NBER WORKING PAPER SERIES

FINTECH LENDING WITH LOWTECH PRICING

Mark J. Johnson
Itzhak Ben-David
Jason Lee
Vincent Yao

Working Paper 31154
http://www.nber.org/papers/w31154

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023

We thank Avramidis Panagiotis, René Stulz, Francesco D'Acunto (discussant), and the seminar audiences at the CFPB Research Conference, Brigham Young University, the OCC FinTech Symposium, The Ohio State University, and the Federal Reserve Bank of Philadelphia for their helpful comments. The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or commissioner. This article expresses the authors' views and does not necessarily reflect those of the Commission, the commissioners, or staff members. Ben-David is a member of the Academic Research Council at the Consumer Financial Protection Bureau (CFPB). This article does not necessarily reflect the views of the CFPB or its staff. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

FinTech Lending with LowTech Pricing
Mark J. Johnson, Itzhak Ben-David, Jason Lee, and Vincent Yao
NBER Working Paper No. 31154
April 2023
JEL No. G21,G23,G50

ABSTRACT

FinTech lending—known for using big data and advanced technologies—promised to break away from the traditional credit scoring and pricing models. Using a comprehensive dataset of FinTech personal loans, our study shows that loan rates continue to rely heavily on conventional credit scores, including 45% higher rates for nonprime borrowers. Other known default predictors are often neglected. Within each segment (prime/nonprime) loan rates are not very responsive to default risk, resulting in realized loan-level returns decreasing with risk. The pricing distortions result in substantial transfers from nonprime to prime borrowers and from low- to high-risk borrowers within segment.

Mark J. Johnson
Brigham Young University
Department of Finance
Provo, UT 84602
markjjohnson@byu.edu

Itzhak Ben-David
The Ohio State University
Fisher College of Business
606A Fisher Hall
Columbus, OH 43210-1144
and NBER
ben-david.1@osu.edu

Jason Lee
U.S. Securities and Exchange Commission
jason.jaehyun.lee@gmail.com

Vincent Yao
Georgia State University
J. Mack Robinson College of Business
35 Broad Street NW
Atlanta, GA 30303
wyao2@gsu.edu

1 Introduction

Over the last decade, FinTech lending has become a popular source of funds for individuals and has drawn in immense investment from capital providers.[1] Compared to traditional lenders who use credit scores (e.g., FICO score) as the primary underwriting tool to screen applicants and price loans, FinTech lenders promise to bring improvement in underwriting by accessing new sources of data and applying advanced statistical techniques.[2] As summarized by U.S. House Representative Trey Hollingsworth:[3]

> "We are finding people who may, by traditional standards, have challenging credit scores or challenging situations, but through new algorithms, new technology, and new capabilities, are saying they might be great credit risks for these type of products."

Indeed, increasing the efficiency of credit allocation and tightening the link between pricing and risk would improve the welfare of borrowers, especially underserved populations with weak credit scores, and FinTech lenders are positioned to make this happen. These lenders use big data and statistical analysis, and can pose a significant competitive advantage

[1] For example, in the $138 billion unsecured personal loan market, the share of online loans increased from 22% to 49% over 2015–2019. Details can be found in a report by Experian: https://www.stlouisfed.org/publications/regional-economist/second-quarter-2019/unsecured-personal-loans-fintech.

[2] See, for example, Chris Lau, Why Hyper-Growth in AI Will Lift Upstart, February 7, 2022, available at https://www.nasdaq.com/articles/why-hyper-growth-in-ai-will-lift-upstart.

[3] See more discussions about FinTech lending in hearing testimony before the Subcommittee on Financial Institutions and Consumer Credit of U.S. House of Representatives Committee on Financial Services (January 30, 2018). For example, Nathaniel Hoopes, executive director of the Marketplace Lending Association, states "There is great evidence that partnerships between originating banks and marketplace lenders are delivering products to underserved communities, places where bank branches have closed and delivered products that are more affordable than the products that were available from traditional institutions and doing so by using advanced techniques that go beyond just looking at a traditional FICO score."

relative to traditional lenders.[4]

In this paper, we show that loan pricing by FinTech lenders is far from utopic risk-based pricing. FinTech lenders heavily rely on traditional credit scoring and do not incorporate other readily-available variable that are known to predict default into their pricing. In particular, loans made in the nonprime segment of the market are 45% more expensive than those at the prime segment of the market (for borrowers with similar risk). We discuss some of the institutional features concerning the market, competition, and regulation that may contribute to these pricing patterns. We then estimate the pricing discrepancies relative to a risk-based pricing counterfactual. Our results show that borrowers in the nonprime segment of the market subsidize borrowers in the prime segment. Within segments, borrowers with high credit quality subsidize those with low credit quality.

We use a unique dataset compiled by a data aggregator that contains most of the unsecured FinTech personal loans made in the U.S. throughout 2014–2020. The original dataset covers about 70% of FinTech unsecured personal loans in the U.S. over that period, originated via online platforms such as LendingClub, Upstart, and Avant. The data is used in real time by institutional investors and the lenders themselves. While these platforms always perform the borrower screening and loan pricing, they have access to different sources of capital, i.e., either their own or investors' capital (e.g., mutual funds and banks) via whole loan sales or securitization. Our data includes information about borrowers (e.g., credit score) and loans (e.g., contract terms and performance), but not applications, investor identity, and the sources of capital. The nature of the data allows us to study pricing decisions in FinTech lending and draw broad conclusions about the emerging industry.

[4] On the intrinsic margin, FinTech borrowers benefited from increased operational efficiency (Buchak, Matvos, Piskorski, and Seru, 2018; Fuster, Plosser, Schnabl, and Vickery, 2019; Berg, Burg, Gombović, and Puri, 2020a). On the extrinsic margin, individuals with limited prior access to the traditional credit markets could be viable borrowers of FinTech credit (Jagtiani and Lemieux, 2019; Berg et al., 2020a; Dolson and Jagtiani, 2021; Di Maggio, Ratnadiwakara, and Carmichael, 2021; Fuster, Goldsmith-Pinkham, Ramadorai, and Walther, 2022). Furthermore, FinTech lenders extend credit access to previously unserved or underserved populations (Hau, Huang, Shan, and Sheng, 2019; Agarwal, Kigabo, Minoiu, Presbitero, and Silva, 2021). They also substitute bank lending to riskier borrowers during the financial crisis of 2007–2009 (Tang, 2019; Balyuk, Berger, and Hackney, 2020; Chava, Ganduri, Paradkar, and Zhang, 2021; Di Maggio and Yao, 2021).

Our main empirical finding is illustrated in Figure 1: The pricing of FinTech loans has not broken away from pricing based on traditional FICO scores. In both panels, the x-axis presents borrowers' FICO score at origination, and the y-axis reflects a simple measure of ex-ante delinquency risk (discussed later in detail). This measure is based on observable characteristics at origination. Panel (a) presents a heat map of the average loan rate. The panel clearly shows that the traditional FICO score is the main determinant of loan pricing. At the same time, our ex-ante default measure based on other variables adds very little explanatory power to interest rates. In addition, the charts include indicators (black lines) marking the interquartile range of the number of loans in each FICO score. Panel (b) presents the ex-post outcomes of loans, measured by the average delinquency rate. It shows that the FICO score and other known default predictors predict loan performance out-of-sample. Together, these charts paint a picture of a market in which traditional FICO scores are the primary determinant of prices and are grossly insensitive to other readily-available information about future performance.

FinTech platforms' over-reliance on FICO scores in their pricing decisions has particular economic consequences regarding the existing segmentation of the personal credit market. Specifically, prime borrowers (FICO $\geq$ 660) generally have access to traditional unsecured and secured credit. Conversely, nonprime borrowers (FICO $<$ 660) do not have access to such credit, and therefore, FinTech lending presented a much-needed opportunity for them. We estimate that nonprime borrowers pay significantly higher interest rates by 7 percentage points (pp) than prime borrowers with the same default risk. These rates constitute a 45% premium over the average rate charged to similar-risk prime loans. This estimate is robust across empirical designs, e.g., regression discontinuity design (RDD) or an out-of-sample default propensity score model. Importantly, the rate gap between the market segments does not change over time, suggesting little impact from lender competition or learning. We explore other explanations for the rate gap, such as expected prepayment risk or unobserved risk that may manifest during a crisis. However, none appears to justify the rate jump.

We discuss some of the institutional factors to which FinTech lenders are exposed that make it difficult to break away from traditional pricing models. For example, the nonprime lending market segment is largely devoid of banks due to regulatory constraints that make originating and investing in nonprime loans prohibitive. As a result, nonprime lending can only be funded by nonbank capital through equity offerings, loans, or securitization, and thus, FinTech lenders face much less competition from banks. In contrast, in the prime segment of the market, FinTech lenders are exposed to greater competition from banks. Due to banks' access to low-cost deposit funding, equilibrium loan rates are substantially lower. Also, when FinTech lenders raise outside capital through securitization they abide by regulatory disclosure rules, which put an emphasis on traditional credit scoring.

Given the estimated gaps in the pricing of FinTech loans and the insensitivity of rates to risk, we set out to estimate the welfare consequences of these pricing distortions. To achieve this goal, we need *a counterfactual rate* that would have been charged by an *ideal* lender that fully implements risk-based pricing had the lender originated the loans. A necessary ingredient for carrying out the task is the expected return that lenders anticipate receiving on their investment. Empirically, loans with the lowest default likelihood appear to generate constant returns to their lenders of about 5% persistently. We assume that riskier loans should be priced such that they generate at least this rate. Therefore, by anchoring the expected return at that level and then using our out-of-sample delinquency risk measure, we can estimate a counterfactual rate for the remaining loans in the prime segment and for all nonprime loans segment. The difference between the actual and counterfactual rates is the implied premium or discount attributable to mispricing, either due to market segmentation or miscalculated risk.

Our analysis of our estimated rate differentials reveals economically significant cross-subsidization or mispricing across- and within-market segments. Specifically, on average, nonprime borrowers subsidize prime borrowers, and low-risk borrowers subsidize riskier borrowers within each segment. Over 83% of the nonprime borrowers overpay relative to the counterfactual rates. On average, all nonprime loans in our sample overpay by 4 pp, and those with similar risk to prime loans overpay by about 8 pp (30% of their average rate). In contrast, most prime borrowers are priced adequately (by design) or underpay relative to their counterfactual rates. The average overpayment by nonprime loans with similar risk to prime loans is about \$500 in the first year and \$1,000 over their loan terms, while prime loans with similar risk receive \$100 in the first year and \$200 through maturity per loan.[5]

Overall, our analysis indicates that FinTech loan pricing is surprisingly simplistic or "LowTech": It relies heavily on borrowers' credit scores while neglecting alternative data sources or other known risk factors. Our findings indicate that some of the longstanding pricing regularities in traditional lending markets (e.g., Keys, Mukherjee, Seru, and Vig, 2010; Agarwal, Chomsisengphet, Mahoney, and Stroebel, 2018; Argyle, Nadauld, and Palmer, 2020) persist in the digital age and have not improved throughout our sample period. We discuss the frictions that could generate these patterns and assess their implications for borrowers. As a result, nonprime individuals, who previously had little access to unsecured credit, have gained access, but most appear to overpay relative to their level of credit risk.

Our findings offer a silver lining for the FinTech industry. While loan pricing appears rather simplistic by the time our sample ends, there is significant scope for effortless improvement in the efficiency of loan pricing. Such an improvement would increase the access to fairly-priced credit to households, especially those underserved, and provide a fair risk-adjusted return to investors. This direction will likely enhance capital allocation across borrowers and increase consumer welfare.

The remainder of the article is organized as follows. Section 2 describes the data and the sample construction and provides summary statistics. Section 3 presents results on the pricing discontinuity around the 660 FICO cutoff using RDD and controlling for credit and market risk. We then propose several institutional features of the consumer credit

[5] Hurst, Keys, Seru, and Vavra (2016) document cross-subsidization of mortgage borrowers across regions. The distortion or mispricing results in transfers similar in magnitude to our estimates here (−\$580 to +\$780), albeit the size of mortgage are an order of magnitude larger than FinTech loans, and their duration is materially longer.

market that can explain the pricing discontinuity. Section 4 assesses the relationship between predicted risk and interest rates within different market segments. We conclude that FinTech platforms' pricing strategy is rather simplistic and inefficient. Section 5 estimates how loan rates should have been priced under counterfactual rates for individual loans and associated cross-transfers. Section 6 concludes.

2 Data and Summary Statistics

2.1 Data

The initial sample includes 7.4 million loans originated between January 2014 and July 2020 by seven of the largest online lenders and available through a leading FinTech loan aggregator. This dataset encompasses the vast majority of all FinTech personal loans made online—roughly 70% in terms of volume.[6] All loans were originated via online platforms such as LendingClub, Upstart, and Avant, which provide screening, pricing, and tracking services. Loans are funded either by lenders' own capital or investors' capital (typically institutional investors).

