FinTech Lending with LowTech Pricing*

Mark J. Johnson, Itzhak Ben-David, Alex Chinco, Jason Lee, and Vincent Yao†

October 10, 2024

Abstract

Technology is disrupting the market for personal loans. This paper investigates how. Nonprime borrowers with credit scores below 660 have traditionally found it difficult to get financing. FinTech lenders claim to be using big data and machine learning to identify and offer bespoke loan terms to the safest nonprime borrowers. However, when we examine a comprehensive dataset of FinTech personal loans, we find that all nonprime borrowers pay roughly the same 25% annual interest rate, no matter what additional personal information they provided at the time of origination. FinTech lenders are approving nonprime loans that lenders have traditionally deemed too risky, but their loan pricing is not risk-based as touted. Perhaps their core disruptive technology is a mobile app, which allows a borrower to get funds in hours rather than days and all without having to leave the house or put pants on. This is a perfectly valid way to shake up an industry, but it is not what is being advertised and comes with a different set of policy concerns.

Keywords: FinTech, Personal Loans, Credit Score, Nonprime Borrowers

JEL Classification: G21, G23, G50

---

*We thank Avramidis Panagiotis, René Stulz, Francesco D'Acunto (discussant), and the seminar audiences at the CFPB Research Conference, Brigham Young University, OCC FinTech Symposium, Purdue FinTech Conference, 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.

†Mark J. Johnson: Brigham Young University, Marriott School of Business, markjjohnson@byu.edu. Itzhak Ben-David: Fisher College of Business, The Ohio State University, and the National Bureau of Economic Research, ben-david.1@osu.edu. Alex Chinco: Baruch College, alexchinco@gmail.com. Jason Lee: U.S. Securities and Exchange Commission, jason.jaehyun.lee@gmail.com. Vincent Yao: Georgia State University, J. Mack Robinson College of Business, wyao2@gsu.edu.

Introduction

FinTech lenders are using technology to disrupt the market for personal loans. Lenders originate around $150b in unsecured personal loans each year in the US. In 2015, only about 22% of this money came from FinTech lenders. By 2019, over 49% of all loan originations were done online. The first FinTech lenders were started in 2005. There are now multiple FinTech lenders with billion-dollar valuations. Upstart is sitting at $3.8b, and LendingClub has a market value of $1.4b.

The rapid growth of the FinTech lending industry has caught the attention of policymakers. When asked, the lenders claim that they are applying advanced statistics to large amounts of data to identify creditworthy borrowers who happen to have bad credit scores (aka, "FICO" scores). Nonprime borrowers (FICO $<$ 660) have traditionally had a hard time getting access to financing. FinTech lenders say that, by looking at more than just a borrower's credit score, they are able to profitably lend to this underserved group.

Here is how Nathaniel Hoopes, executive director of the Marketplace Lending Association, put it: "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."

This is an attractive narrative for policymakers. In the words of US House Representative Trey Hollingsworth: "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 types of products." If FinTech lenders are using big data and machine learning to give underserved groups better access to financing, then the best thing a regulator can do is get out of their way.

It is also an attractive narrative for academics. We like analyzing large amounts of data and developing new statistical methods. This is where our

passions are. It would be convenient for us if FinTech lenders had found a way to use these tools to make the world a better place. It would suggest researchers could benefit society as a whole by pursuing their own interests.

But does this narrative describe what FinTech lenders are actually doing? To answer this question, we analyze a database containing 70% of the US personal loans made by FinTech lenders from 2014 to 2020. Our data contain the terms of each loan (e.g., loan amount and interest rate), the loan's subsequent performance (e.g., delinquency and default), and various borrower characteristics (e.g., FICO score, income, loan purpose, home-ownership status, number of credit checks in the past 6 months, etc). We describe these data in Section 1.

Crucially, our data only includes originated loans. As a result, there are some questions that we cannot speak to. For example, should policymakers make it easier for more people to take out personal loans from a FinTech lender? This question concerns the loans FinTech lenders might have made but did not—i.e., offers that were turned down by the borrower, applications that were rejected by the lender, and loans that were needed but never applied for. It does not make sense to extrapolate from the loans in our sample to these unobserved cases. There was a reason these unobserved loans were not originated.

But hypothetical policy expansions are not the only interesting angle worth exploring! In Section 2, we show that FinTech lenders clearly have lower lending standards than traditional banks. They could be offering bespoke loan terms to the safest nonprime borrowers by looking at more than just an applicant's credit score. But are FinTech lenders actually doing this? This is an important question that is right in our data's proverbial wheelhouse. It is about whether the observed data are consistent with a particular story.

FinTech lenders collect lots of personal information from each applicant—e.g., income, loan purpose, home-ownership status, number of recent credit checks, etc. Many of these variables are highly predictive of future default, even within groups of borrowers with identical FICO scores. If FinTech lenders are using this information to offer better terms to the safest nonprime borrowers, then we should see interest rates vary across nonprime loans. For example, the average nonprime borrower with a $3k personal loan should have a lower

interest rate than the typical nonprime borrower with a $15k loan. By contrast, our data indicate that all nonprime FinTech personal loans have roughly the same 25% annual interest rate.

To provide a measuring stick, we start in Section 3 by studying loans within ±5 points of the prime/nonprime threshold. The average almost-prime borrower (FICO $\in 655, 656, 657, 658, 659$) is hardly any riskier than the average barely-prime borrower (FICO $\in 660, 661, 662, 663, 664$). Nevertheless, the left panel of Figure 1 shows that borrowers who are just below the threshold get charged 8.1%pt higher annual interest rate, 16.9% vs 25.0%. As a result, the right panel shows that almost-prime borrowers are 3.9%pt more likely to miss a payment during the first twelve months, 8.0% vs 11.9%.

To put this in perspective, note that the average loan in our sample is $12k. This implies that an almost-prime borrower with a 25% annual interest rate would have to pay $477 per month to take out an average-sized personal loan; whereas, the average barely-prime borrower with a 16.9% rate would only have to pay $427 per month. In other words, first-year delinquency rates go up by 3.9%pt in response to a $50 bump in the monthly payment.

Among the group of borrowers with FICO $\in \{655, 656, 657, 658, 659\}$, some people are much safer than others. The right panel of Figure 1 shows that almost-prime borrowers with small loans (below 25%ile; roughly $3k) are $15.9\% - 9.3\% = 6.6$%pts less likely to go delinquent in the first year than those with large loans (above 75%ile; roughly $15k). Almost-prime borrowers with two or more credit checks in the past 6 months are $13.4\% - 10.9\% = 2.5$%pts more likely to miss one of the first twelve payments compared to an almost-prime borrower with no recent credit inquiries. Among almost-prime borrowers, homeowners are $12.4\% - 11.1\% = 1.3$%pts less likely to go delinquent in the first twelve months than renters. In the final row, we use all these variables to predict how likely each almost-prime borrower is to go delinquent. Those in the riskiest quartile are $16.5\% - 8.3\% = 8.2$%pts more likely to be delinquent in the first year than the safest quartile.

Given how sensitive delinquency rates are, it makes sense why it would be helpful to collect additional data about each applicant's financial situation. It

