)
$female$			-0.082***	-0.050***	-0.024***	-0.021***
			(0.011)	(0.0085)	(0.0030)	(0.0026)
$female \times lastweek$			0.017***	0.016***	0.011***	0.011***
			(0.0023)	(0.0021)	(0.0016)	(0.0016)
$female \times firstweek$			-0.0043***	-0.0034***	-0.0031***	-0.0032***
			(0.00091)	(0.00084)	(0.00087)	(0.00076)
Loan-Level Controls	NO	NO	NO	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES	YES	YES
Month-Year FE	YES	YES	YES	YES	YES	NO
County FE	YES	YES	YES	YES	YES	NO
Lender FE	NO	NO	NO	NO	YES	NO
Month-Year-County	NO	NO	NO	NO	NO	YES
Month-Year-Lender	NO	NO	NO	NO	NO	YES
$last - first$	0.095	0.074	0.074	0.064	0.061	0.060
$p - value$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$last - first\ (female)$			0.096	0.084	0.075	0.074
$p - value\ (female)$			0.0000	0.0000	0.0000	0.0000
$last - first\ (black - other)$			0.022	0.020	0.015	0.014
$p - value\ (black - other)$			0.0000	0.0000	0.0000	0.0000
N	7090705	12550419	19641147	18464529	18463277	17898971
r2	0.043	0.036	0.044	0.092	0.23	0.32

Table A.2: The table repeats the analysis in Table 3, but the $black$ indicator variable is replaced with a $female$ indicator variable equal to one for female applicants. All regressions are based on individual loan-level regression estimates of abnormal origination rates in the last and first week of the month (see equation 8). The dependent variable is a dummy that takes value 1 if a loan application is originated and 0 if it is denied. In columns (1) and (2), the sample is restricted to, respectively, Blacks and whites or other race applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$ ($lastday$ and $firstday$), and for the difference of the interaction coefficients for female applicants, along with the *p-value* of tests of the null that each difference is equal to 0. Standard errors are clustered by lender and year. Estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) black	(2) black	(3) black	(4) approval	(5) approval
$lastweek$	-0.0014**	-0.00072	-0.00073	0.31***	0.32***
	(0.00058)	(0.00053)	(0.00057)	(0.0014)	(0.0015)
$firstweek$	-0.00063	-0.00020	-0.000075	-0.12***	-0.12***
	(0.00071)	(0.00050)	(0.00055)	(0.0015)	(0.0015)
$approval$	-0.080***	-0.049***	-0.050***		
	(0.0033)	(0.0017)	(0.0018)		
$approval \times lastweek$	0.0053***	0.0019***	0.0021***		
	(0.00066)	(0.00060)	(0.00069)		
$approval \times firstweek$	0.00011	0.00057	0.00041		
	(0.00063)	(0.00051)	(0.00060)		
$black$				-0.63***	-0.55***
				(0.0026)	(0.0028)
$black \times lastweek$				0.070***	0.064***
				(0.0046)	(0.0047)
$black \times firstweek$				-0.0086*	-0.0074
				(0.0049)	(0.0051)
Loan Level Controls	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	NO	YES
Day-of-Week FE	YES	YES	YES	NO	YES
Month-Year FE	YES	YES	NO	NO	YES
Lender FE	NO	YES	NO	NO	NO
County FE	NO	YES	NO	NO	NO
State FE	NO	NO	NO	NO	YES
Month-Year-County	NO	YES	YES	NO	NO
Month-Year-Lender	YES	NO	YES	NO	NO
$last - first\ (approval)$	0.0052	0.0013	0.0017		
$p - value\ (approval)$	0.0000	0.042	0.035		
N	14510888	14030895	13465036	19064873	18464655
r2	0.037	0.19	0.26	-	-

Table A.3: The table reports several alternative tests for discriminatory behavior based on high-frequency (within-month) variation. All estimates are based on loan level data, including all approved and denied loan applications for Black and white applicants. In columns (1), (2), and (3) the dependent variable is a dummy equal to one for Black applicants. The variable $approval$ is a dummy equal to one for originated loans, $lastweek$ and $firstweek$ are dummies equal to one for loans approved or denied in the last and first week of the month. The table also reports, for the first three columns, estimates of the difference in the coefficients for the interactions between the approval dummy and the last and first week dummies, along with their *p-values* for the null that the difference between the coefficients is equal to 0. In the first three columns, standard errors are clustered by lender and year. In columns (4) and (5), we estimate logit models, which predict approval probabilities at the loan level. Across all columns, estimates are based on a 5% random sample of the HMDA data from 1994 to 2018.