For each loan, we observe several key borrower characteristics available to lenders during underwriting and pricing, including FICO score, annual income, credit utilization ratio, loan purpose (e.g., debt consolidation), and homeownership status. We also observe the loan terms, including amount, maturity, scheduled payments, and annual interest rate. Last, the dataset includes performance information, including month-by-month principal and interest payments. In addition, the dataset includes delinquency status by month, categorized as current, delinquent, in default, or charged off. In the event of a charge-off, actual recoveries

[6] This estimate is based on quarterly loan volumes aggregated in the U.S. Fintech Market Report by S&P Global Market Intelligence in February 2021.

and losses are reported.[7]

2.2 Sample

Several filters are applied to the data to ensure that the analysis is not affected by idiosyncrasies in data reporting. We first address our inability to verify whether reported interest rates include or exclude upfront fees. One could back out the APR if given the gross loan amount, origination fees, and scheduled loan payments. However, in the data, it is unclear whether loan amounts are gross or net of fees; for many observations, origination fees are missing. Therefore, we restrict the sample to loans with multiples of \$500 or \$1,000, as these round loan amounts are likely to exclude fees. This filter reduces the sample by 1.7 million loans. An additional 100,000 loans are dropped when we require the reported interest rate to equal the interest rate implied by the loan terms.[8]

We further filter the sample based on loan maturity and apparent censoring of FICO scores. We limit the sample to loans with maturities of 36 months—the most popular term—which reduces our sample by 1.6 million. Inspecting the data, it appears that some borrowers have precise FICO scores while others are censored. This censoring means that in a histogram of FICO scores, we observe spikes at FICO scores ending in 2 or 7 and every other 9, suggesting that some lenders report all borrowers within a range of FICO scores using a single value (e.g., 652 for 650–654). We assume that FICO scores ending in 2 or 7 correspond to a five-point bin average and leave them in the data set. We exclude those borrowers censored at every 20-point FICO ending in 9 because our identification requires us to compare loans with FICO bins in the 655–659 and 660–654 ranges. Doing so reduces

[7] "Recoveries" refers to the total cash payment that the lender receives when the loan is sold off to other investors or debt collection agencies. This variable is critical in computing the return to the investor, and in many cases, it is an unobserved variable that limits the ability to make precise calculations. For example, De Roure, Pelizzon, and Thakor (2021) estimate returns, but they have to assume different recovery rates and verify that results are insensitive to these rates. Di Maggio et al. (2021) calculate realized returns to the lender but assume full repayment after 12 months. Due to our long time series, we can calculate returns to matured loans.

[8] They could differ if the reported rate were an all-in APR, but the loan amount was reported gross of fees.

our sample by an additional 765,000 observations.

Last, we drop loans for which any variables needed in regressions are missing, including loan amount, payment, default, state, and origination time. Unless otherwise noted, we also drop any loans originated after July 2019 to give us a performance window of at least 12 months.

These restrictions result in a sample of 2.3 million loans. Our results are largely insensitive to all of these restrictions, but we choose to be conservative to alleviate any concerns arising from including these observations.[9]

We calculate the ex-post realized returns for each loan as the internal rate of return (IRR) using all the previously described cash flows.[10] By focusing on loans with a single term (36 months), we avoid comparability issues arising from comparing cash flows with different terms.[11]

2.3 Summary Statistics

Table 1 provides summary statistics for loan terms, borrower characteristics, and loan performance. These unsecured loans' average interest rate and loan size are 16.3% and $11,898, respectively. Borrowers have a median income of $62,500, which is very similar to the national median household income in 2018 of $63,179.[12] The average (median) FICO score is slightly above prime at 684 (677). The average revolving credit utilization of 53% is

[9] To illustrate the differences between the final sample and the entire dataset, see interest rates by FICO depicted in Figure A.1 in the Appendix. The filtered sample has a slightly higher jump around the non-prime/prime cutoff than the unfiltered data by about 2 pp.

[10] To provide a more concrete example of how returns are calculated, consider $10,000 loan with a 20% interest rate that requires equal payments $371 for 36 months. In this example, the IRR would equal the interest rate of 20%. However, if the payments stopped after ten months and the loan was charged off in month 13, we include the monthly recoveries as one of the cash flows. Hence, if recoveries in month 13 are $3,000, then the IRR would be −57%. Last, we use the IRR to compute the effective annual interest rate of the loan. We do so to ensure that annualized returns are not below −100%.

[11] E.g., loans with the same origination date and different maturities may be exposed to different economic shocks.

[12] See https://www.census.gov/library/stories/2019/09/us-median-household-income-not-significantly-diff html.

above the average for households in the U.S. of 42%–45%.[13] This relatively high utilization rate is expected, given that these loans are predominantly made to individuals who intend to consolidate their debt or pay off high credit card balances. Almost three in four borrowers (76%) take out a FinTech loan for this purpose, while other cited purposes are other (11%), home improvement (7%), and medical expenses (1%).

3 The Prime/Nonprime Rate Gap

Does FinTech lending break away from traditional lending that focuses heavily on FICO scores? In this section, we estimate a substantial pricing gap between otherwise similar borrowers categorized as prime or nonprime borrowers based on FICO scores. We then discuss institutional factors that could lead to an equilibrium where rates differ substantially for these two market segments.

3.1 RDD Regressions

Panel (a) of Figure 2 plots point estimates and confidence intervals from regressions of interest rates on FICO score bins while controlling for origination year-month fixed effects. The five-point FICO bins begin at FICO 600 and end at 720, with an increment of five points (e.g., 650–654). The estimated coefficients for each FICO bin represent the average level of these variables for all loans contained in the bin. Panel (a) shows a striking jump at FICO 660—the threshold for prime borrower status, with the average interest rate on the loans shifting from about 17% in the 660–664 FICO bin to 25% in the 655–659 FICO bin, increasing by 8 pp. In contrast, there is no discontinuity across FICO bins above 660 and much smaller jumps among bins below 660. Figure A.2 in the Appendix plots estimates from regressions of interest rates on other variables using similar specifications. While Panel (a) shows discontinuously smaller loan sizes for nonprime borrowers, Panels (b) and (c) show no

[13] See "revolvers" utilization in reports by the American Bankers Association. See https://www.aba.com/news-research/research-analysis/credit-card-market-monitor/.

discontinuity in either income or monthly payment in FICO scores. Thus, the discontinuity in pricing is not simultaneously a change in FICO and some other innate characteristic of the borrower.

Panel (b) of Figure 2 shows a moderate jump in the delinquency rate at FICO 660, while continuous patterns generally persist elsewhere. This discontinuity in performance is not entirely unanticipated, given the significant price jump. Adverse selection and moral hazard might induce both the selection of borrower types (on unobservable characteristics) and subsequent changes in repayment behavior. Both of these effects result from higher interest rates and not because there is anything uniquely risky about borrowers who happen to be at FICO 659 at origination instead of 660. What is surprising, however, is that despite having higher rates of delinquency, these nonprime borrowers consistently deliver returns to the lender that are substantially higher (6.3 pp or a 260% increase) than prime borrowers, as seen in Panel (c) of Figure 2.

Exploiting the discontinuity, we next use an RDD approach to quantify the magnitude of the jump in interest rates and ex-post performance at FICO 660. This approach allows us to test the sensitivity in this jump to include other fixed effects and covariates within a narrow window. For example, loan amounts also change at the cutoff and could explain some of the disparities if lower loan amounts are associated with higher loan rates.

To identify the differences in outcomes (e.g., rate), r, between the otherwise similar prime and nonprime borrowers, we estimate the following model:

$$\begin{aligned} r_{it} =& \beta_1 \cdot I(FICO_{it} < 660) + \beta_2 \cdot I(FICO_{it} < 660) \cdot (FICO_{it} - 660) \\ &+ \beta_3 \cdot (FICO_{it} - 660) + \theta_{it} + \gamma_i + \mu_t + \nu_{t,s} + \epsilon_{it}, \end{aligned} \tag{1}$$

where $\text{FICO}_{it} \in [660 - h, 660 + h]$. The variable $\text{I}(\text{FICO}_{it} < 660)$ is an indicator equal to one if the FICO score is less than 660 and zero otherwise. The primary coefficient of interest, β_1, is the estimated jump in the outcome for individuals below prime. The bandwidth h

identifies the window around the cutoff for which observations must fall to be included in the estimation. While not essential, especially in a very localized window around the threshold, the second and third terms in the equation allow the slope to differ above and below 660. We also include the loan amount (θ_{it}) and fixed effects for origination month (μ_t), state (γ_i), and state-by-year ($\nu_{t,s}$).

Table 2 reports the results of the RDD regressions. Columns (1)–(3) show that interest rates jump by 8.6–9.4 pp for nonprime borrowers, compared to similar prime borrowers within a narrow window. Magnitudes differ by less than one pp when the bandwidth around FICO 660 goes from 40 in Column (1) to 5 in Column (3), which excludes FICO controls, given the tight window around the cutoff.[14] Columns (4)–(6) indicate that delinquency rates jump by 4.4–4.9 pp for nonprime borrowers. This is a relatively large increase from an average delinquency rate of 8.2% for prime borrowers. The magnitude of the estimated pricing gap is economically large.

We note, however, two important caveats to the RDD analysis. First, RDD can only be used to compare loans within a narrow band near the FICO cut off, making it hard to generalize to a larger sample. Second, we cannot decompose the estimated price jump into what can be justified by expected risk and what cannot since lenders may price expected risk differently.

3.2 Interest Rate Gap Unrelated to Credit Risk

To address the RDD analysis's caveats, we estimate a model of delinquency risk in this subsection and use the predicted delinquency rate to assess the interest rate gap unrelated to ex-ante risk.

[14] Furthermore, because many loans have FICO scores that are censored at 2s and 7s, the inclusion of RDD controls may be largely influenced by this bunching.

3.2.1 Assessing Credit Risk

We estimate a logit model that regresses the delinquency indicator on borrower and loan characteristics in the following specification:

$$I(DLQ)_{it} \quad = \quad \beta \cdot X_{it} + \epsilon_{it}, \tag{2}$$

where I(DLQ)_{it} is an indicator of whether the loan, taken out by the borrower i, becomes delinquent in the first 12 months after origination date t, and X_{it} is a set of borrower and loan attributes at origination date, including 20-point FICO bins, $5,000-increment income bins, $1,000-increment loan amount bins, and indicators for the loan purpose as well as the age of the oldest credit, revolving utilization, and the number of credit inquiries in the last six months from the credit bureau data. We also include $\Delta FICO$, which is the difference between the FICO bin and the actual FICO score, to account for FICO variation within bins. To obtain an ex-ante measure of risk for each loan originated in month t, we run rolling regressions using loans originated between months $t - 36$ and $t - 12$. The 12-month gap between estimation and prediction ensures we only use the information available at origination.

Figure 3 plots selected coefficients for four regressions in the time series, while others are plotted in Figure A.3 in the Appendix. The coefficients are relatively stationary across the four year-months that are plotted.[15] The two panels show that risk decreases monotonically with FICO and income, as anticipated. To save space, loan amount bins, loan purpose, and other credit bureau coefficients are plotted in the Appendix.

We evaluate the performance of the model in two ways. The first is to look at the area under the curve (AUC), a commonly used measure for prediction accuracy. An AUC of 50% implies that the model does no better than random chance, while 100% has perfect predictive capabilities. It is generally accepted that 70% is considered good, especially

[15] The period 2017m1 refers to the regression that predicts loan delinquency for January 2017. This regression includes data from January 2014 to December 2015.

in information-scarce environments. The average AUC in our regressions is close to that benchmark at 66.5%, which squares with other unsecured debt models documented in the literature.[16] The second approach is to plot the average predicted delinquency $I(\widehat{DLQ})_{i,t}$ against realized delinquency separately for the prime and nonprime markets. If the model performs well, we expect to see little difference in performance between the two markets and for the slopes to line up close to the 45-degree line. We plot the average delinquency for 50 predicted delinquency bins separately for prime and nonprime loans in Figure 4, Panel (a). The standard errors between the two markets are mostly overlaid in the region where prime and nonprime loans have similar ex-ante risks. The figure confirms that realized delinquency aligns well with predicted delinquency (lies on the 45-degree line). There is also a consistent overlap between the two markets, which we will use in later tests. We will refer to this region as the "overlapping region" of risk.

Panel (b) of Figure 4 shows the relation between ex-ante risk and default for the subset of loans that have reached maturity by July 2020. The results confirm that the predicted delinquency and the realized default are closely related in this sample. Since default can happen at any time during the life of the loan, it is not surprising that default is higher than 12-month delinquency for any level of predicted delinquency. Surprisingly, prime borrowers tend to default at higher rates than nonprime borrowers for similar levels of ex-ante delinquency risk.

3.2.2 Estimating the Interest Rate Gap While Controlling For Credit Risk

Using a reliable measure of ex-ante risk, we compare the pricing of loans with similar risks. The exercise is relatively straightforward. We plot average interest rates over the predicted delinquency rate (in bins) separately for prime and nonprime borrowers in Figure 5, Panel (a). The difference in rates between the two is statistically and economically significant. This gap is plotted in Panel (b) for the overlapping region where we can comfortably compare

[16] For example, Berg et al. (2020a) find the AUC to be 68.3% using only credit bureau information in their data, and Berg, Puri, and Rocholl (2020b) calculate an AUC of 66.6% using data from a German bank.

loans with similar risk, as opposed to the far left or far right region, where we only see loans in one of the two segments. On average, rates for nonprime loans are about 10 pp higher than those for similar-risk prime borrowers, which is very similar to what we found in RDD regressions. The gap stays mostly the same over time and may even get a little bigger, which suggests that lenders do not learn much and that competition is not likely to close this gap.