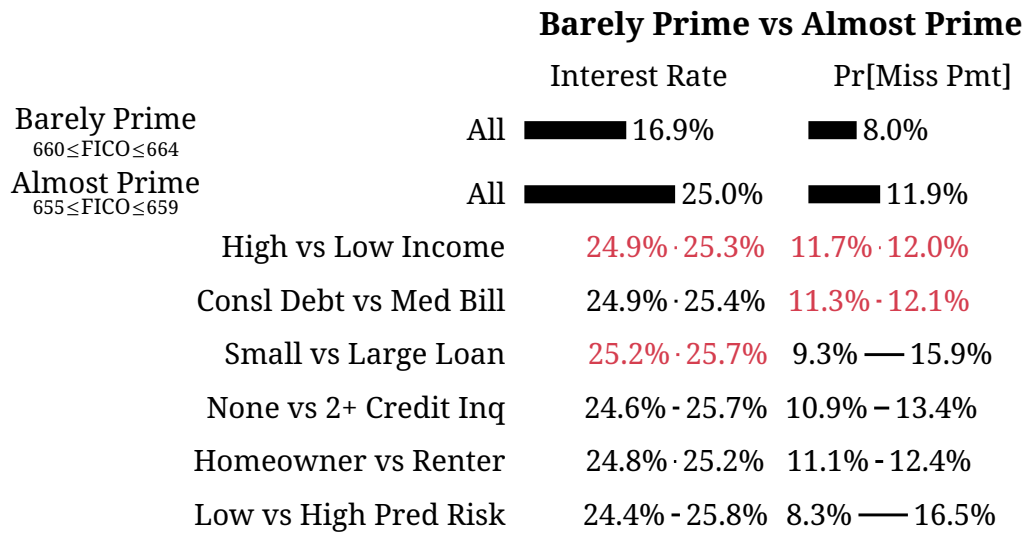


Figure 1. *The top row shows the average interest rate of FinTech personal loans to barely-prime borrowers (FICO* $\in$ 660, 661, 662, 663, 664*) and the fraction of these loans that missed a payment during the first year. The second row shows the same statistics for almost-prime borrowers (FICO* $\in$ 655, 656, 657, 658, 659*). The remaining rows show the differences in average interest and delinquency rates between various subsets of almost-prime borrowers. High vs Low Income: borrowers above 75th percentile of the almost-prime income distribution (high income) vs those below the 25th percentile (low income). Consl Debt vs Med Bill: loans used to consolidate debt vs those used to pay medical expenses. Small vs Large Loan: loans below the 25th percentile of the almost-prime loan-size distribution (small loans) vs those above the 75th percentile (large loans). None vs 2+ Credit Inq: borrowers with zero credit checks during past* 6 *months vs those with two or more. Homeowner vs Renter: borrowers who own their home vs those who rent. Low vs High Pred Risk: borrowers below the 25th percentile of the predicted delinquency-rate distribution when using all available borrower characteristics (safe) vs those above the 75th percentile (risky). Red indicates that the safer group (listed first) had a higher average.*

is possible to use information other than a borrower's credit score to identify more creditworthy almost-prime borrowers using the data FinTech lenders collect at the time of origination. FinTech lenders could be offering these people better loan terms. However, there is just no evidence that they are.

In the left panel of Figure 1, we see hardly any difference between the rates charged to safe and risky almost-prime borrowers. Comparing the first and last rows in Figure 1 really drives the point home. The bottom right row shows that the safest quartile of almost-prime borrowers are no more likely to miss one of their first twelve payments than the average barely-prime borrower, 8.3% vs 8.0%. Nevertheless, the typical barely-prime borrower pays a 24.4% − 16.9% = 7.5%pt lower annual interest rate. Borrowers have to provide lots of personal data when applying for a loan. There is very little evidence that FinTech lenders are using any of it for pricing purposes. All FinTech personal loans involving almost-prime borrowers have roughly the same 25% annual interest rate.

No matter how selection works, it is hard to square this evidence with FinTech lenders' claims about going out of their way to offer better loan terms to the safest nonprime borrowers. In Section 4, we model the endogenous selection problem to highlight precisely why. When almost-prime borrowers submit an application, sometimes a FinTech lender will return to them with an unacceptable loan offer. FinTech lenders could offer borrowers different terms on the loan offers that get refused. In principle, this could explain why the average annual interest rate and delinquency rate are the same for both high- and low-income almost-prime borrowers in Figure 1. However, it cannot explain why the average annual interest is the same for both $3k and $15k almost-prime loans, even though smaller loans are 6.6%pts less likely to miss a payment.

In Section 5, we extend these results to the entire population of nonprime FinTech personal loans—i.e., for all values of FICO < 660. We build a simple delinquency-risk model for nonprime personal loans that incorporates other information like income, loan amount, and homeownership in addition to a borrower's credit score. Our model yields accurate and unbiased first-year delinquency predictions for the entire nonprime borrowers population. We find

that twelve-month delinquency rates are 3.5× higher for the riskiest nonprime borrowers, 7% vs 24%. However, they pay just 4%pts most in annual interest, 25% vs 29% per year.

Our paper contains important lessons for both policymakers and researchers. Because we only observe originated loans, we cannot directly answer questions about the welfare implications of giving more people personal loans from FinTech lenders. However, it is noteworthy FinTech lenders have chosen to advertise a completely different business model to regulators. The lenders themselves clearly want to be viewed as companies that give underserved groups access to financing by applying cutting-edge machine-learning algorithms to big data. In fact, our data suggest that their core business model involves using a mobile app to offer the traditional FICO-based loans in a faster and more convenient package.

FinTech lenders could be using big data and machine learning to profitably lend to nonprime borrowers. We are not the first to show that some of these individuals are creditworthy (Jagtiani and Lemieux, 2019; Berg, Burg, Gombović, and Puri, 2020; Dolson and Jagtiani, 2021; Di Maggio, Ratnadiwakara, and Carmichael, 2021; Fuster, Goldsmith-Pinkham, Ramadorai, and Walther, 2022). FinTech lenders might be offering this underserved group an important new financing option (Hau, Huang, Shan, and Sheng, 2019; Di Maggio and Yao, 2021; Agarwal, Kigabo, Minoiu, Presbitero, and Silva, 2023). But they do not have to be. Previous research has documented other benefits to online lending, such as increased operational efficiency (Buchak, Matvos, Piskorski, and Seru, 2018; Fuster, Plosser, Schnabl, and Vickery, 2019; Berg, Burg, Gombović, and Puri, 2020).

Our results also contain an important lesson for financial economists. At the moment, FinTech researchers take it for granted that, to revolutionize an industry, a company must do something sophisticated, such as develop a self-driving car. But it is worth remembering that Uber's key technology is a garden-variety mobile app. Our data suggest that this is also the key technology behind FinTech lenders' business model. Moreover, we suspect that they are not the only kind of FinTech firm that is doing something much more mundane

than researchers would like to believe.

1 Data Description

This section describes where our data comes from, how we cleaned it, and what the summary statistics look like. The raw data includes 7.4m loans originated between January 2014 and July 2020 by seven of the largest online lenders. The data comes from a leading FinTech loan aggregator and includes loans that are funded by the lenders themselves as well as those capitalized by outside investors. We estimate that our data contains roughly 70% of all FinTech personal loans, given the quarterly volumes provided by S&P Global Market Intelligence in February 2021.

For each originated loan, we observe three kinds of variables. First, we see data about the borrower's characteristics that was collected by the FinTech lender prior to making the loan: FICO score, annual income, credit utilization rate, home-ownership status, and the intended purpose of the loan purpose. While we observe a relatively rich set of borrower characteristics, it is only a subset of what lenders collect. Second, we see the terms of each loan: amount borrowed, maturity, and annual interest rate. Third, we see the loan's subsequent performance. Each month we see whether a borrower is current, delinquent on at least one payment, or has already defaulted on the loan.

In the event of default, a lender typically sells the loan to another investor or a collection specialist. Our data contains the size and timing of the cash payment to the lender in the event that a loan is charged off. This value is critical in computing the internal rate of return (IRR) on a loan, and most other papers cannot directly observe it.

To see why, consider a 3-year $12.0k personal loan with a 25% annual interest rate. First, suppose the borrower makes all 36 monthly payments of $477.12. In

this case, the lender's IRR would be equal to the annual interest rate

$$\$12{,}000 = \sum_{m=1}^{36} \frac{\$477.12}{(1 + \text{IRR}/12)^m} \qquad \Rightarrow \qquad \text{IRR} = 25\% \tag{1}$$

Now, think about a world where the borrower stops making payments after 10 months and the lender subsequently sells the loan for \$3.0k to a collection specialist in month 13. In that case, the lender would suffer a 57.6% loss

$$\$12{,}000 = \sum_{m=1}^{10} \frac{\$477.12}{(1 + \text{IRR}/12)^m} + \frac{\$3{,}000}{(1 + \text{IRR}/12)^{13}} \qquad \Rightarrow \qquad \text{IRR} = -57.6\% \tag{2}$$

Our data would show that the lender collected a \$3.0k recovery payment in month 13. By contrast, other papers like De Roure, Pelizzon, and Thakor (2022) are only able to estimate returns under different recovery rates, which is problematic because small changes in the recovery payment can have big effects on the implied return. For example, if the recovery payment had been just \$500 smaller in the above example, \$2.5k vs \$3.0k, the lender's IRR would have dropped to −70.0%. Di Maggio, Ratnadiwakara, and Carmichael (2021) are able to calculate realized returns but only see the first 12 months of payments. They would miss the \$3.0k recovery payment in month 13.