	(1) 5-Year Default	(2) 5-Year Default FICO < 660	(3) 5-Year Default LTV $> 80\%$	(4) 5-Year Default Low Docs
$lastweek$	-0.00011	-0.00016	-0.00031	0.00042
	(0.00027)	(0.00047)	(0.00033)	(0.00077)
$firstweek$	0.00044	0.0011	0.00097*	0.00049
	(0.00027)	(0.00064)	(0.00054)	(0.00041)
$black$	0.0033**	0.0011	0.0017	0.0075***
	(0.0012)	(0.0012)	(0.0015)	(0.0024)
$black \times lastweek$	-0.0013**	-0.0020	-0.0022***	-0.0039**
	(0.00059)	(0.0011)	(0.00062)	(0.0016)
$black \times firstweek$	0.0020***	0.0011	0.0025**	0.0033***
	(0.00060)	(0.00096)	(0.00095)	(0.0011)
Loan-Level Controls	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES
$last - first$	-0.0006	-0.0012	-0.0013	-0.0001
$p - value$	0.2800	0.2300	0.1400	0.9400
$last - first\ (black)$	-0.0038	-0.0043	-0.0060	-0.0073
$p - value\ (black)$	0.0015	0.0036	0.0023	0.0003
$last - first\ (black - other)$	-0.0033	-0.0031	-0.0047	-0.0072
$p - value\ (black - other)$	0.0003	0.0380	0.00080	0.0020
N	20732913	3729582	6606008	5617655
r2	0.12	0.16	0.15	0.22

Table A.4: The table reports regression estimates of the difference in performance between mortgages originated in the last and first week of the month. The dependent variable is a dummy equal to one for mortgages that defaulted within 5 years after origination. In column (2), the sample is restricted to subprime loans (FICO < 660). In column (3) the sample is restricted to high loan-to-value loans (LTV$> 80\%$), and in column (4) to loans with low documentation. $lastweek$ and $firstweek$ are dummies equal to one in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with their *p-value*s. Standard errors are clustered by lender and year. Estimates are based on the merged sample of HMDA and Black Knight McDash data from 1994 to 2018.

	(1) 5-Year Termination	(2) 5-Year Termination FICO < 660	(3) 5-Year Termination LTV $> 80\%$	(4) 5-Year Termination Low Docs
$lastweek$	-0.0026	-0.00060	-0.0041*	0.0023
	(0.0022)	(0.0012)	(0.0020)	(0.0038)
$firstweek$	0.0010	0.00032	0.0018	-0.00028
	(0.0014)	(0.0020)	(0.0016)	(0.0017)
$black$	-0.054***	-0.038***	-0.061***	-0.044***
	(0.0083)	(0.0063)	(0.0091)	(0.0086)
$black \times lastweek$	-0.0024*	-0.0043***	-0.0014	-0.0088**
	(0.0014)	(0.0012)	(0.0015)	(0.0039)
$black \times firstweek$	0.0018	0.0013	0.0033	0.0036**
	(0.0018)	(0.0019)	(0.0021)	(0.0014)
Loan-Level Controls	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES
Day-of-Week FE	YES	YES	YES	YES
Month-Year-County	YES	YES	YES	YES
Month-Year-Lender	YES	YES	YES	YES
$last - first$	-0.0036	-0.00092	-0.0059	0.0026
$p - value$	0.2900	0.7700	0.0780	0.6200
$last - first\ (black)$	-0.0078	-0.0066	-0.0110	-0.0099
$p - value\ (black)$	0.1100	0.0440	0.0067	0.1200
$last - first\ (black - other)$	-0.0042	-0.0057	-0.0047	-0.0120
$p - value\ (black - other)$	0.1200	0.0410	0.1000	0.0047
N	20732913	3729582	6606008	5617655
r2	0.22	0.28	0.25	0.32

Table A.5: The table reports regression estimates of the difference in performance between mortgages originated in the last and first week of the month. The dependent variable is a dummy equal to one for mortgages that were terminated (due to default or refinancing) within 5 years after origination. In column (2), the sample is restricted to subprime loans (FICO $<$ 660). In column (3) the sample is restricted to high loan-to-value loans (LTV$>$ 80%), and in column (4) to loans with low documentation. $lastweek$ and $firstweek$ are dummies equal to one in the first and last week of the month. $black$ is a dummy equal to one for Black applicants. The table also reports estimates of the difference between the coefficients for the dummies $lastweek$ and $firstweek$, and for the difference of the interaction coefficients for Black applicants, along with their *p-value*s. Standard errors are clustered by lender and year. Estimates are based on the merged sample of HMDA and Black Knight McDash data from 1994 to 2018.