By focusing on loans in the overlapping region, this ex-ante risk measure allows us to compare borrowers who may be far from the FICO 660 cutoff but with similar credit risk, which makes our findings more generalized than the RDD analysis. In other words, the overlapping region includes 848,000 borrowers, whereas the 10-point window around 660 includes only 213,000 borrowers. We make the measurement even more precise by regressing interest rates on the nonprime loan indicator while controlling for loan size and predicted delinquency risk, either as a continuous or a categorical variable.

$$r_{it} = \beta_1 \cdot I(FICO_{it} < 660) + \beta_2 \cdot I(\widehat{DLQ})_{it} + \theta_{it} + \mu_t + \epsilon_{it}, \tag{3}$$

where $I(\widehat{DLQ})_{it}$ is the ex-ante predicted risk obtained from rolling regressions specified in Equation (2); θ_{it} is the loan amount; μ_t is the origination time fixed effects. Table 3 presents the analysis. Our baseline specification is in Column (1), where we control for year-month fixed effects, loan amount, and $I(\widehat{DLQ})_{it}$ as a continuous variable. Column (2) controls $I(\widehat{DLQ})_{it}$ as a categorical variable, and Column (3) adopts the same specification as in Column (1) but based on only loans with FICO score 20 points above and below the 660 cutoff. Results show that the rate gap between the two segments ranges from 6.2 to 7.3 pp and is statistically significant at the 1% level.

3.3 Controlling for Other Risk

In addition to credit risk, lenders may charge higher interest rates on loans with higher prepayment risk, which can result in the loss of revenue for lenders due to reinvestment risk and thus justify higher rates on these loans. In Table 4, we present the analysis using as the dependent variable a prepayment indicator variable that equals 100 if the loan is prepaid before the maturity date and 0 otherwise. Columns (1)–(3) adopt similar specifications as in Table 3 where we control for predicted delinquency risk, loan amount, and year-month fixed effects. Results are similar across different specifications. Column (1) suggests that nonprime loans are 2.8 pp more likely to be prepaid before maturity than prime loans of comparable risk and size, representing 7% of the average prepayment rate (39%).

To account for the difference in prepayment risk between nonprime and prime loans, we obtain the component of the expected prepayment risk that is orthogonal to the expected credit risk. We include it as an additional control in Equation (3). Results are reported in Columns (4)-(6) of Table 4. They show that the estimated rate gap between nonprime and prime segments is still 7.3 pp using our baseline specification in Column (4), which is identical to the magnitude found in Table 3, suggesting that the estimated rate gap is unrelated to prepayment risk as well.

Magnitude of the Rate Gap Using the estimates from Column (4) in Table 4, the estimated rate gap of 7.3 pp is a 45% premium over the average rate for prime loans (15.5%) in the overlapping region. The gap accounts for 28% of the average rate for nonprime loans (26.5%) in the sample. With an average loan amount of \$5,344, these borrowers pay about \$400 ($= 5344 \times 7.33\%$) in interest payments during the first year of the loan more than they would in the absence of the rate premium. The overpayment would be \$800 during the entire term assuming that the expected duration is two years for a 3-year loan, equivalent to a week's gross pay for someone making \$42,000 a year.

3.4 Performance During COVID-19

The performance of loans in our sample is updated through July 2020, including several months during the COVID-19 period. If nonprime borrowers were riskier than prime borrowers based on unobservable characteristics, these factors would manifest when facing an unanticipated shock such as COVID, resulting in much worse performance for nonprime borrowers. To test the difference in performance during a crisis, we focus on the repayment behavior of FinTech loans before and during the COVID-19 crisis period, i.e., from October 2019 through July 2020. Specifically, our analysis includes loans that remain outstanding as of December 2019 and have predicted delinquency between 6% and 13% (the overlapping region).

Panel (a) of Table A.1 in the Appendix presents summary statistics of the panel data sample. It shows that loans in the nonprime sample have slightly higher ex-ante delinquency risk $I(\widehat{Delinquency})$ than those in the prime sample, primarily due to compositional differences. Their loan size is also much smaller. Following the onset of the COVID-19 pandemic, the delinquency rate of loans in both samples increased sharply to a very similar level, near 13%. We estimate the OLS regression in the following difference-in-difference (DID) specification based on the panel data:

$$I(DLQ)_{i,t} = \beta \cdot I(FICO_i < 660) \cdot I(COVID_t) + \gamma_i + \gamma_t + \theta_{i,t} + \epsilon_{i,t}, \tag{4}$$

where $\text{I(DLQ)}_{i,t}$ indicates whether the loan is delinquent at calendar time t; $\text{I(FICO}_i < 660)$ indicates whether the loan is nonprime; $\text{I(COVID}_t)$ is an indicator that equals 1 for months since March 2020, and 0 otherwise. As in a standard DID specification, we control for individual (γ_i) and time (γ_t) fixed effects. Hence, β captures the change in loan performance within-borrower before and after the onset of COVID-19. We also implement more stringent specifications by controlling for state $\times$ time, origination date (t_0) $\times$ time, and our predicted delinquency measure $\times$ time fixed effects ($\theta_{i,t}$). These extra steps ensure that we only compare the performance of loans within the same state and month of origination with the same ex-ante risk level.

Table 5 presents the results. Column (1) reports the standard DID results, and the other columns report the results with additional controls. Column (1) shows that nonprime loans perform slightly better during the crisis, with a 0.4 pp lower delinquency rate after March 2020 compared to prime loans during that same period and nonprime loans before the pandemic. In Columns (2) and (3), where we control for ex-ante delinquency and fixed effects for state and origination time, the difference in performance is much more pronounced, with nonprime borrowers having a 2 pp lower delinquency rate compared with prime borrowers with similar ex-ante risk during COVID-19 as well as nonprime loans before the onset of the pandemic. The magnitude is economically meaningful, as a raw decrease in delinquency of this size would have meant a 19% improvement from the pre-COVID-19 level.

These results show that nonprime loans had lower rates of delinquency during the COVID-19 period than similar prime loans before the COVID period. When we compare the performance of loans with similar ex-ante risk, the relative over-performance of nonprime loans is even more pronounced. The results suggest that nonprime loans do not have a significantly higher unobservable risk that would manifest when unanticipated shocks occur.

3.5 Why Does the Pricing Discontinuity Exist and Persist?

Why do lenders persistently charge different rates for loans around the 660 FICO cutoff with similar risks? In this subsection, we offer several explanations based on institutional features of the consumer credit market.

3.5.1 Limited Competition in the Nonprime Segment

During and after the global financial crisis, lending by large banks to small businesses and nonprime borrowers fell sharply and has not fully recovered (Chen, Hanson, and Stein, 2017; Buchak et al., 2018). To provide direct evidence of limited bank credit supply in the nonprime segment, we hand-collect the minimum FICO requirements disclosed by banks and

FinTech lenders online.[17] We start by identifying the 15 largest banks in the U.S. by asset size. We restrict our list to those offering unsecured personal loans. Among the nine that offer this type of loan, we first search on the bank website for minimum qualifications and, if unavailable, then search on NerdWallet, WalletHub, or Bankrate. Only one (PNC Bank) of these nine had no qualifications listed anywhere. We report the other eight in boldface font in Table 6. In addition to the largest banks, we include the minimum FICO score requirements reported for prominent bank lenders like American Express, Discover, Barclays, and USAA. These banks appear to be significant players in the personal loan space.

Next, we conduct a similar search among FinTech lenders. These online lenders, including the seven largest ones reported in boldface font, are shown in the last row of Table 6.[18] The contrast in lending requirements is stark between banks and non-banks. Online lenders often accept people with credit scores below 660, but none of the banks on this list do. The lack of bank lending to nonprime borrowers, corroborated by others in the literature,[19] leads to limited competition and higher market power for FinTech lenders in the nonprime segment. This bifurcated market may explain part of the puzzle in disjointed rates at the prime cutoff. FinTech platforms that lend below prime do not have to compete with big banks and can more easily take advantage of borrowers, whereas in the prime space, they have to compete with bank lending.

[17] Banks often post credit requirements on their websites. Personal finance websites like WalletHub, NerdWallet, and Bankrate summarize each lender's loan products and often list these requirements to help people decide where to apply for loans. For examples of how these websites display personal loan information, see https://wallethub.com/best-banks-for-personal-loans and https://www.nerdwallet.com/best/loans/personal-loans/best-personal-loans.

[18] Avant, Best Egg, LendingClub, LendingPoint, Prosper, SoFi, and Upstart are identified as the most prominent lenders according to the U.S. FinTech Market Report produced by S&P Global Market Intelligence in February 2021. For an older version of this report available without a specific request, see the 2017 version: https://pages.marketintelligence.spglobal.com/rs/565-BDO-100/images/DigitalLending_Public_Web.pdf.

[19] See, for example, the rejection rate for borrowers by FICO score among large banks in Di Maggio et al. (2021).

3.5.2 Regulatory Requirements

The regulatory environment in which lenders operate can impose constraints on the credit quality of borrowers and the pricing of loans. Consistent with this assertion, Buchak et al. (2018) present evidence for the central role of regulation in shaping the landscape of consumer lending. The authors focus on the mortgage market and show that two forces, regulatory differences and technological advantages, explain the simultaneous decline of banks and dramatic growth among FinTech lenders. In particular, the Dodd-Frank Act's changes to regulatory capital requirements and the restructuring of the financial regulatory framework have led to a significant drop in bank lending.

Banking regulations Banks must follow clear and strict rules when originating nonprime loans. For example, rules for nonprime lending (defined by borrowers below the 660 FICO cutoff) are spelled out in the FDIC Risk Management Manual for Examination Policies:[20] "Nonprime lending should only be conducted by institutions that have a clear understanding of the business and its inherent risks and have determined these risks to be acceptable and controllable given the institution's staff, financial condition, size, and level of capital support. In addition, nonprime lending should only be conducted within a comprehensive lending program that employs strong risk management practices to identify, measure, monitor, and control the elevated risks inherent in this activity." Basel III regulations are another example of how rules can raise the cost of loans to borrowers with impaired credit. In particular, the rule says banks must set aside more capital for riskier loans, with loan riskiness measured by external credit ratings such as FICO. Lower ratings are linked to lower credit scores, costing banks more to make and keep these loans on their balance sheets.

Even though FinTech companies are not constrained by banking regulations directly, they can still be affected by them in two ways. One is the wholesale funding channel, which shadow banks and FinTech lenders use to get money from banks or other investors with higher

[20] See more details about the 660 FICO cutoff at https://www.fdic.gov/regulations/safety/manual/manual_examinations_full.pdf.

regulatory costs. For example, several FinTech lending platforms, like Prosper, also make traditional direct-to-consumer loans through WebBank, an FDIC-insured, state-chartered industrial bank that has to follow the same rules. Second, FinTech lenders compete with banks to provide unsecured credit.[21] If competing banks charge higher rates on nonprime loans due to higher regulatory costs, FinTech lenders can also charge higher rates and remain competitive in this segment. In this case, bank rates on nonprime loans could become a pricing benchmark for FinTech lenders. Jiang (2019) uses mortgage data to demonstrate the two channels. There should be a hefty interest rate gap between nonprime and prime segments through both channels.

Originate-to-distribute (OTD) business model Many non-bank fintech lenders utilize the originate-to-distribute (OTD) business model, in which they sell loans they originate to third parties or securitize them. As a result, these lenders aim to maximize loan origination volume and related fees while complying with the Securities and Exchange Commission (SEC) disclosure guidelines.

The OTD business model can intensify the impact of FICO cutoffs on origination quantities and pricing in two notable ways. First, like mortgage originators before the Global Financial Crisis, FinTech lenders engaging in the OTD model primarily earn income from origination commissions. This makes them more sensitive to risk criteria set by institutional investors (often expressed in FICO terms) rather than predicted default probabilities. Second, the SEC's disclosure guidelines for securitized assets mandate issuers to provide loan-level information in a standardized format, enabling investors to evaluate the risk of the underlying assets (Securities and Commission, 2014). This information is presented in categories such as credit score ranges (e.g., 620–639, 640–659, 660–679) and loan sizes (e.g., \$1,500–\$1,999). As the success of securitization issuances (and originators' fees) depends on these presentations, FICO scores may receive too much attention relative to other predictors

[21] Our sample contains unsecured personal loans exclusively screened by FinTech platforms, whose investors include both banks and nonbanks. We do not have data on loans directly originated by banks.

of default.