We apply five different kinds of filters to ensure that our empirical analysis will not be driven by data artifacts. Our results are largely insensitive to these restrictions, but we choose to be conservative to alleviate any concerns arising from including these observations. After imposing these five filters, we are left with a final sample of 2.3m FinTech personal loans.

Filter #1: Loan Amount.

While we observe the size of each loan, the data do not specify whether this value is gross or net of fees. To alleviate this confusion, we restrict our sample to loans for round multiples of \$500 or \$1,000 since these are unlikely to include fees. This filter reduces the sample by 1.7m loans. We also exclude an additional 100k loans where the reported monthly payments do not match

the values implied by the stated annual interest rate.

Filter #2: Maturity.

Most personal loans fully amortize over 36 months. We remove 1.6m loans that have some other maturity. By focusing on loans with a single term (36 months), we avoid comparability issues by comparing cash flows with different terms. For example, loans with the same origination date and different maturities may be exposed to different economic shocks.

Filter #3: FICO Score.

Sometimes a borrower's FICO score will be given in 5-point bins. When this happens, the reported value ends in either a 2 or a 7. For example, a lender would submit FICO $= 652$ for a loan with FICO $\in [650, 655)$ and FICO $= 657$ for a loan with FICO $\in [655, 660)$. We include these observations.

However, our data also contains 765k observations where the FICO score has been rounded to the nearest 20-point bin. When this happens, the reported value ends in a 9. For example, all loans with FICO $\in [640, 659)$ would have a value of FICO $= 649$. We exclude these loans because, while it is unlikely that a ± 2 point error in a borrower's FICO score would matter much, it is plausible that being off by ± 10 could matter.

Filter #4: Missing Data.

We drop observations where any of the following variables have missing values: the loan amount, its required monthly payment, the borrower's current/delinquent/default status, the state of origination, the date of origination.

Filter #5: Sample Period.

We drop any loans originated after July 2019 to give us a performance window of at least 12 months.

Summary Statistics

		#Obs	Avg	Med	SD	Min	Max
FICO		2,313,105	684	677	44	600	850
InterestRate	(%)	2,313,105	16.3	13.6	8.5	0.9	52.4
LoanAmount	($)	2,313,105	11,898	10,000	8,386	500	44,500
Payment	($/m)	2,313,105	410	334	274	44	1,235
Income	($/yr)	2,313,105	73,446	62,500	43,594	16,812	250,000
$\frac{12\times\text{Payment}}{\text{Income}}$	(%)	2,313,105	0.07	0.07	0.04	0.01	0.21
CreditHistLength	(yr)	2,313,105	16.2	14.6	7.7	3.4	40.3
#Inquiries		2,313,105	0.7	0.0	1.0	0.0	5.0
CreditUtilization	(%)	2,313,105	53.2	54.7	25.0	1.3	99.3
PurposeIsDebtConsl	(%)	2,313,105	76.1				
PurposeIsHomeImp	(%)	2,313,105	7.3				
IsHomeowner	(%)	2,313,105	55.7				
WillDefault	(%)	2,313,105	14.4				
WillBeDelinq	(%)	2,313,105	8.12				
WillPrePay	(%)	2,313,105	45.9				
IRR	(%)	1,138,365	6.2	12.1	29.7	-100.0	41.8

Table 1. *This table presents summary statistics for our sample of FinTech personal loans. The data come from a loan aggregator specializing in FinTech lending. The sample is restricted to* 36*-month loans. FICO: borrower's credit score. InterestRate (%): annual interest rate. LoanAmount ($): size of the loan. Payment ($/m): monthly payment amount in dollars. Income ($/yr): borrower's annual income.* $\frac{12\times Payment}{Income}$ *(%): size of loan payments as a percent of a borrower's annual income. CreditHistLength (yr): age of the oldest item in a borrower's credit report. #Inquiries: number of credit inquiries the borrower made as reported by the credit bureau in the six months before origination. CreditUtilRate (%): percent of the borrower's available revolving credit utilized at the origination time. PurposeIsDebtConsl (%): stated purpose of loan is debt consolidation. PurposeIsHomeImp (%): stated purpose of loan is home improvement. IsHomeowner (%): borrower owns their own house. WillDefault (%): loan is charged off within its duration. WillBeDelinq (%): loan will have a delinquent payment within one year of origination. WillPrePay (%): loan will be paid in full before reaching 36-month term. IRR (%): internal rate of return using all cash flows at the loan level, excluding any financial or operating costs on the lender side.*

Table 1 provides loan terms, borrower characteristics, and performance statistics. The average annual interest rate and loan size are 16.3% and $11,898, respectively. Borrowers have a median income of $62,500, which is close to the national median household income in 2018 of $63,179 (US Census, 2019). The average credit utilization rate in our sample, 53.2%, is well above the national average. This relatively is to be expected given that 76.1% of FinTech personal loans are used to consolidate debt.

2 Institutional Background

FinTech lenders are much more willing to lend to nonprime borrowers. For all intents and purposes, a borrower with FICO = 660 is just as risky as a borrower with a FICO = 659. However, because the first borrower is above the prime/nonprime threshold, she will have a much easier time getting a personal loan from traditional banks (Chen, Hanson, and Stein, 2017; Buchak, Matvos, Piskorski, and Seru, 2018). This section provides direct evidence of this difference in lending standards (Subsection 2.1) and discusses some of the key reasons why it exists (Subsection 2.2).

2.1 Direct Evidence

We hand-collect the minimum FICO requirements disclosed by banks and FinTech lenders. Banks often post credit requirements on their websites. In cases where they do not, we turn to personal finance websites like WalletHub, NerdWallet, and Bankrate. These platforms often list a bank's menu of loan products and the minimum requirements for each offering. For example, see `https://wallethub.com/best-banks-for-personal-loans` or `https://www.nerdwallet.com/best/loans/personal-loans/best-personal-loans`.

We look at the 15 largest banks in the US by asset size. We restrict our list to those offering unsecured personal loans. Among the 9 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 1 (PNC Bank) of these 9 had no qualifications listed anywhere. We report the other 8 in Panel (a) of Table 2. 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.

Panel (b) in Table 2 shows the results of a similar exercise for FinTech lenders. The US FinTech Market Report produced by S&P Global Market Intelligence in February 2021 labels Avant, Best Egg, LendingClub, LendingPoint, Prosper, SoFi, and Upstart as the most important FinTech lenders. There is a stark difference in minimum-lending requirements in Panels (a) and (b). 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 (e.g., Di Maggio, Ratnadiwakara, and Carmichael, 2021), limits competition and gives FinTech lenders market power. 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.

2.2 Underlying Causes

There are several reasons why traditional banks have been unwilling to lend to nonprime borrowers.

Bank Regulations. 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, Matvos, Piskorski, and Seru (2018) presents 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

Minimum FICO Requirements For Different Lenders

Panel (a): Traditional Banks

580	600	620	640	660	680	700
				Wells Fargo Am Express* SunTrust LightStream Discover* Goldman Sachs Marcus	Citibank US Bank Fifth Third	TD Bank HSBC Barclays*

Panel (b): Online Lenders

580	600	620	640	660	680	700
Avant	Best Egg LendingClub LendingPoint Upstart	Freedom Plus* Upgrade*	Payoff* Rocket Loans* Prosper		SoFi	

Table 2. *This table identifies the minimum FICO scores for the* 15 *largest banks (Panel a) and the* 7 *largest online lenders (Panel b) when underwriting unsecured personal loans.* 6 *of the largest* 15 *banks do not offer personal unsecured loans: Chase Bank, Bank of America, Capital One, Bank of New York Mellon, and State Street Bank. Institutions with asterisks in each panel denote other lenders with a significant presence in the credit market. These data come from 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.*

the restructuring of the financial regulatory framework have led to a significant drop in bank lending.

Banks must follow clear and strict rules when originating nonprime loans. For example, here is how the FDIC Risk Management Manual for Examination Policies describes the rules for nonprime lending:

> "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." https://www.fdic.gov/risk-management-manual-examination-policies.

The manual stresses that bank examiners should closely monitor subprime loans, which are characterized, among other factors, by a FICO score of 660, due to their higher probability of delinquency.

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 directly constrained by banking regulations, they can still be affected by them in two ways. First, there is the wholesale funding channel. Shadow banks and FinTech lenders often get capital from banks and other investors who face higher regulatory costs. For example, several FinTech lending platforms, like Prosper, also make traditional direct-to-consumer loans through WebBank. This FDIC-insured, state-chartered industrial bank has to follow the same rules.

Second, FinTech lenders compete with banks to provide unsecured credit. 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. 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 (2023) uses mortgage data to demonstrate the two channels.

Originate To Distribute (OTD). FinTech lenders are likely more reliant on the originate-to-distribute (OTD) business model. They aim to maximize loan origination volume and related fees while complying with the Securities and Exchange Commission (SEC) disclosure guidelines. Lenders achieve this by securitizing the loans they originate and selling them to institutional investors. These investors provide the capital that FinTech lenders use for new loan originations.

However, this fact should make them less likely to lend to nonprime borrowers. 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, 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., $620 \leq \text{FICO} < 640$, $640 \leq \text{FICO} < 660$, $660 \leq \text{FICO} < 680$) and loan sizes (e.g., $\$1{,}500 \leq \text{LoanAmount} < \$2{,}000$).

The success of a securitized product depends on meeting these criteria. As a result, a borrower's FICO score may receive more attention relative to other predictors of default. Taken to the extreme, a lender may simply refuse to lend to any borrower who falls outside of certain prespecified guidelines. It is noteworthy that FICO scores in our database are often only reported to the

level of precision required by securitized products, the nearest 5- or 20-point bin. The market forces "legibility" on borrowers (Scott, 2020).

State Usury Laws. An additional regulatory constraint that all lenders must abide by is usury laws that limit the interest rate charged on loans. The existence of these laws means that, for the riskiest nonprime borrowers, it is not possible to lend at a profitable interest rate.

While these laws are determined at the state level, most of the time financial institutions can bypass these state-level requirements by incorporating their lending arm in states with high usury limits. As a result, the effective usury limit is 36% in almost all states. The only exceptions to this rule are Colorado, Iowa, New York, West Virginia, and Vermont 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.

Usury laws do not 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. Federal laws such as the Equal Credit Opportunity Act and the Fair Housing Act prohibit lenders from incorporating information correlating with race, religion, national origin, sex, marital status, age, source of income, etc. Such regulations effectively limit the use of many variables that predict default in ways that happen to be correlated with protected classes. Previous research has documented important effects in mortgage lending (Fuster, Goldsmith-Pinkham, Ramadorai, and Walther, 2022).

Again, these Federal restrictions should make it harder for FinTech lenders to use big data and machine learning to give underserved groups better access to financing. It is also worth remembering that many FinTech lenders have no local branches. As a result, they may lack the place-based knowledge needed to institute an effective regional pricing scheme.

Funding 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 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.

3 Almost-Prime Loans

In this section, we study loans close to the prime/nonprime threshold at FICO = 660. Within a narrow range, an almost-prime borrower with a FICO score just below 660 will not be much riskier than a barely-prime borrower with a FICO score just above the cutoff. Nevertheless, in Subsection 3.1, we find that FinTech lenders charge almost-prime borrowers much higher annual interest rates. In Subsection 3.2, we then show that FinTech lenders could use the additional non-FICO data they collect about each borrower to identify the safest of the almost-prime loans. Yet, there is no evidence that FinTech lenders use this information for pricing purposes. They originate all almost-prime loans at more or less the same 25% annual interest rate.

3.1 vs Barely-Prime Loans

We start by comparing the annual interest rates and first-year delinquency rates of FinTech personal loans on either side of the prime/nonprime threshold. We want to emphasize that our main results do not involve comparing prime and nonprime loans. We are primarily interested in whether FinTech lenders are pricing predictable differences in risk within the group of nonprime borrowers. We begin with this comparison because it is the standard regression-discontinuity design (RDD) researchers are used to seeing. The results in this first subsection also provide a helpful point of comparison when trying to interpret the economic magnitudes of our effect sizes.

The first thing we look at are some simple averages. Let ℓ denote an arbitrary FinTech personal loan. In the regressions that follow, we will use Outcome_ℓ as a generic stand-in variable

$$\text{Outcome}_\ell \in \{\text{InterestRate}_\ell, 100 \times \text{WillBeDelinq}_\ell\} \tag{3}$$

that denotes either a loan's interest rate or a 0/100 indicator variable for whether the borrower missed one of the first twelve payments.

The top two rows of Figure 1 show that, even though almost-prime borrowers are not much riskier, they pay a much higher interest rate than the typical barely-prime borrower and, as a result, are more likely to miss a payment during the first twelve months. Table 3 replicates this finding while controlling for observables. Columns (1) and (4) look at borrowers within ±5pts of the prime/nonprime threshold using the specification below

$$\begin{aligned} \text{Outcome}_\ell \overset{\text{OLS}}{\sim} \hat{\alpha} \cdot 1\{\text{FICO}_\ell < 660\} \\ + \hat{\delta} \cdot \log(\text{LoanAmount}_\ell) \\ + \text{State}\times\text{Year FEs} \\ + \text{Origination Month FEs} \end{aligned} \tag{4}$$

Column (1) indicates that the average annual interest rate is $\hat{\alpha} = 9.43$%pts higher for FinTech personal loans with $655 \leq \text{FICO} < 660$ than for those

with $660 \leq \text{FICO} < 665$. The log(LoanAmount) control ensures that we are not comparing apples to oranges in terms of loan size. The state×year and origination month FEs ensure that the difference is not coming from comparing loans made in different geographic areas or times.

Column (4) reports that the average FinTech personal loan with $655 \leq \text{FICO} < 660$ is around $\hat{\alpha} = 4.88\%$pts more likely to miss one of the first twelve monthly payments when compared to a loan with $660 \leq \text{FICO} < 665$. The control variables in Equation (4) ensure that the difference in delinquency rates cannot be attributed to differences in loan size, geography, or time trends.

As discussed in the introduction, the $\hat{\alpha} = 9.43\%$pt estimate from Column (1) implies that an average almost-prime borrower would have to pay $50 more each month to borrow $12k. The $\hat{\alpha} = 4.88\%$pt difference in Column (4) tells us that this relatively small change in monthly payments substantially increases the borrowers' likelihood of missing a payment during the first twelve months. Only one in $1/8.0\% \approx 13$ barely-prime borrowers will go delinquent during the first year; whereas, one out of $1/(8.0\% + 4.9\%) \approx 8$ almost-prime borrowers will miss a payment right out of the gate. Small differences in a borrower's free cash flow each month can have a big impact on loan performance.

Almost-Prime vs Barely-Prime Loans

Dep Variable:	InterestRate			100 × WillBeDelinq		
FICO Range:	660 ± 5 [655, 665)	660 ± 20 [640, 680)	660 ± 40 [620, 700)	660 ± 5 [655, 665)	660 ± 20 [640, 680)	660 ± 40 [620, 700)
	(1)	(2)	(3)	(4)	(5)	(6)
1{FICO < 660}	9.43***	9.13***	8.57***	4.88***	4.53***	4.45***
	(0.12)	(0.12)	(0.12)	(0.22)	(0.23)	(0.19)
FICO Interaction	N	Y	Y	N	Y	Y
Controls + FEs	log(LoanAmount), State×Year FE, Origination Month FE					
#Obs	213,017	822,252	1,457,316	213,017	822,252	1,457,316
Adj R^2	64.1%	48.7%	55.3%	12.7%	1.9%	2.4%

Table 3. *This table estimates the difference in annual interest and 12-month delinquency for loans on either side of the prime/nonprime threshold at FICO* = 660*. Columns (1) and (4) look at loans within* ±5 *points of the prime/nonprime threshold. Columns (2) and (5) look at loans within* ±20 *points of the threshold. Columns (3) and (6) look at loans with* ±40 *points. InterestRate (%): annual interest rate on a loan. WillBeDelinq: indicator variable that is* 1 *for loans that miss a payment during the first year and* 0 *otherwise.* 1{*FICO* < 660}*: indicator variable that is* 1 *if a loan is nonprime and* 0 *otherwise. Each column reports the results of a separate regression. Columns (1) and (4) use the specification in Equation* (4)*. Columns (2), (3), (5), and (6) use the specification in Equation* (5)*, which includes* (*FICO* − 660) *and the interaction term* 1{*FICO* < 660} × (*FICO* − 660)*. All regressions include controls for* log(*LoanAmount*) *as well as both state-year and origination-month fixed effects. Numbers in parentheses are standard errors clustered by origination month.* ⋆*,* ⋆⋆*, and* ⋆⋆⋆ *denote statistical significance at the* 10%*,* 5%*, and* 1% *levels.*