State usury laws An additional regulatory constraint that all lenders must abide by is usury laws that limit the interest rate charged on loans. These laws are determined at the state level, with maximum rates varying widely across states. However, with only a few exceptions, most financial institutions can bypass these state-level requirements by incorporating their lending arm in states with high usury limits. The outcome is that in almost all states, the effective usury limit is 36%.[22]

Usury laws do not appear to play an essential role in loan terms in our sample. Only a small fraction of loans (12%) have rates that exceed 30%. Removing states where usury caps bind (Colorado, Iowa, New York, West Virginia, and Vermont) does not affect the paper's findings.

Federal fair lending laws Another factor is the federal fair lending laws—the Equal Credit Opportunity Act and the Fair Housing Act—that prohibit lenders from incorporating information correlating with race, religion, national origin, sex, marital status, age, source of income, etc. These laws effectively limit the use of many variables, such as income, that are predictive of default but correlated with some protected classes. Regarding location, many FinTech lenders operate a centralized platform without local branches like commercial banks; thus, they lack the local knowledge necessary to institute regional pricing.

3.5.3 Funding Costs and Cross-Subsidization

FinTech lenders differ considerably from banks on several critical dimensions, including how they obtain funding and the regulatory environments in which they operate. Banks mainly depend on low-cost deposits insured by the government and are not tied to the risk of how banks lend money. In contrast, FinTech firms have to raise funds from wholesale

[22] Colorado, Iowa, New York, West Virginia, and Vermont are the exceptions to the rule where the usury cap is set below 36%—Laws allowing financial institutions to export usury laws from out-of-state are not in effect in these states.

funding or captive markets at a higher cost that compensates for the risk of their assets. In other words, FinTech platforms tend to charge higher rates that vary with risk in both prime and nonprime market segments. On the other hand, for FinTech firms to compete with banks in the prime segment, they cannot charge competitively high rates and thus have to cross-subsidize the rates on prime loans using a premium extracted from the nonprime segment. This sector structure partly helps explain why the cost of credit in our sample of FinTech loans is so high, especially for nonprime loans, and why one segment subsidizes the other.

4 Within-Segment Pricing and Loan Returns

The previous section showed that FinTech pricing continues to be dictated by FICO scores, with an especially significant average pricing gap between prime and nonprime loans. The results are not justified by underlying risk, suggesting large deviations from risk-based pricing at the market segment level. In this section, we explore possible differences in the pricing-risk slope through which risk is priced to assess the efficiency of FinTech pricing within each segment. A steeper positive slope indicates higher sensitivity of interest rates to changes in risk. We also test the sensitivity of loan-level returns to risk to assess whether the risk is adequately priced.

4.1 Relationship Between Expected Risk and Rates

To begin our analysis, we estimate the relationship between predicted risk and interest rates for different market segments:

$$r_{it}^s = \beta_0^s + \beta_1^s \cdot I(\widehat{DLQ})_{it}^s + \mu_t^s + \epsilon_{it}^s, \tag{5}$$

where r_{it}^s, $I(\widehat{DLQ})_{it}^s$ and μ_t^s are defined the same as in Equation (2) except with a superscript s that is indexed to prime or nonprime loans. We also estimate Equation (5) using only those loans in the overlapping region (expected delinquency rates between 6% and 13%), which enable us to compare the pricing of nonprime and nonprime loans with similar risks. Results are reported in Panel A of Table 7. Columns (1) and (2) are based on all nonprime and prime loans, respectively; Columns (3) and (4) are based on nonprime and prime loans in the overlapping region, respectively; Column (5) estimates a DID specification by including the interaction of $I(\widehat{DLQ})^s_{it}$ and a nonprime dummy using both prime and nonprime loans in the overlapping region.[23] Finally, Column (6) estimates the relationship between price and risk for safe prime loans that we will later use in calibrating counterfactual rates.

The results highlight three important facts. First, consistent with results in Section 3, average rates in the nonprime market, captured by the coefficient on the nonprime indicator in Column (5), are significantly higher than those in the prime market by 14 pp. The estimated gap is higher than 7.3 pp in Table 4 because here we control for the interaction of nonprime dummy and predicted delinquency rate. Second, FinTech loan rates in the prime market are much more responsive to expected risk than those in the nonprime segment. Columns (1) and (3) show that a one-pp increase in the predicted delinquency rate raises the prime loan rate by 89 bp and 61 bp, respectively, compared to only 22 bp in Column (2) and 26 bp in Column (4) for nonprime loans, respectively.

Third, loans in the overlapping and non-overlapping markets within the prime and nonprime markets are also priced differently. Generally, the pricing slope is steeper for loans with lower risk. It decreases along the risk spectrum of the loans, i.e., from non-overlapping prime (the least risky) $\rightarrow$ overlapping prime $\rightarrow$ overlapping nonprime $\rightarrow$ non-overlapping nonprime loans (the most risky). For example, Columns (1) and (3) show that the risk gradient of price in the overlapping and all nonprime markets differs by 3.4 bp (=25.8-22.4)

[23]Panel (b) of Table A.1 in the Appendix reports statistics by four segments: non-overlapping prime, overlapping prime, overlapping nonprime, and non-overlapping nonprime. Across the four segments, interest rates vary from 11.2% for the non-overlapping prime loans, 15.5% for the overlapping prime loans, and 26.5% for the overlapping nonprime loans to 27.9% for the non-overlapping nonprime loans. The largest difference is 16.7% between the non-overlapping prime and non-overlapping nonprime loans, followed by 11% between the overlapping prime and overlapping nonprime loans. Loan amounts obtained by prime borrowers are generally higher than those obtained by nonprime borrowers. Loan sizes for non-overlapping nonprime are only 44% of those for non-overlapping prime.

per one one-pp increase in the predicted delinquency rate, while Columns (2) and (4) show that the differential risk gradient in the prime markets is 28.4 bp (= 89.3-60.9).

4.2 Realized Return

Interest rate sensitivity to risk is very different for prime and nonprime borrowers. However, without examining ex-post performance, it is difficult to interpret which (if any) of these market segments appropriately prices risk. If rates are sufficiently sensitive to risk, we would expect realized returns to remain constant with risk or perhaps increase. In this subsection, we test this relationship for different market segments for matured loans only using loan-level IRRs, which summarizes the total cash flows during the loan term, net of actual losses.[24] The analysis may help address the concern that higher rates charged on nonprime loans, although not justifiable with the expected risk, may be justifiable with their higher loss severity.

Panel B of Table 7 reports the results. Similar to Panel A, Columns (1) and (2) are based on all nonprime and prime loans, respectively; Columns (3) and (4) are based on nonprime and prime loans in the overlapping region, respectively; Column (5) estimates a DID specification based on all loans in the overlapping market. The negative and significant coefficients on predicted delinquency in all five columns indicate that IRR decreases with expected risk, suggesting that lenders generally do not price loans according to their expected risk. This is especially true for nonprime and prime loans in the overlapping market with coefficients of -106 and -95, respectively, both significant at the 1% level. The insignificant coefficient on the interaction term on predicted delinquency and the nonprime dummy suggests that the difference in the IRR-risk slope between nonprime and prime loans in the overlapping market is not statistically different from zero. In addition, Column (5) shows that nonprime

[24] There are several important caveats to using this IRR. First, it is only available for matured loans, which limits the sample to loans originated before July 2017. Second, the particular economic scenario during our sample period is relatively benign and may not represent a full business cycle. Third, we do not observe origination fees, servicing fees, or the lender's cost of capital and debt used to finance the loans. Thus, we cannot make inferences about lenders' profitability.

loans are associated with an IRR of 12.5 pp higher than prime loans with similar risks.

To visually illustrate this result, we plot the average IRR against expected risk in Figure 6. It shows that nonprime loans deliver substantially higher returns, about 11% on average, in the overlapping region. Furthermore, in both segments, returns always decrease in risk, except for the safest loans in the prime alone segment shown in Column (6), which exhibit constant returns in risk. The pattern of decreasing returns is suggestive of inefficient pricing. Even worse, the riskiest tails of both prime and nonprime segments show negative returns.

4.3 Summary

Put together, we find that the pricing of FinTech loans is rather simplistic and inefficient based on three findings. First, nonprime loans are charged a 45% premium that is not justifiable with their underlying risk, including default and prepayment risks, compared to prime loans with similar risk. Second, the pricing of nonprime loans is less responsive to expected risk than that of prime loans, with its pricing-risk slope smaller by 56%. Third, loan-level returns generally decrease with expected risk for prime and nonprime loans. This suggests lenders realize excess returns from less risky loans in both segments while losing money on more risky ones. The patterns in FinTech pricing imply a great deal of cross-subsidization on the online lending platforms, from nonprime to prime segments and, within each segment, from low-risk and high-risk loans. The distortion looks pretty persistent, with little change over time.

5 Implications of FinTech Pricing

What are the consequences of the seemingly simplistic and inefficient pricing of FinTech loans that we have documented? FinTech platforms may be helping some borrowers obtain low rates at the expense of other borrowers. This section aims to quantify the effective transfers that occur due to mispricing. We begin this analysis by determining counterfactual rates for each loan and then quantify the average overpayment or underpayment by market segment.

5.1 Estimating Counterfactual Rates

In an ideal world, loan rates would reflect the risk so that returns would be constant, if not increasing, with risk. Any attempt to identify this counterfactual rate is difficult from a practical standpoint for at least two main reasons. First, ascertaining the appropriate return on a portfolio of consumer loans is difficult without knowing the individual beta that translates expected risk into an expected rate. The data shows that most loans do not achieve appropriate returns since their returns decrease with risk. Second, while we observe the performance of loans conditional on their transacted loan rates, it is difficult to know how risk or repayment behavior will change under new counterfactual rates. This can be particularly problematic when adverse selection or moral hazard exists.

To address these challenges, we look for loans that exhibit non-decreasing returns in risk in our sample. From Figure 6, most loans in the non-overlapping prime segment—the least risky loans with an expected delinquency rate between 0% and 4%—pay an interest rate from around 7.5% to nearly 14%. Yet, despite these differences, their ex-post IRRs are persistently constant at around 5%. This pattern is further confirmed by results in Column (6) of Table 7, Panel B, where the coefficient on predicted delinquency using these loans (calibration sample) is not statistically significant, suggesting that they exhibit constant returns to risk. Given these facts, we assume that these prime loans show the most appropriate risk-based pricing based on our existing data, and riskier loans should be priced to generate at least this rate. We apply the estimated relationship between interest rates and risk in this segment, found in Panel A of Table 7, Column (6), to estimate a counterfactual rate for the remaining loans in the prime segment as well as for all loans in the nonprime segment.

We make two assumptions in this counterfactual exercise. First, all loans can be appropriately priced based on the risk-rate relationship among the safest prime loans. This assumption may be less realistic with extremely risky loans due to, for example, the interest rate caps placed by the state usury laws. But for most loans, we find no reason to believe that the relationship should change dramatically. Second, borrower risk is not affected by the loan's interest rate. This assumption would be violated if the counterfactual rate influenced the selection of borrowers or moral hazard behavior on their part. For most loans, the negative impact of adverse selection will be a non-issue because the counterfactual rates are lower than the actual rates, implying that the true counterfactual may be even lower. For loans with estimated counterfactual rates higher than their actual rates, our results from the RDD at FICO 660 show that higher rates probably affect delinquencies more than IRRs.

5.2 Magnitude of Rate Differentials

Based on the actual and counterfactual rates for individual loans, we calculate the rate differential as

$$\Delta\widehat{r}_{it} \quad = \quad r_{it,\ actual} - r_{it,\ cfactual}. \tag{6}$$

$\Delta\widehat{r}_{it}$ quantifies the degree to which the loan interest rate differs from the counterfactual. A positive number indicates overpayment, and a negative number indicates underpayment relative to the rate the borrower would have received.

Panel (a) of Figure 7 plots the average $\Delta\widehat{r}_{it}$ over the predicted risk for prime and nonprime loans separately. The rate differentials for all prime and nonprime loans are -0.5 pp (discount) and 4.3 pp (premium), respectively, primarily decreasing in predicted risk from 12.3 pp to -4.6 pp among nonprime loans.[25] Over 83% of the nonprime loans—those with an expected delinquency rate between 6% and 20%—are overpaying. In particular, those in the overlapping nonprime segment—with an expected delinquency rate between 6% and 13%—overpay by about 8 pp, which is slightly higher than the estimated rate gap 7.3 pp

[25] Some of this insensitivity may be due to state usury laws that restrict the maximum rates on consumer loans. Figure A.4 in the Appendix shows that note rates on nonprime loans are mostly capped at 36%.

in Section 3. In contrast, only 48% of prime loans have positive $\Delta\hat{r}_{it}$, and magnitudes of overpayment are very small. By construction, the average rate differential for most loans in the non-overlapping prime segment is close to zero. However, the riskier loans in the overlapping prime segment are underpaying by 1.1 pp.

We also calculate the implied transfer in dollars over the first year of the loan arising from mispricing by multiplying the loan amount by the calculated differential,

$$\widehat{Transfer}_{it} = \Delta\hat{r}_{it} \times LoanAmount_{it}. \tag{7}$$

Panel (b) of Figure 7 plots the density distribution of $\widehat{Transfer}_{it}$. On average (weighted by loan size), nonprime borrowers pay \$290 to the other market while prime borrowers receive \$150 from the other market, but both with wide dispersion.[26] If the loan is paid through maturity, the transfer amount would be roughly two times $\widehat{Transfer}_{it}$, i.e., \$580 and -\$300 for nonprime and prime loans, respectively. In particular, those in the overlapping nonprime segment overpay by \$500 in the first year and \$1,000 through maturity per loan to other borrowers. In contrast, those in the overlapping prime segment receive \$100 in the first year and \$200 through maturity per loan above and beyond the risk-based pricing.

How big are these transfers? We can use two benchmarks. First, Hurst et al. (2016) estimate mispricing in the mortgage market due to the national pricing policy of the government-sponsored enterprises (GSEs) to be −\$580 (subsidy) for high-risk states to +\$780 (taxes) for low-risk states. Second, following the onset of COVID-19, the U.S. Congress authorized the \$2.2 trillion CARES Act, which provided one-time direct cash payments of \$1,200 to low-income individuals making up to \$75,000 and another \$500 per dependent. Our estimated average interest payment by nonprime borrowers of \$500 in one year, or \$1,000 over the loan term, is comparable in magnitude to mortgage market mispricing despite loan maturities that are 5–10 times shorter. This average overpayment would represent roughly 82% of the

[26]Nonprime borrowers at the 75^{th} percentile pay \$600 in transfers, while prime borrowers at the same percentile pay only \$170.

subsidy relative to the CARES Act.

6 Conclusion

FinTech lenders pride themselves on using innovative data and superior statistical techniques to screen borrowers and price loans. One would expect that the rise of these new lenders would result in loan rates that reflect the riskiness of borrowers and loans.

Our results show that reality is still far from ideal. We discovered that borrowers' FICO scores are still the most critical determinant of pricing after analyzing 2.3 million unsecured personal loans made by FinTech platforms. Notably, other variables known to predict default are not adequately priced. Finally, we document that even in this market, there is a substantial interest rate gap of 7 pp between prime and nonprime loans with very similar risks after controlling for both credit and prepayment risks.

We highlight several institutional features of the market that are likely to contribute to the tight reliance of lenders on FICO scores. In particular, banks refrain from making nonprime loans due to banking regulations. As a result, competition is lower in the nonprime segment of the market. Furthermore, the capital available for lending in the prime segment is significantly cheaper since banks rely on low-cost deposits.

Our results demonstrate that the hysteresis in traditional lending markets persists in the digital age. While the current credit scoring system provides an imperfect proxy for default likelihood, even sophisticated lenders find it very hard to break away from the current framework and cannot implement true risk-based pricing using big data and advanced technologies. The continued overreliance on traditional credit scores for unsecured pricing has societal implications. Our results show that nonprime borrowers—especially those with low expected risk—cross-subsidize prime borrowers, especially those with high expected risk, leading to more expensive credit provisions for underserved populations.

References

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel, 2018, Do banks pass through credit expansions to consumers who want to borrow?, *Quarterly Journal of Economics* 133, 129–190.

Agarwal, Sumit, Thomas Kigabo, Camelia Minoiu, Andrea F Presbitero, and André F Silva, 2021, Serving the underserved: Microcredit as a pathway to commercial banks, *Review of Economics and Statistics* 1–45.

Argyle, Bronson S., Taylor D. Nadauld, and Christopher J. Palmer, 2020, Monthly payment targeting and the demand for maturity, *Review of Financial Studies* 33, 5416–5462.

Balyuk, Tetyana, Allen N. Berger, and John Hackney, 2020, What is fueling fintech lending? The role of banking market structure, Working paper, University of South Carolina.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri, 2020a, On the rise of fintechs: Credit scoring using digital footprints, *Review of Financial Studies* 33, 2845–2897.

Berg, Tobias, Manju Puri, and Jörg Rocholl, 2020b, Loan officer incentives, internal rating models, and default rates, *Review of Finance* 24, 529–578.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.

Chava, Sudheer, Rohan Ganduri, Nikhil Paradkar, and Yafei Zhang, 2021, Impact of marketplace lending on consumers' future borrowing capacities and borrowing outcomes, *Journal of Financial Economics* 142, 1186–1208.

Chen, Brian S, Samuel G Hanson, and Jeremy C Stein, 2017, The decline of big-bank lending to small business: Dynamic impacts on local credit and labor markets, Technical report, National Bureau of Economic Research.

De Roure, Calebe, Loriana Pelizzon, and Anjan V. Thakor, 2021, P2P lenders versus banks: Cream skimming or bottom fishing?, Working paper, SAFE.

Di Maggio, Marco, Dimuthu Ratnadiwakara, and Don Carmichael, 2021, Invisible primes: Fintech lending with alternative data, Working paper, Harvard Business School.

Di Maggio, Marco, and Vincent Yao, 2021, Fintech borrowers: Lax screening or cream-skimming?, *Review of Financial Studies* 34, 4565–4618.

Dolson, Erik, and Julapa Jagtiani, 2021, Which lenders are more likely to reach out to underserved consumers: Banks versus fintechs versus other nonbanks?, Working paper, FRB of Philadelphia.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, 2022, Predictably unequal? The effects of machine learning on credit markets, *Journal of Finance* 77, 5–47.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, 2019, The role of technology in mortgage lending, *Review of Financial Studies* 32, 1854–1899.

Hau, Harald, Yi Huang, Hongzhe Shan, and Zixia Sheng, 2019, How fintech enters China's credit market, in *AEA Papers and Proceedings*, volume 109, 60–64.

Hurst, Erik, Benjamin J. Keys, Amit Seru, and Joseph Vavra, 2016, Regional redistribution through the US mortgage market, *American Economic Review* 106, 2982–3028.

Jagtiani, Julapa, and Catharine Lemieux, 2019, The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform, *Financial Management* 48, 1009–1029.

Jiang, Erica Xuewei, 2019, Financing competitors: Shadow banks' funding and mortgage market competition, Working paper, USC Marshall.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? Evidence from subprime loans, *Quarterly Journal of Economics* 125, 307–362.

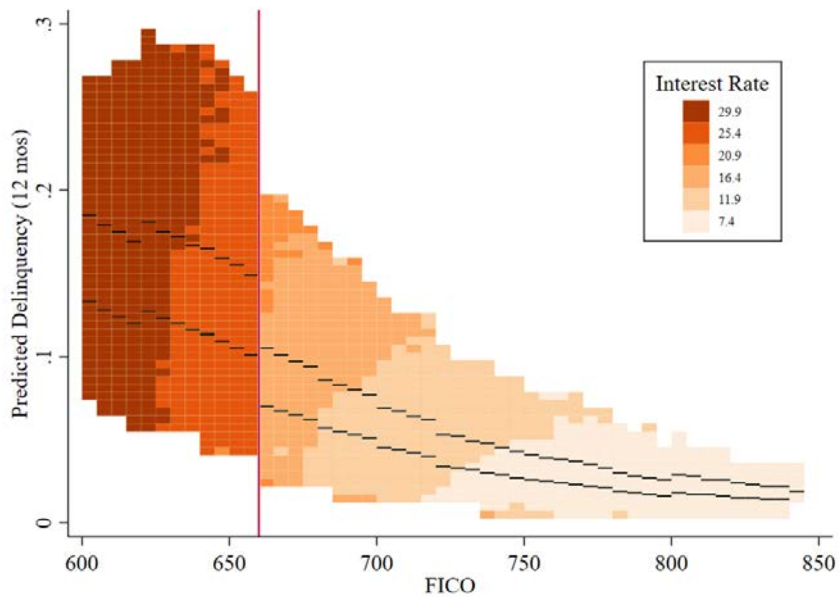
Securities, and Exchange Commission, 2014, Asset-backed securities disclosure and registration, Final rule, Securities and Exchange Commission.

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

Figure 1. Performance and Interest Rates, by FICO and Predicted Risk

This figure presents heat maps of the average interest rate (Panel (a)) and 12-month delinquency rate (Panel (b)) based on all loans in the sample. Loans are sorted by FICO score (x-axis; binned at five-point intervals) and by a measure of ex-ante delinquency (y-axis; binned at 0.5-point intervals). The delinquency risk measure is the out-of-sample prediction of a 12-month delinquency rate based on observable characteristics at origination. For each FICO bin, the black lines indicate the 25^{th} and 75^{th} percentile breakpoints in loan volume.

(a) Interest Rate (%)



(b) Delinquency (%)

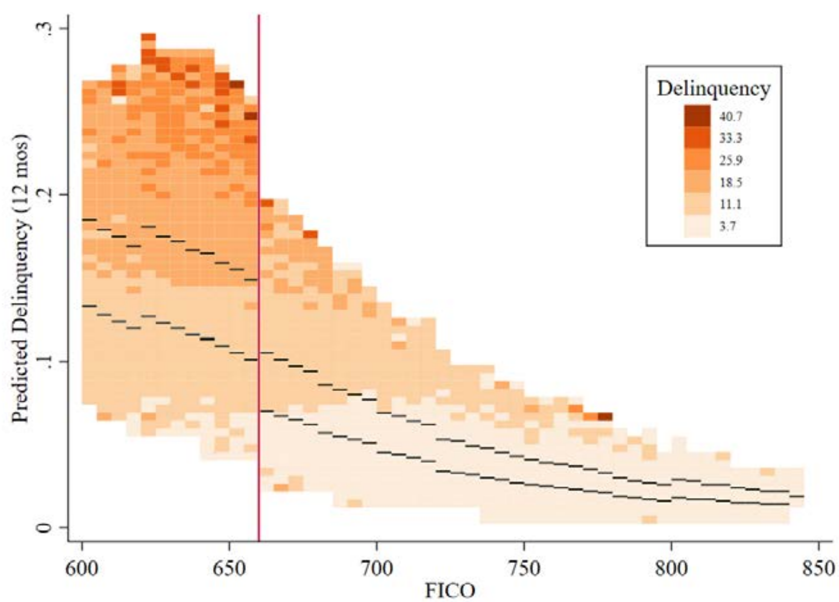
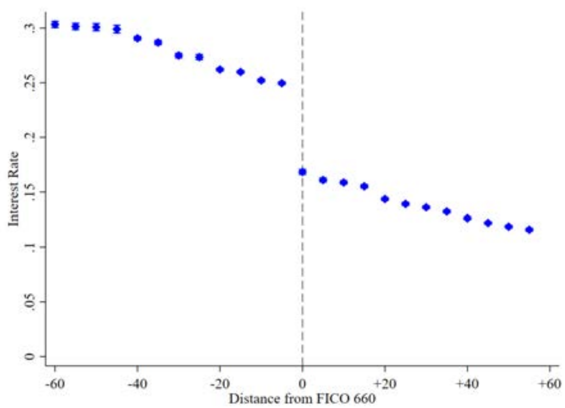


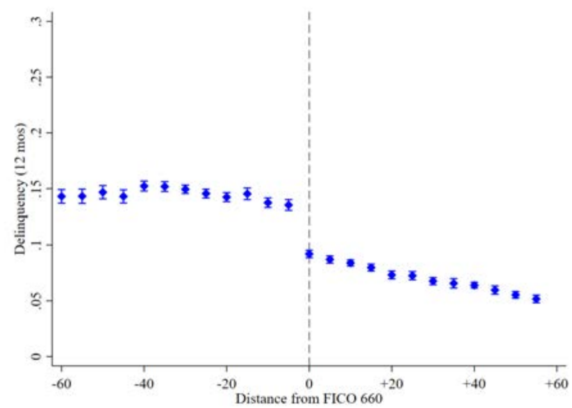
Figure 2. Discontinuities at the FICO 660 Cutoff

The figures show coefficients and standard errors from ordinary least squares regressions for five-point FICO indicators. The dependent variable in Panel (a) is the interest rate on loans. In Panel (b), the dependent variable indicates delinquency in the first 12 months. Lastly, in Panel (c), the dependent variable is the internal rate of return (IRR). In addition to FICO bins, the regressions include origination year-month fixed effects. All regressions are based on all loans in the sample. Standard errors are clustered by origination year-month.