The remaining columns in Table 3 replicate the same finding in larger windows around the prime/nonprime threshold. Columns (2) and (5) study loans where $640 \leq \text{FICO} < 680$. Columns (3) and (6) look at an even larger range, $620 \leq \text{FICO} < 700$. Because a 40pt difference in FICO scores represents a meaningful difference in risk, we use an interaction term to control for distance to the prime/nonprime threshold

$$\begin{aligned} \text{Outcome}_\ell \overset{\text{OLS}}{\sim} & \hat{\alpha} \cdot 1\{\text{FICO}_\ell < 660\} \\ & + \hat{\beta} \cdot (\text{FICO}_\ell - 660) \\ & + \hat{\gamma} \cdot [1\{\text{FICO}_\ell < 660\} \times (\text{FICO}_\ell - 660)] \\ & + \hat{\delta} \cdot \log(\text{LoanAmount}_\ell) \\ & + \text{State}\times\text{Year FEs} \\ & + \text{Origination Month FEs} \end{aligned} \tag{5}$$

Our estimates for $\hat{\alpha}$ hardly change as we extend the window.

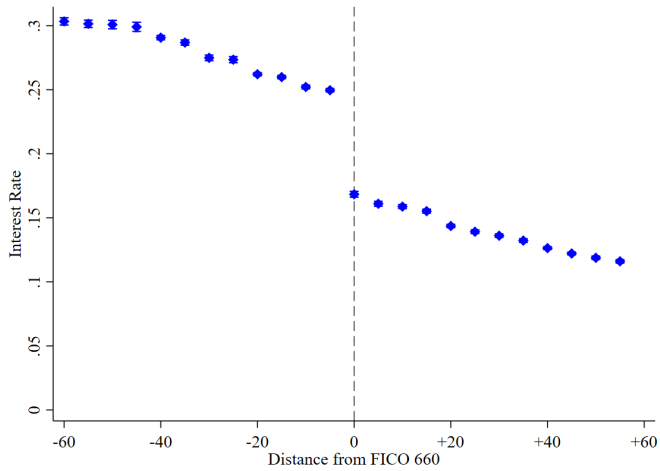
Figure 2 extends this finding across an even-wider ±60pt neighborhood of FICO = 660. The figure shows averages for 24 different 5pt FICO bins

$$\text{Bins} \overset{\text{def}}{=} \Big\{ [600, 605), [605, 610), \ldots, [655, 660), [660, 665), \ldots, [715, 720) \Big\} \tag{6}$$

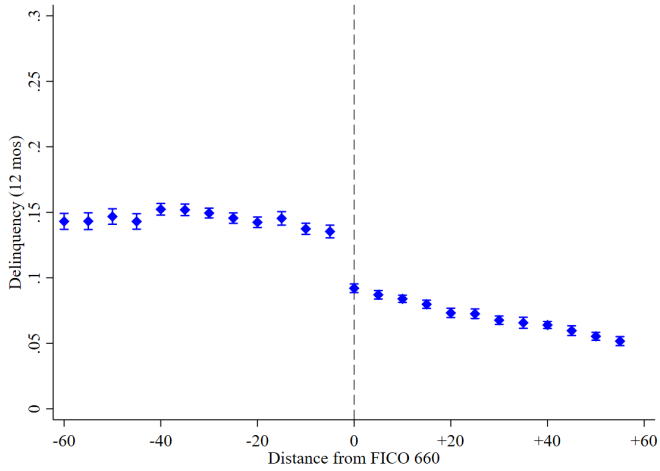
The blue dots are estimated $\hat{\mu}_b$ values from the following regression specification

$$\begin{aligned} \text{Outcome}_\ell \overset{\text{OLS}}{\sim} & \sum_{b \in \text{Bins}} \hat{\mu}_b \cdot 1[\,\text{FICO}_\ell \in b\,] \\ & + \text{State}\times\text{Year FEs} \\ & + \text{Origination Month FEs} \end{aligned} \tag{7}$$

The vertical blue crossbars denote 95% confidence intervals for each estimated value of $\hat{\mu}_b$ using standard errors clustered by origination month. Both panels in Figure 2 display a sharp change between the almost-prime bin, $[655, 660)$, and the barely-prime bin, $[660, 665)$.



(a) Interest Rate



(b) Delinquency Rate

Figure 2. *This figure shows averages for FinTech personal loans within* ±60*pts of the prime/nonprime threshold in* 5*pt FICO-score bins. Each panel represents the results of estimating the specification in Equation* (7) *using a different outcome variable. Panel (a) shows results for the annual interest rate. Panel (b) characterizes how the twelve-month delinquency rate changes with FICO score. In addition to FICO bins, the regression behind each panel includes state×year and origination-month FEs. All regressions are based on all loans in the sample. Standard errors are clustered by origination month.*

3.2 Within-Group Variation

This paper is not mainly about comparing the pricing and performance of almost-prime borrowers to that of barely-prime borrowers. Our key observation has to do with the within-group variation (or lack thereof) among nonprime borrowers. Column (4) in Table 4 shows that FinTech lenders could be using additional non-FICO variables, such as the number of credit inquiries and the size of the loan, to identify the safest almost-prime loans.

Almost-prime borrowers with large loans (above 27%ile; roughly $15k) are 15.9% − 9.3% = 6.6%pts more likely to go delinquent in the first year than those with small loans (below 25%ile; roughly $3k). Among almost-prime borrowers, homeowners are 12.4% − 11.1% = 1.3%pts less likely to go delinquent in the first twelve months than renters. Almost-prime borrowers with two or more credit checks in the past 6 months are 13.4% − 10.9% = 2.5%pts more likely to miss one of the first twelve payments compared to an almost-prime borrower with no recent credit inquiries.

It makes sense why it might be helpful to collect additional data about a borrower's financial situation. At a 25% annual interest rate, the monthly payment on a 3-year $3k personal loan would be $119.28. To borrow $15k at the same terms, a borrower would have to pay $596.40 per month. It does not take an advanced machine-learning algorithm to figure out that an almost-prime borrower might be more likely to miss a monthly payment when required to pay $477.12 more each month.

This is just one of many non-FICO variables that FinTech lenders collect. They could be offering better loan terms to almost-prime homeowners with no recent credit inquiries who are applying for a $3k personal loan. We find very little evidence that they such better terms. Column (3) of Table 4 shows all almost-prime FinTech personal loans have the same 25% rate, regardless of other predictors. The largest difference in interest rates is 1.1%pts, and none of them is statistically distinguishable from zero. Borrowers have to provide lots of personal data to FinTech lenders when they apply for a personal loan. It does not seem like FinTech lenders are using any of it for pricing purposes.