(a) Interest Rate (%)



(b) Delinquency (%)



(c) IRR

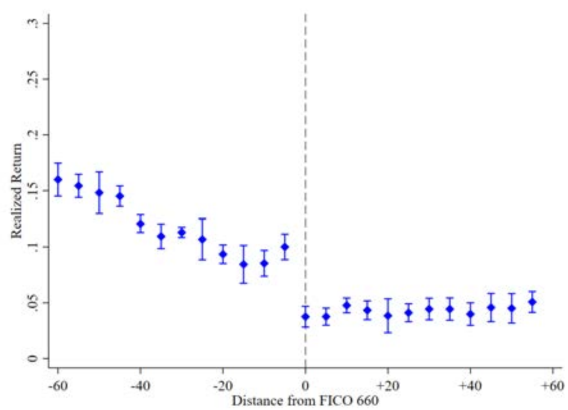
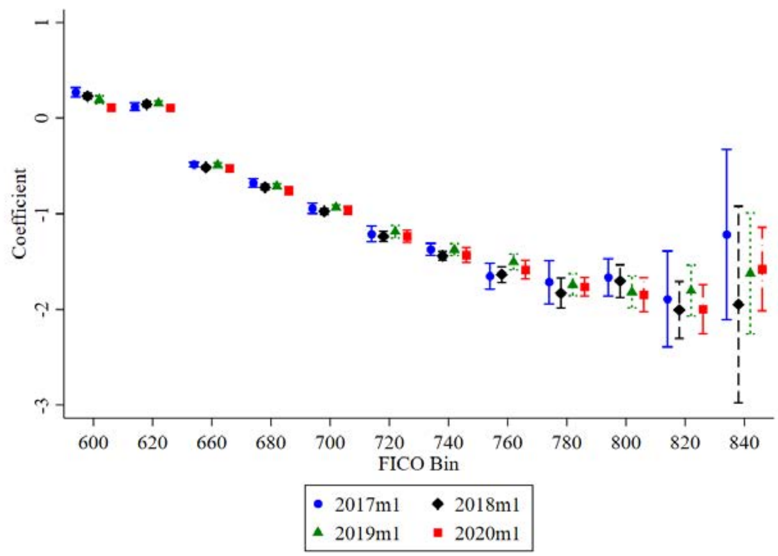


Figure 3. Credit Risk Model Coefficients over Time

This figure plots coefficients from rolling logit regressions used to predict the ex-ante probability of the loan becoming delinquent in the first 12 months for all loans originated from 2017 to 2020. The sample period for the regressions consists of 24 months of loan originations, with the most recent being 12 months before the prediction month. The dependent variable is an indicator that equals 1 if the loan became delinquent in the first 12 months. Select independent variables are plotted in two panels, with others shown in the Appendix, Figure A.3. Panel (a) shows the coefficients on 20-point FICO bin indicators, with the 640–659 FICO bin omitted. Panel (b) plots select coefficients from $5,000 income bins, with $55,000 as the omitted bin. Standard errors are clustered by origination month.

(a) FICO bins



(b) Income bins

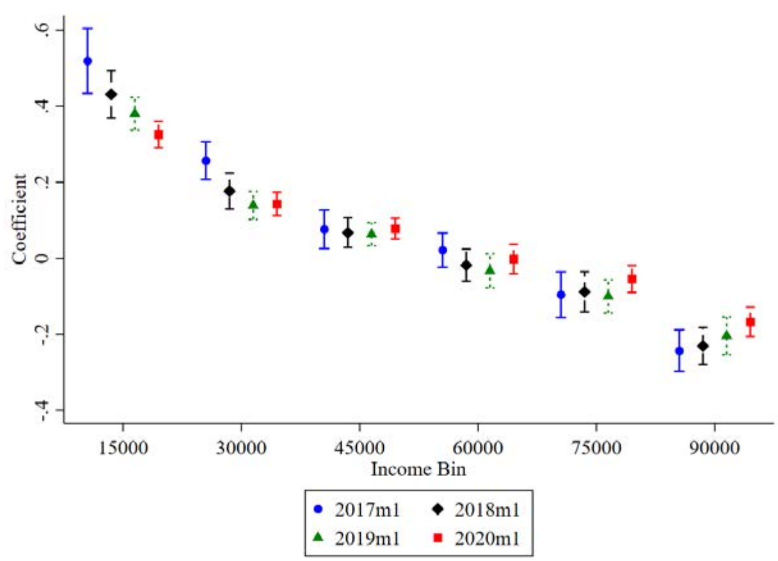
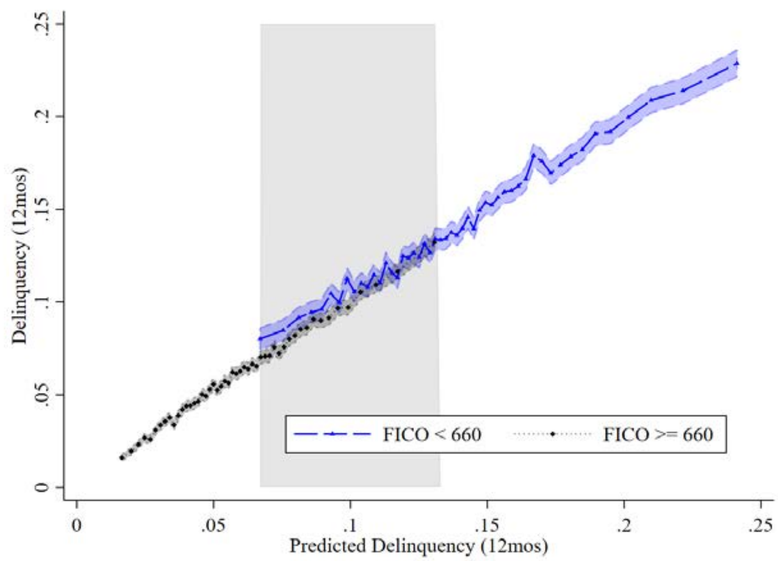


Figure 4. Performance of the Credit Risk Model

This figure compares the predicted loan performance relative to the actual performance for all loans in the sample, but separately for loans above and below FICO score 660. Predicted delinquency is estimated using rolling logit regressions from the prior three years of originations such that the prediction is out of sample. Panel (a) shows predicted delinquency relative to actual delinquency, with the overlapping region of risk between prime and nonprime highlighted in gray. Panel (b) shows the relationship between ex-ante risk and default for the subset of loans that have reached maturity. Default occurs when the lender ultimately charges off the loan.

(a) Delinquency (%)



(b) Default (%)

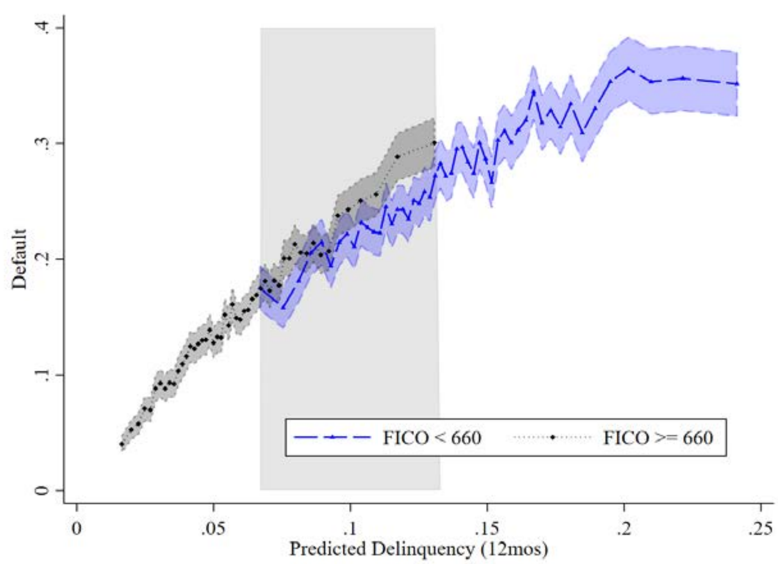
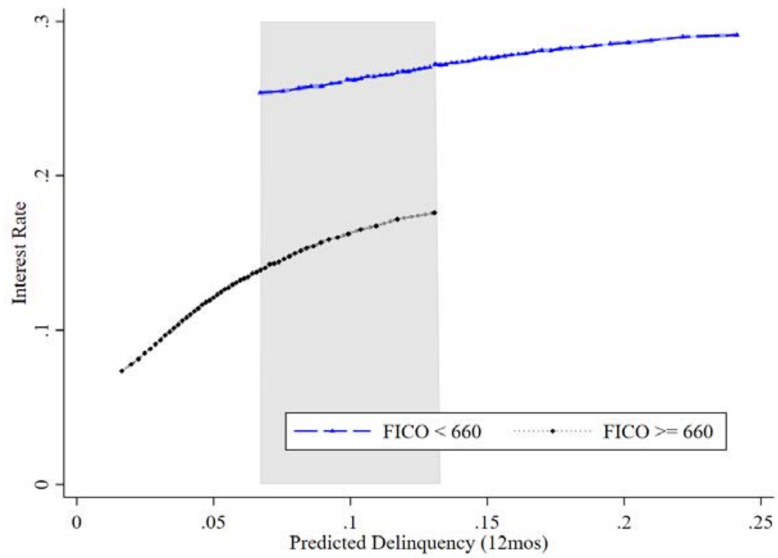


Figure 5. Ex-ante Predicted Risk and Rate

This figure shows the relationship between expected risk and pricing for prime and nonprime loans. Predicted delinquency is estimated using rolling logit regressions from the prior three years of originations such that the prediction is out of sample. Panel (a) plots the average interest rate and standard errors by expected risk separately for all prime and nonprime loans in the sample, with the overlapping region highlighted in gray. Panel (b) plots the loan rate difference between nonprime and prime loans in the overlapping region separately for 2017, 2018, and 2019.

(a) Interest Rate (%)



(b) Interest Rate Gap (%)

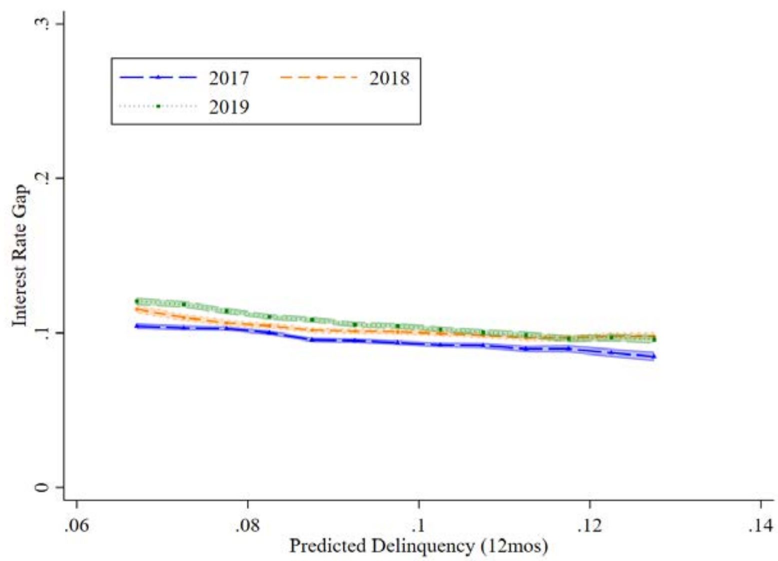


Figure 6. Ex-ante Predicted Risk and Realized IRR

This figure shows the relationship between realized loan-level IRR and predicted risk separately for all prime and nonprime loans in the sample. Plotted is the average IRR by predicted delinquency bins, along with standard errors. IRR is calculated using the initial outlay (loan amount) and subsequent cash flows (loan repayments) observed in the panel data. These cash flows also include loan recoveries when the loan is charged off. Only loans that reached maturity as of July 2020 are included in this sample.

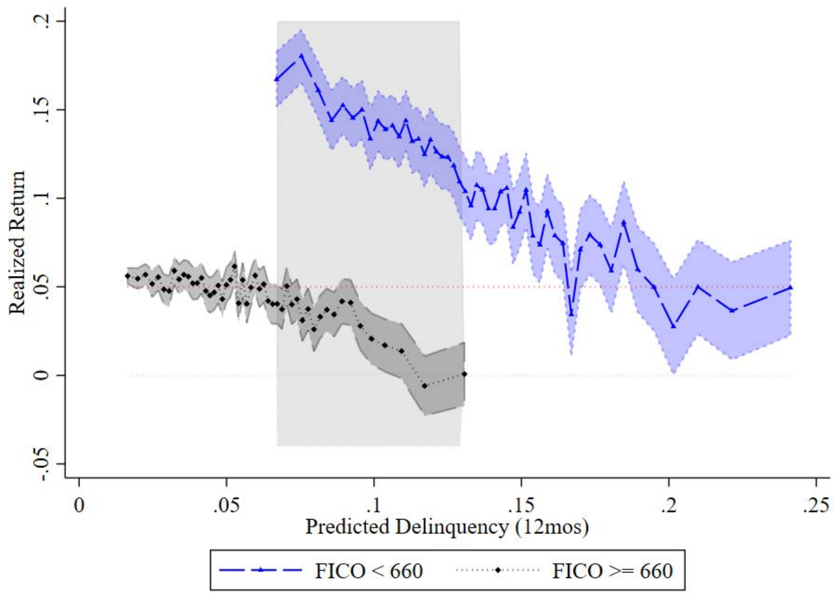
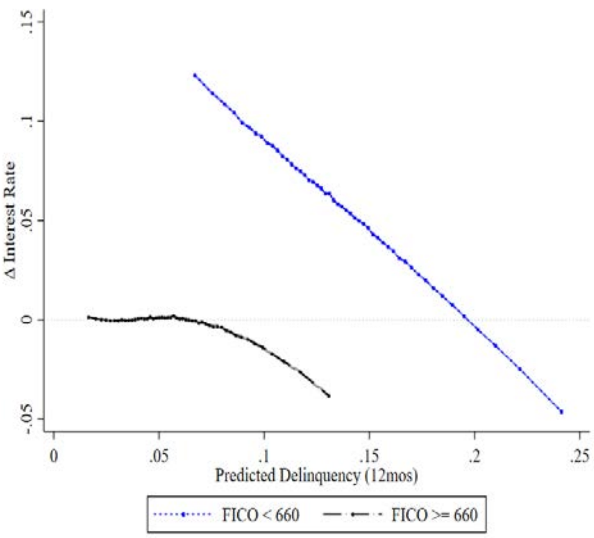


Figure 7. Magnitude and Distribution of Rate Differentials ($\Delta\widehat{r}_{it}$)

This figure shows the distribution of rate differentials $\Delta\widehat{r}_{it}$, defined as the difference between the actual rate and the calculated counterfactual rate for all prime and nonprime loans in the sample, respectively. The counterfactual interest rate is calculated using the coefficients from the pricing model of "safe" prime loans that exhibit ex-post constant returns in risk. Panel (a) shows the relationship between $\Delta\widehat{r}_{it}$ and predicted risk. Panel (b) shows the distribution of $\Delta\widehat{r}_{it}$ multiplied by the loan amount to demonstrate the dispersion in interest differentials in annual interest payments.

(a) Average Rate Differentials



(b) Distribution of Rate Differentials

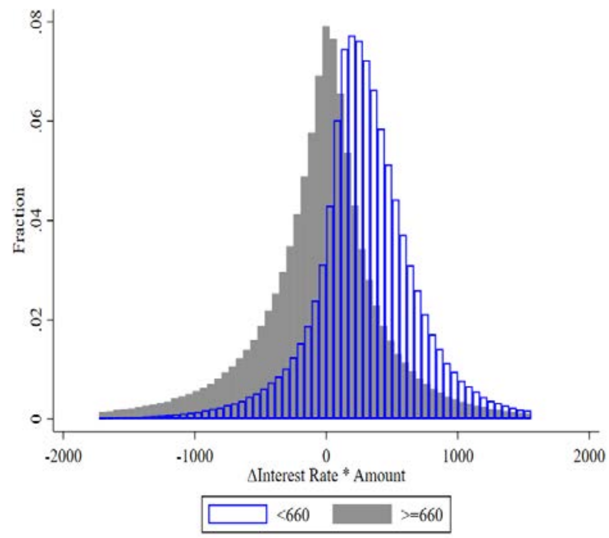


Table 1. Summary Statistics

This table presents summary statistics of borrower and loan characteristics. The data come from a FinTech loan aggregator specializing in FinTech lending. The sample is restricted to 36-month loans. *Rate (%)* is the annual interest rate. *Payment* is the monthly payment amount in dollars. *Age of Credit* is the age of their oldest credit in the credit report. *No. of Inquiries* is the number of credit inquiries the borrower made as reported by the credit bureau in the 6 months before origination. *Credit Utilization* is the fraction of the borrower's available revolving credit utilized at the origination time. *Default* occurs when a loan is charged off within its duration, and 12-month *Delinquency* occurs when the loan has a delinquent payment within one year since origination. *IRR* is the calculated internal rate of return using all cash flows at the loan level, excluding any financial or operating costs on the lender side.

	N	Mean	Median	SD	Min	Max
Interest Rate (%)	2,313,105	16.29	13.59	8.48	0.91	52.41
Loan Amount ($)	2,313,105	11,898	10,000	8,386	500	44,500
Term	2,313,105	36.00	36.00	0.00	36.00	36.00
Payment	2,313,105	410	334	274	44	1235
FICO	2,313,105	684	677	44	600	850
Income	2,313,105	73,446	62,500	43,594	16,812	250,000
Payment/Income, annual %	2,313,105	0.07	0.07	0.04	0.01	0.21
Age of credit	2,313,105	194	175	92	41	483
No. of Inquiries	2,313,105	0.71	0.00	1.03	0.00	5.00
Credit Utilization	2,313,105	0.53	0.54	0.25	0.01	0.99
I(Loan Purpose: Debt Consolidation)	2,313,105	0.76	1.00	0.43	0.00	1.00
I(Loan Purpose: Home Improvement)	2,313,105	0.07	0.00	0.25	0.00	1.00
I(Homeowner)	2,313,105	0.55	1.00	0.50	0.00	1.00
Default (%)	2,313,105	14.44	0.00	35.15	0.00	100.00
Delinquency (%)	2,313,105	8.12	0.00	27.31	0.00	100.00
Pre-Payment (%)	2,313,105	45.85	0.00	49.83	0.00	1.00
IRR (%)	1,138,365	6.15	12.12	29.74	-100.00	41.80

Table 2. RDD Regression

This table estimates the difference in loan interest rates and delinquency for borrowers who fall just below the prime threshold of FICO 660. *I(FICO <660)* identifies borrowers with FICO scores below 660, and the coefficient represents the jump in loan rates for borrowers below the threshold. RDD controls include *FICO−660* and its interaction with the indicator, which allow for the slopes to differ on either side of the threshold with respect to FICO. All coefficients are multiplied by 100 for ease of reporting. Columns (1) and (4) use a bandwidth of ±40 FICO points; Columns (2) and (5) use ±20 points; and Columns (3) and (6) use ±5 points. All regressions include a control for the log of loan amount and fixed effects for the state, origination year-month, and state-by-year. Standard errors are clustered by origination month. *** indicates significance at the 1% level.

Dep Var	Interest Rate (%)			I(Delinquency)		
	(1)	(2)	(3)	(4)	(5)	(6)
I(FICO < 660)	8.57***	9.13***	9.43***	4.45***	4.53***	4.88***
	(0.12)	(0.12)	(0.12)	(0.19)	(0.23)	(0.22)
Sample Restriction	620-699	640-679	655-664	620-699	640-679	655-664
RD Controls	Yes	Yes	No	Yes	Yes	No
Loan Attributes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Origination YM FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1,457,316	822,252	213,017	1,457,316	822,252	213,017
R^2	0.55	0.49	0.64	0.02	0.02	0.13

Table 3. Rate Gap in the Overlapping Sample

This table shows estimates from ordinary least squares regressions of loan rates on nonprime dummy based on loans in the overlapping region, i.e., those with predicted delinquency between 6% and 13%. $I(\widehat{Delinquency})$ is the predicted delinquency rate based on a real-time credit risk model specified in Equation (2). The sample includes loans with predicted delinquency between 6% and 13%, where prime and nonprime loans are overlapped. The dependent variable is the loan interest rate. Specifications (1) and (3) include $I(\widehat{Delinquency})$ as a continuous variable, while specification (2) uses $I(\widehat{Delinquency})$ as a categorical variable, with groups at every .05% of predicted delinquency. Standard errors are clustered by origination year-month. *** indicates significance at the 1% level.

Dep Var	Interest Rate (%)		
	(1)	(2)	(3)
I(FICO $<$ 660)	7.33***	7.16***	6.22***
	(0.08)	(0.07)	(0.09)
$I(\widehat{Delinquency})$	71.37***		64.99***
	(1.19)		(1.67)
Log(Loan Size)	-2.72***	-2.71***	-2.66***
	(0.06)	(0.06)	(0.06)
$I(\widehat{Delinquency})$ Bin FE	No	Yes	No
Origination YM FE	Yes	Yes	Yes
Sample	All Loans	All Loans	FICO 640-680
N	848,552	848,552	512,484
R^2	0.57	0.57	0.44

Table 4. Prepayment Risk and Rates

This table shows estimates of the effect of prepayment risk on rate differential between prime and nonprime loans based on loans in the overlapping region, i.e., those with predicted delinquency between 6% and 13%. The dependent variable in Columns (1)-(3) is an indicator equal to 100 if the loan is paid off in full before maturity at 36 months. In Columns (4)-(6), the dependent variable is the loan interest rate. The primary independent variable is an indicator of nonprime status while controlling for predicted delinquency, loan size and predicted prepayment risk. All specifications include origination year-month fixed effects. Standard errors are clustered by origination year-month. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Dep var	I(Prepayment)			Interest Rate (%)		
	(1)	(2)	(3)	(4)	(5)	(6)
I(FICO < 660)	2.85***	2.94***	3.12***	7.33***	7.05***	6.22***
	(0.25)	(0.27)	(0.25)	(0.08)	(0.08)	(0.09)
$\widehat{I(Delinquency)}$	-58.08***	-59.16***	-59.47***	71.38***	72.93***	64.99***
	(7.34)	(7.53)	(8.58)	(1.17)	(1.13)	(1.64)
Log(Loan Size)	-1.41***		-1.39***	-2.72***		-2.66***
	(0.15)		(0.19)	(0.06)		(0.06)
$\widehat{I(Prepayment)}$				0.00	0.00	0.00
				(0.00)	(0.00)	(0.00)
Loan Size Bin FE	No	Yes	No	No	Yes	No
Origination YM FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All Loans	All Loans	FICO 640-680	All Loans	All Loans	FICO 640-680
N	848,552	848,552	512,484	848,552	848,552	512,484
R^2	0.13	0.13	0.13	0.57	0.58	0.44

Table 5. Performance of Nonprime Loans During COVID

This table estimates the impact of COVID-19 on delinquency in a difference-in-differences framework using loan panel data. The sample tracks the history of any loan outstanding as of October 2019 that are in the overlapping region from October 2019 through July 2020. The dependent variable is an indicator equal to 100 if the loan misses a payment in that month and 0 otherwise. The main explanatory variable, *I(<660)* × *COVID*, is the interaction of the nonprime loan dummy with a post-COVID-19 dummy equal to 1 if the performance date is after February 2020. All regressions include a borrower and performance date fixed effects. Other specifications include state-by-month, predicted delinquency-by-month, and origination date-by-month fixed effects. Standard errors are clustered by performance date-by-state. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Dep Var	$\text{I(Delinquency)}_{i,t}$		
	(1)	(2)	(3)
$\text{I}(\text{FICO}_i < 660) \times \text{I}(\text{COVID}_t)$	-0.43*	-2.08***	-2.24***
	(0.24)	(0.30)	(0.20)
Borrower FE	Yes	Yes	Yes
Calendar Time FE	Yes	Yes	Yes
$\widehat{\text{I(Delinquency)}}_{\text{i,0}}$ × Calendar Time FE	No	Yes	Yes
State × Calendar Time FE	No	No	Yes
Origination YM × Calendar Time FE	No	No	Yes
N	5,182,793	5,182,793	5,182,793
R^2	0.32	0.32	0.52

Table 6. Hand-Collected Information on Lender's Minimum FICO Rules

This table identifies the minimum FICO scores that banks and online lenders will consider in credit underwriting decisions for unsecured personal loans. The 15 largest banks by asset size are reported in bold, and other bank lenders with a significant presence in this credit market are included. For online lenders, the seven largest lenders by origination volume are in bold. These minimums are identified using personal finance websites (e.g., WalletHub or NerdWallet) when the data are unavailable through the institution's own website. Data collection was performed in November 2021.

Min. FICO	580	600	620	640	660	680	700
Bank					**Wells Fargo** American Express **SunTrust (LightStream)** Discover **Goldman Sachs (Marcus)**	**Citibank** **U.S. Bank** **Fifth Third Bank**	**TD Bank** **HSBC** Barclays
Online Lender	**Avant**	**Best Egg** **LendingClub** **LendingPoint** **Upstart**	Freedom Plus Upgrade	Payoff Rocket Loans **Prosper**		**SoFi**	

Bold represents the largest banks (15) and largest FinTech lenders (7).
The following do not offer personal unsecured loans: Chase Bank, Bank of America, Capital One, Bank of New York Mellon, and State Street Bank.

Table 7. Relationship Between Interest Rates, Returns, and Predicted Risk

This table estimates the sensitivity of loan interest rates and returns to predicted risk for various samples. The dependent variable is the loan interest rate in Panel A and the loan-level realized return in Panel B. The independent variable is the predicted delinquency from the credit risk model. Columns (1) and (2) use all nonprime and prime loans, respectively. Columns (3)–(5) use only loans in the overlapping region where prime and nonprime loans overlap in risk. Column (5) includes an indicator for nonprime loans and their interaction with predicted delinquency. Column (6) shows the calibration sample of safe prime loans used to determine counterfactual rates. In Panel B the sample is restricted to loans that have full performance data for the calculation of returns. Origination year-month fixed effects are included in all regressions. Standard errors are clustered by origination year-month. *** indicates significance at the 1% level.