Within-Group Variation

	#Obs	Sorting Variable	Interest Rate	Delinq Prob
	(1)	(2)	(3)	(4)
1{Income ≤ 25%ile}	22,122	$28,007	25.3%	11.7%
1{Income > 75%ile}	22,108	$80,180	24.9% [+0.4%]	12.0% [−0.3%]
PurposeIsMedicalBill	1,318		25.4%	11.3%
PurposeIsDebtConsl	61,148		24.9% [+0.5%]	12.1% [−0.8%]
1{#Inquiries ≥ 2}	20,625	2.8	25.7%	13.4%
1{#Inquiries = 0}	42,403	0.0	24.6% [+1.1%]	10.9% [+2.5%]
IsRenter	44,887		25.2%	12.4%
IsHomeowner	36,922		24.8% [+0.4%]	11.1% [+1.3%]
1{LoanAmount > 75%ile}	22,219	$15,184	25.2%	15.9%
1{LoanAmount ≤ 25%ile}	22,618	$2,863	25.7% [−0.5%]	9.3% [+6.6%]

Table 4. *This table show the differences in the average interest rate and twelve-month delinquency rate for different subsets of almost-prime borrowers, FICO* $\in \{655, 656, 657, 658, 659\}$*. #Obs: number of loans in the subgroup. Sorting Variable: average value of the sorting variable for loans in a given subgroup. Interest Rate (%): average annual interest rate for loans in a given subgroup. Delinq Prob (%): percent of loans in a subgroup that go delinquent within the first twelve months. Numbers in square brackets denote the difference between the average in each subgroup.* 1{*Income* ≤ 25%*ile*}*: borrowers below the 25th percentile of the almost-prime income distribution (low income).* 1{*Income* > 75%*ile*}*: borrowers above the 75th percentile of the almost-prime income distribution (high income). PurposeIsMedicalBill: almost-prime loans where the stated purpose is paying a medical bill. PurposeIsDebtConsl: almost-prime loans where the stated purpose is debt consolidation.* 1{*#Inquiries* ≥ 2}*: almost-prime loans where the borrower has had two or more credit checks during the past* 6 *months.* 1{*#Inquiries* = 0}*: almost-prime loans where the borrower has had no credit checks during this time. IsRenter: almost-prime loans taken out by a borrower who does not own a home. IsHomeowner: almost-prime loans taken out by a borrower who does.* 1{*LoanAmount* > 75%*ile*}*: loans above the 75th percentile of the almost-prime loan-size distribution (large loans).* 1{*LoanAmount* ≤ 25%*ile*}*: loans below the 25th percentile of the almost-prime loan-size distribution (small loans).*

4 Endogenous Selection

Our primary data source only includes FinTech personal loans that were originated. We do not observe loan offers that FinTech lenders made but were subsequently turned down by the borrower. So perhaps FinTech lenders are actually offering bespoke loan terms to each almost-prime borrower? Perhaps it is just that the borrowers who accept the offers all happen to get quoted a 25% annual interest rate? This section introduces a simple model of endogenous selection within a potential outcomes framework. In Subsection 4.1, we first highlight what this selection story can explain. Then, in Subsection 4.2, we show that it cannot explain key patterns in our data.

4.1 What It Can Explain

The first two rows of Table 4 offer a pattern that an endogenous selection story could potentially explain. It would be reasonable to think that it is safer to lend to an almost-prime borrower making $80k a year (HasHighIncome = 1) than to an almost-prime borrower with an annual income of just $28k (HasHighIncome = 0). We encode this idea in our model as

$$\Pr[\text{WillBeDelinq}] = \kappa - \lambda \cdot \text{HasHighIncome} + \mu \cdot \text{IsUnobsrvblyRisky} \tag{8a}$$

$$\text{InterestRate} = r_f + \nu \cdot \mathbb{E}[\text{WillBeDelinq}] \tag{8b}$$

$r_f, \kappa \geq 0$ are non-negative, $\lambda, \mu, \nu > 0$ are strictly positive, and IsUnobsrvblyRisky $\in \{0, 1\}$ indicates that a borrower is less creditworthy than they appear on paper.

The first two entries in column (4) of Table 4 seem to violate this economic intuition. They show that the average interest and delinquency rate is roughly the same on the personal loans that FinTech lenders make to each group. The average annual interest rate is 25% for both sets of personal loans, and roughly one in $1/12\% \approx 8$ borrowers will fail to make a payment during the first year in both groups. However, this could be because the FinTech personal loans in our data have a non-random assignment of IsUnobsrvblyRisky.

Suppose that, for high-income almost-prime borrowers, a FinTech personal loan is the option of last resort. Borrowers in this group only accept a FinTech lender's offer (AcceptsOffer = 1) if no one else will lend to them because they are riskier on unobservable dimensions (IsUnobsrvblyRisky = 1)

$$\Pr[\,\text{AcceptsOffer} \,|\, \text{HasHighIncome} = 1\,] = \begin{cases} 1 & \text{if IsUnobsrvblyRisky} = 1 \\ 0 & \text{if IsUnobsrvblyRisky} = 0 \end{cases} \tag{9}$$

By contrast, the exact opposite is true for low-income almost-prime borrowers

$$\Pr[\,\text{AcceptsOffer} \,|\, \text{HasHighIncome} = 0\,] = \begin{cases} 0 & \text{if IsUnobsrvblyRisky} = 1 \\ 1 & \text{if IsUnobsrvblyRisky} = 0 \end{cases} \tag{10}$$

In order to make it into this calculation (IsInOurData = 1), the FinTech lender's offer must be accepted (AcceptsOffer = 1). Hence, if we use $\hat{\mathbb{E}}[\cdot] \overset{\text{def}}{=} \mathbb{E}[\,\cdot \,|\, \text{IsInOurData} = 1\,]$ to denote the sample average computed using the observed data, then we can write

$$\hat{\mathbb{E}}[\,\text{WillBeDelinq} \,|\, \text{HasHighIncome} = 1\,] = \kappa - \lambda + \mu \tag{11a}$$

$$\hat{\mathbb{E}}[\,\text{WillBeDelinq} \,|\, \text{HasHighIncome} = 0\,] = \kappa \tag{11b}$$

We can write the difference between Equations (11a) and (11b) as

$$\begin{aligned} \mu - \lambda = \hat{\mathbb{E}}[\,&\text{WillBeDelinq} \,|\, \text{HasHighIncome} = 1\,] \\ &- \hat{\mathbb{E}}[\,\text{WillBeDelinq} \,|\, \text{HasHighIncome} = 0\,] \end{aligned} \tag{12}$$

Notice that there are two effects to consider. On one hand, if it is safer to lend to high-income almost-prime borrowers, then it would lower the sample average delinquency rate for this subpopulation ($-\lambda$ in Equation 12). But, on the other hand, the observed high-income almost-prime borrowers in this example are the worst of the worst (μ in Equation 12). The identical delinquency rates in the first two rows of Table 4 column (4) could stem from these two effects precisely canceling out, $\mu = \lambda$.

4.2 What It Cannot Explain

However, notice that it is not possible to explain any arbitrary pattern in the data with this selection story. In column (3) of Table 4, we see that the average annual interest rate is 25% for FinTech personal loans made to both high- and low-income almost-prime borrowers. By plugging Equation (12) into the rule for determining a loan's interest rate in Equation (8b), we see that this pattern is perfectly consistent with a selection story

$$\begin{aligned} \nu \cdot (\mu - \lambda) = \hat{\mathbb{E}}[\,\text{InterestRate} \,|\, \text{HasHighIncome} = 1\,] \\ - \hat{\mathbb{E}}[\,\text{InterestRate} \,|\, \text{HasHighIncome} = 0\,] \end{aligned} \tag{13}$$

However, the final two rows in Table 4 show something very different. In column (3), we see that the average annual interest rate is 25% for both large and small almost-prime FinTech loans. But, unlike the previous case where we split by borrower income, the final two values in column (4) indicates that a \$3k almost-prime loan (IsSmallLoan = 1) is 6.6%pts less likely to miss a payment than a \$15k almost-prime loan (IsSmallLoan = 0).

This cannot happen in a model of endogenous selection. If the two groups have identical average annual interest rates, they cannot have different average delinquency rates. Endogenous selection can distort the apparent riskiness of originated almost-prime loans. But conditional on a loan being originated, it should still be priced correctly. If a selection story is responsible for

$$\begin{aligned} 0 = \nu \cdot (\mu - \lambda) = \hat{\mathbb{E}}[\,\text{InterestRate} \,|\, \text{IsSmallLoan} = 1\,] \\ - \hat{\mathbb{E}}[\,\text{InterestRate} \,|\, \text{IsSmallLoan} = 0\,] \end{aligned} \tag{14}$$

in the last two entries of column (3) in Table 4, then we cannot observe

$$\begin{aligned} 0 \neq \mu - \lambda = \hat{\mathbb{E}}[\,\text{WillBeDelinq} \,|\, \text{IsSmallLoan} = 1\,] \\ - \hat{\mathbb{E}}[\,\text{WillBeDelinq} \,|\, \text{IsSmallLoan} = 0\,] \end{aligned} \tag{15}$$

when comparing the last two values in column (4). This toy model is obviously a caricature of reality, but the basic logic holds for any $\nu > 0$.