Panel A: Interest Rate

Dep Var	Interest Rate (%)					
Sample	All		Overlap Region			Calibration
	Nonprime	Prime	Nonprime	Prime	Both	Sample
	(1)	(2)	(3)	(4)	(5)	(6)
$I(\widehat{Delinquency})$	22.37***	89.36***	25.78***	60.92***	60.52***	116.63***
	(1.09)	(0.97)	(2.39)	(1.19)	(1.13)	(4.52)
× I(FICO < 660)					-33.81***	
					(2.41)	
I(FICO < 660)					13.98***	
					(0.33)	
Origination YM FE	Yes	Yes	Yes	Yes	Yes	Yes
N	827,441	1,272,659	340,924	507,628	848,552	293,596
R^2	0.04	0.23	0.01	0.06	0.49	0.07

Panel B: Realized Return

Dep Var	IRR (%)					
Sample	All		Overlap Region			Calibration
	Nonprime	Prime	Nonprime	Prime	Both	Sample
	(1)	(2)	(3)	(4)	(5)	(6)
$I(\widehat{Delinquency})$	-105.79***	-57.81***	-105.86**	-95.10***	-95.31***	-7.28
	(6.56)	(5.33)	(15.48)	(8.78)	(8.76)	(9.69)
× I(FICO < 660)					-9.97	
					(17.80)	
I(FICO < 660)					12.53**	
					(1.77)	
Origination YM FE	Yes	Yes	Yes	Yes	Yes	Yes
N	80,292	122,992	37,995	46,015	84,010	29,224
R^2	0.01	0.00	0.00	0.00	0.02	0.00

Appendix

Figure A.1. Interest Rates by FICO Based on an Alternative Sample

This figure plots the average loan interest rates by FICO score bins separately for the entire dataset ("All Loans") and the sample that we use throughout the paper ("Filtered Sample"). See subsection 2.2 for a complete description of the construction of the filtered sample. *Interest Rate* is the interest rate charged on loans and does not include origination fees in the filtered sample but may include fees in the "all loans" sample. FICO scores are binned at intervals of five points.

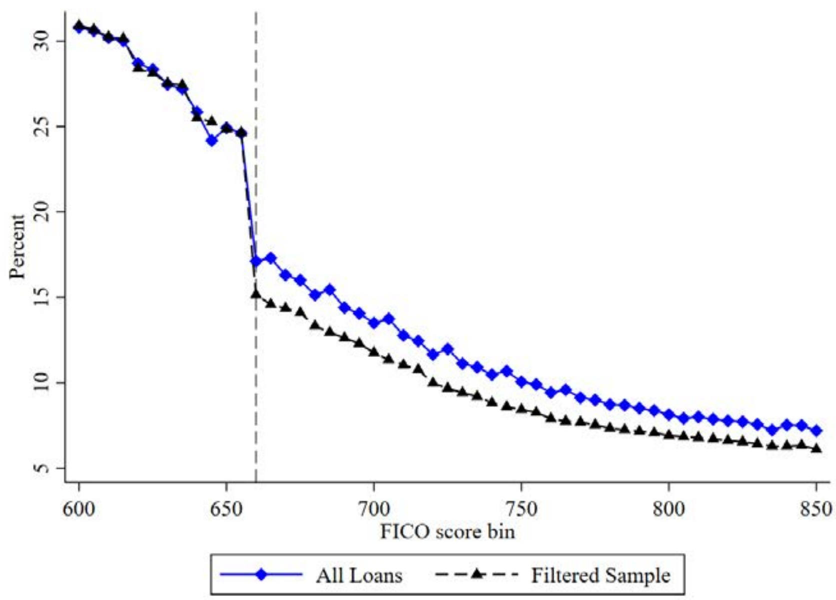
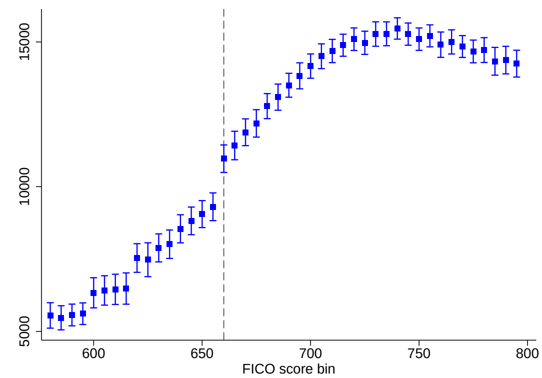


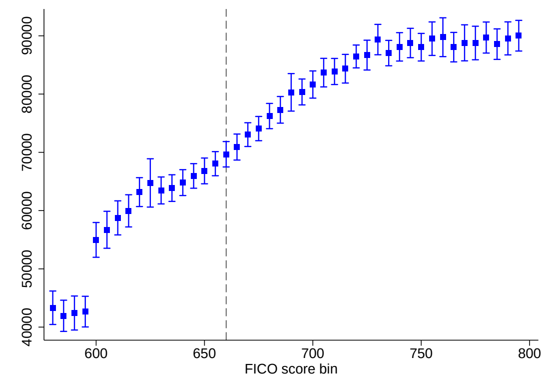
Figure A.2. More RDD Regressions

This figure shows the discontinuity at FICO 660 for loan characteristics and the borrowers' monthly income. The dependent variables are loan size in Panel (a), borrower income in Panel (b), and monthly payment in Panel (c). The corresponding dependent variable is regressed on FICO score bins for each panel, controlling for originating month fixed effects. All regressions are based on all loans in the sample. Standard errors are clustered at the monthly level. The coefficients are plotted on credit score bins, along with the confidence intervals at the 95% level.

(a) Loan Amount ($)



(b) Income ($)



(c) Monthly Payment ($)

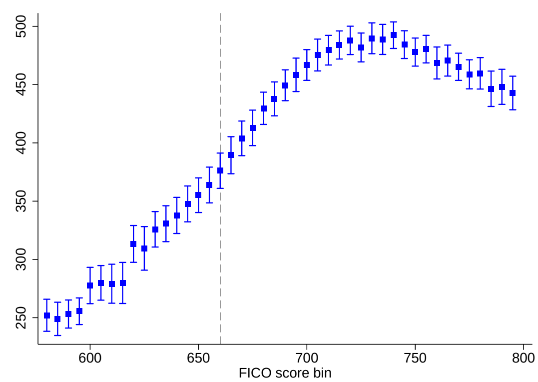
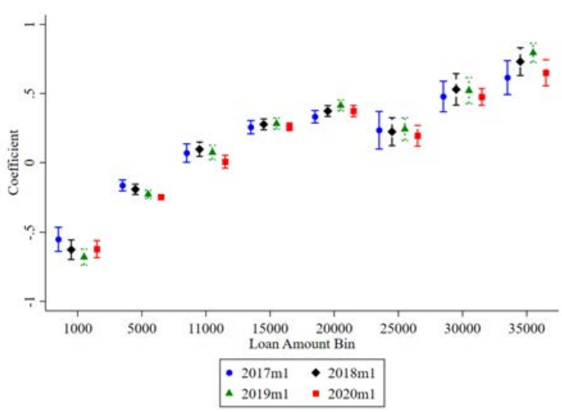


Figure A.3. More Credit Risk Model Coefficients over Time (Figure 3 Continued)

This figure plots other coefficients from logit rolling regressions used to predict the ex-ante delinquency risk. The estimation sample consists of 24 months of loan originations, the most recent being 12 months before the prediction month. The dependent variable is an indicator equal to 1 if the loan became delinquent in the first 12 months. Panel (a) plots select coefficients for $1,000 loan amount bins, with the $10,000 bin omitted. Panel (b) plots coefficients for other credit bureau information and indicators for the largest reported loan purposes—credit card refinancing, home improvement, and medical—with debt consolidation as the omitted category. $\Delta FICO$ is the borrower's FICO score minus the midpoint of their 20-point FICO bin. *Cr. Age* is the age of their oldest credit in the credit report. *# Inq.* is the number of credit inquiries the borrower made in the six months prior to origination. *Cr. Util.* is the ratio of the borrower's revolving credit utilization at the time of origination. Standard errors are clustered by origination month.

(a) Loan Amount



(b) Loan Purpose & Credit Bureau Variables

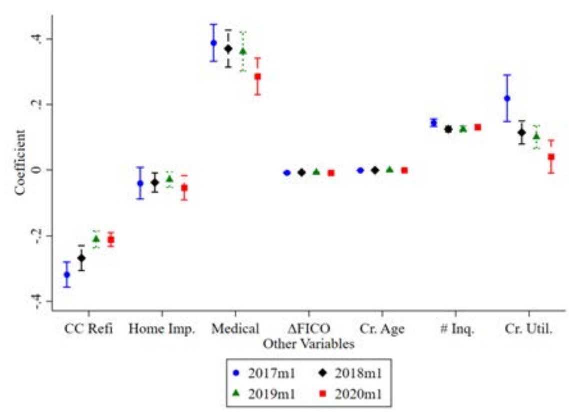


Figure A.4. Distribution of Interest Rates by Market Segments

This figure shows the distribution of interest rates separately for prime and nonprime loans.

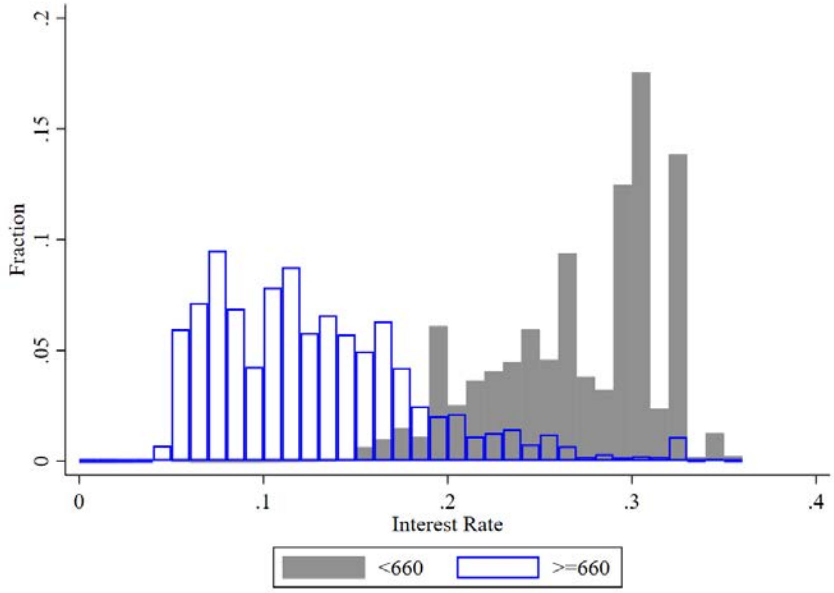


Table A.1. Additional Summary Statistics

This table presents summary statistics of borrower and loan characteristics.Panel (a) reports summary statistics of the loan panel used in Section 3.4. The sample tracks the history of any loan outstanding as of October 2019 that are in the overlapping region from October 2019 through July 2020. The subsample is further divided into before and during/post-COVID periods. Panel (b) reports statistics on all loans. The sample is divided into four segments: non-overlapping prime, overlapping prime, overlapping nonprime and non-overlapping nonprime. Overlapping prime and overlapping nonprime loans make up the overlapping region/sample.

Panel A: Loan by Month Panel Sample

	Overlapping Prime				Overlapping Nonprime			
	Pre-COVID		Post-COVID		Pre-COVID		Post-COVID	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I(Delinquency)$_{i,t}$ (%)	10.18	30.24	12.67	33.26	11.56	31.98	12.90	33.52
Rate$_{i,0}$	15.04	5.65	14.98	5.64	26.32	4.79	26.29	4.80
Loan Amount$_{i,0}$	12,265	7,597	12,374	7,623	5,150	3,202	5,135	3,191
FICO$_0$	681.43	17.37	681.76	17.50	637.91	14.98	637.75	15.09
$\widehat{\text{I(Delinquency)}}_{i,0}$	8.81	1.58	8.81	1.59	10.90	1.54	10.90	1.54
N	1,922,547		1,205,912		1,277,572		776,962	

Panel B: Individual Loan Sample by Subsamples

	Non-overlapping Prime		Overlapping Prime		Overlapping Nonprime		Non-overlapping Nonprime	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Rate (%)	11.17	4.96	15.45	5.70	26.45	4.70	27.90	4.43
Loan Amount (\$)	9,479	6,788	12,102	7,583	5,344	3,345	7,868	4,382
Term	36.00	0.00	36.00	0.00	36.00	0.00	36.00	0.00
Payment	307	215	417	253	216	133	324	179
FICO	719	37	680	17	638	15	633	15
Income	58,408	20,318	53,443	20,218	54,035	19,819	47,621	19,318
Payment/Income	0.07	0.04	0.10	0.05	0.05	0.03	0.09	0.04
Age of credit	200	98	165	82	181	87	149	73
No of Inquiries	0.43	0.74	0.82	1.12	0.57	0.85	1.18	1.35
Credit Utilization	0.40	0.24	0.53	0.23	0.63	0.25	0.65	0.24
I(Debt Consolidation)	0.76	0.43	0.76	0.42	0.68	0.47	0.66	0.48
I(Home Improvement)	0.07	0.25	0.04	0.20	0.07	0.25	0.05	0.21
I(Homeowner)	0.56	0.50	0.47	0.50	0.45	0.50	0.42	0.49
I(Default) (%)	7.54	26.41	14.05	34.75	15.89	36.56	22.68	41.88
I(Delinquency) (%)	4.60	20.95	8.98	28.59	11.36	31.73	16.57	37.18
I(Prepayment) (%)	41.41	49.26	35.52	47.86	39.02	48.78	34.27	47.46
Prepayment Month	15.09	9.17	14.89	9.11	13.94	9.11	13.57	8.83
IRR (%)	5.10	22.95	3.11	31.64	8.19	42.49	13.62	36.71
N	765,031		507,628		340,924		486,517	