5 All Nonprime Loans

This section extends our earlier finding to the entire population of nonprime FinTech personal loans. In Subsection 5.1, we build a simple delinquency-risk model to account for the wide variation in riskiness across nonprime FinTech personal loans. Our model incorporates non-FICO information like income, loan amount, and homeownership and yields accurate unbiased predictions for the likelihood of first twelve-month delinquencies. In other words, we do what FinTech lenders say they are doing. We find that twelve-month delinquency rates are 3.5× higher for the riskiest nonprime borrowers, 7% vs 24%. However, they pay just 4%pts most in annual interest, 25% vs 29% per year. In Subsection 5.3, we then show that all this unpriced predictable variation in delinquency rates can generate a significant difference in realized returns.

5.1 Predicting Delinquency

FinTech lenders claim to be using big data and machine learning to identify and offer bespoke loan terms to the safest nonprime borrowers. We now perform a naïve version of the exercise. Each month m, we fit the logit specification below to data on FinTech personal loans originated during the 24 month window, $\{(m-36), \ldots, (m-13)\}$

$$\text{WillBeDelinq}_\ell \overset{\text{Logit}}{\sim} \hat{\boldsymbol{\theta}}^\top \textbf{Characteristics}_\ell \tag{16}$$

The vector $\textbf{Characteristics}_\ell$ contains a flexible array of loan and borrower characteristics for the ℓth FinTech personal loan: 20pt FICO bins, \$5k income bins, \$1k loan-amount bins, indicators for the loan purpose, the length of the borrower's credit history, the percent of the borrower's available revolving credit utilized at the time of origination, and the number of credit inquiries during the last six months. Then, we use the fitted values to predict the likelihood of delinquency for each loan originated in month m. This is where the predicted delinquency rate comes from in the final row of Figure 1.

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 in information-scarce environments. The average AUC in our regressions is close to that benchmark at 66.5%, which is comparable to other unsecured debt models in earlier studies (e.g., Berg et al., 2020,?).

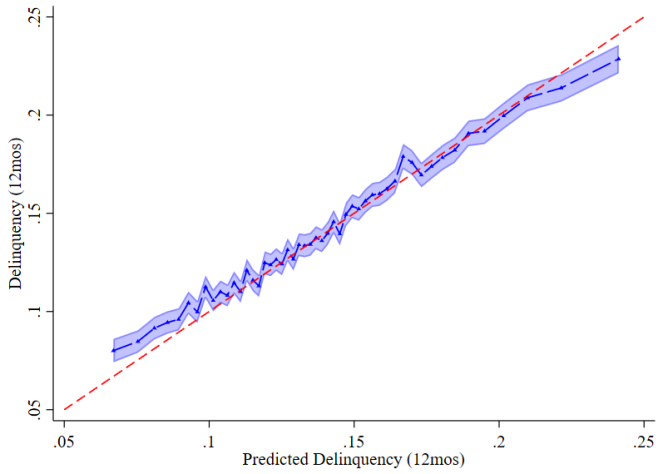
The second approach is to use a binscatter plot to show the average realized delinquency for nonprime borrowers against the expected delinquency, calculated using pre-origination data, $\hat{\mathbb{E}}[\text{WillBeDelinq}_\ell]$. If the prediction model performs well, we expect the slope to line up close to the 45-degree line. Indeed, Figure 3(a), panel (a), shows that the average realized delinquency aligns well with predicted delinquency. In Figure 3(b), Panel (b), we replace the average realized delinquency with the average default. Even though our prediction model was estimated with an ex-post delinquency indicator, it also captures default risk well.

Next, we assess the degree to which the pricing of nonprime loans varies with our measure of predicted delinquency risk. The exercise is relatively straightforward. Figure 4 shows a binscatter plot of the average interest rates FinTech lenders charge on loans, per their predicted delinquency rate.

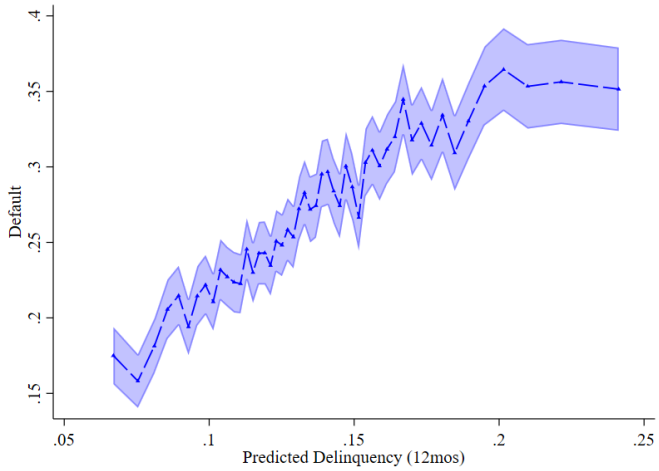
The average interest rate charged to nonprime borrowers does not meaningfully vary with their predicted delinquency. For example, a borrower with a 10% delinquency likelihood pays an average rate of 26%, only slightly lower than the 28% paid by a borrower with a 20% delinquency likelihood. This minimal difference suggests that the pricing of loans does not adequately reflect the differences in predicted credit risk among nonprime borrowers.

5.2 FICO-based Pricing

Our results basically say that the FICO score is primarily relevant to pricing and that FinTech lenders do not use the information available to them. We can



(a) Delinquency



(b) Default

Figure 3. *This figure compares predicted loan performance relative to the actual performance for all nonprime FinTech personal loans. We predict delinquency for nonprime loans originated in month m by fitting a logit regression to the observed first-year delinquency outcomes of all nonprime loans originated in months* $\{(m-36),\ldots,(m-13)\}$. *Panel (a) shows a binscatter plot of the predicted likelihood of missing one of the first twelve payments vs the observed delinquency rate. 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.*

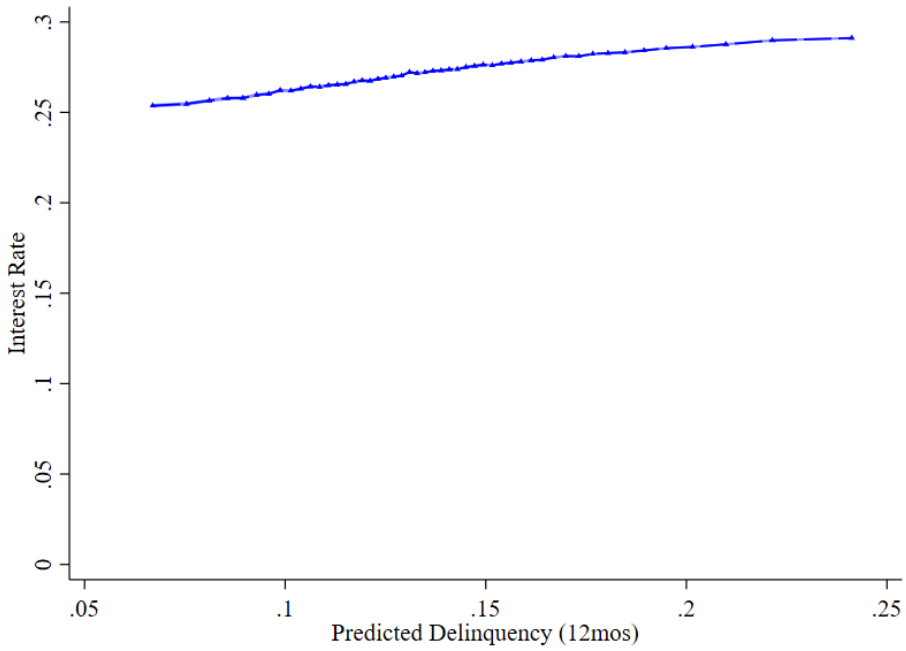


Figure 4. *This figure shows the relationship between a nonprime FinTech personal loan's predicted delinquency rate and its annual interest rate. We predict delinquency for nonprime loans originated in month m by fitting a logit regression to the observed first-year delinquency outcomes of all nonprime loans originated in months* $\{(m-36),\ldots,(m-13)\}$.

visualize this result for the entire market by showing how loans are priced with respect to FICO and expected delinquency.

Figure 5 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. The colors, reflecting the average interest rate, are best described as vertical stripes. In other words, they vary with FICO scores, and are almost invariant with respect to delinquency probability. Interestingly, the little variation in loan rates is much more pronounced in the prime market than in the nonprime.

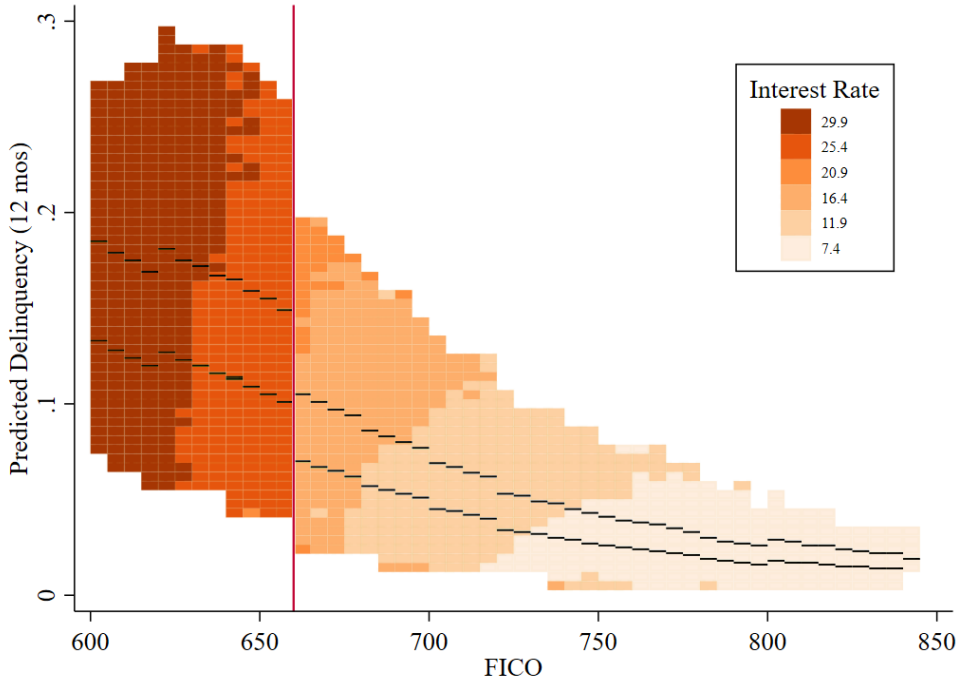


Figure 5. *This figure shows a heat map of the average interest rate (darker red = higher interest rate) based on all FinTech personal loans. Loans are sorted by FICO score (x-axis; binned at* 5*pt intervals) and by predicted delinquency (y-axis; binned at* 0.5*pt intervals; minimum of 100 loans per bin). We predict delinquency for nonprime loans originated in month m by fitting a logit regression to the observed first-year delinquency outcomes of all nonprime loans originated in months* $\{(m-36),\ldots,(m-13)\}$*. For each FICO bin, the black lines indicate the* 25*th and* 75*th percentile breakpoints.*

5.3 Realized Returns

A fundamental principle in finance states that taking on higher risk is rewarded with higher returns in financial markets. However, capital providers in the nonprime unsecured personal loan market do not experience this. The fact that nonprime loan prices are insensitive to expected delinquency risk implies that realized returns will predictably vary: low-risk loans earn high returns, and vice versa.

To see this result, we plot the average IRR against expected risk in Figure 6.

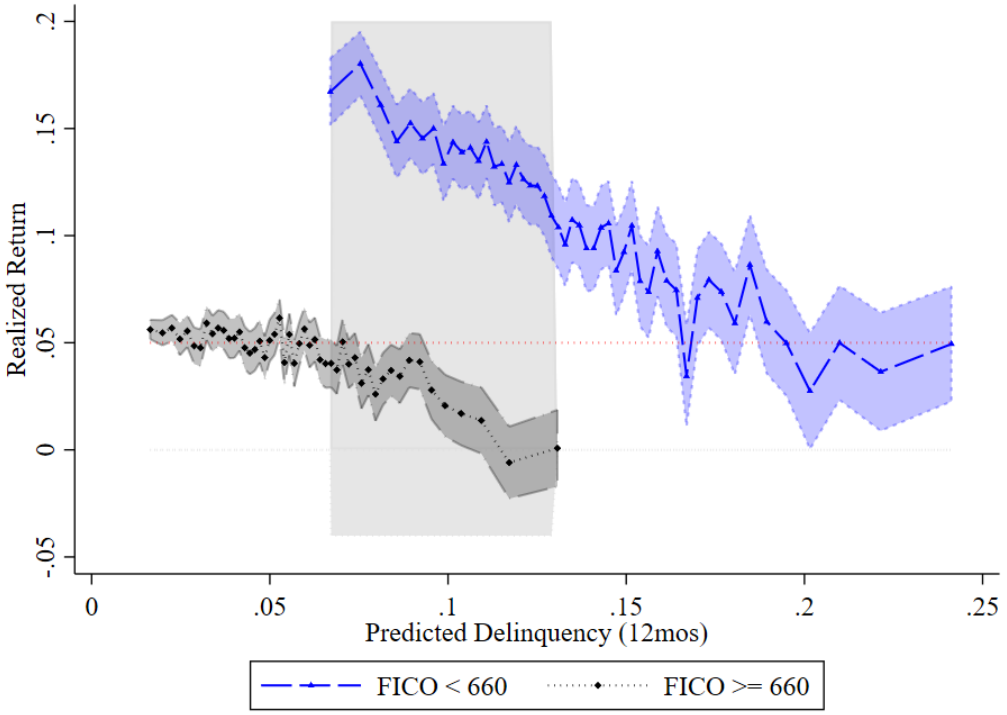


Figure 6. *This figure shows the relationship between realized loan-level IRR and predicted risk separately for all prime and nonprime loans in the sample. Plotted is the average IRR by predicted delinquency bins, along with standard errors. IRR is calculated using the initial outlay (loan amount) and subsequent cash flows (loan repayments) observed in the panel data. These cash flows also include loan recoveries when the loan is charged off.*

IRR is computed at the loan level, as in Equation (1), annualized. The plot shows that low-expected risk nonprime loans deliver an annualized IRR of about 11% on average. Defying the fundamental principle of finance, our results show that the realized return declines as the loans' expected risk increases. The loans in the lowest decile of expected delinquency earn over 15% return. In contrast, those in the ex-ante riskiest decile earn less than 5% a year.

A lender who invests its own capital likely would have invested in sophisticated algorithms (or even not-so-sophisticated as we did) to avoid the riskiest loans. However, the incentives of some FinTech firms may differ, given that their business model is based on fees earned from loan securitization.

Conclusion

The benefits of technological advances in financial markets materialize in different forms. To date, academics and regulators have been told that FinTech firms bring value to consumers by using sophisticated pricing algorithms that allow risk-based loan pricing. This would have been a true innovation for nonprime borrowers, who were primarily barred from accessing credit from traditional lenders like banks. Lenders would relax the financial constraints of a large population by offering risk-based pricing, as opposed to FICO-based term sheets.

Using a dataset containing most of the unsecured personal loans made in the U.S. up to 2020, we report that this is not the case. While FinTech lenders make loans to nonprime borrowers, these loans are at rates in the mid-20 percentage points, and they vary little with expected default risk. We show that around the prime cutoff (FICO 660) nonprime borrowers pay rates higher by 9% despite their default rate is higher only by 4.5%. Within the nonprime borrower population, rates are flat with respect to easy-to-calculate delinquency risk measures.

The absence of risk-based beyond FICO score means that capital providers invest in pools of loans of different qualities. FinTech lenders could have improved the risk-return relationship in these portfolios by using a simple risk-based pricing algorithm. They may have chosen not to do it because the originate-to-distribute model typically rewards origination volume rather than performance and quality.

Although risk-based pricing is not the primary way FinTech lenders provide value, they excel in making credit more accessible through online and mobile platforms, simplifying the application process, and expediting access to funds. Our findings suggest that the core disruptive technology of FinTech lenders is not advanced statistical modeling but the convenience of mobile apps that streamline the borrowing experience. While this approach is a valid market innovation, it diverges from the industry's promise of democratizing credit through superior risk assessment.

These findings raise important policy considerations since the credit constraints that FinTech intended to relax are still there. Regulators and researchers should, therefore, critically assess whether FinTech's advertised innovation aligns with its actual practices.

References

Agarwal, S., T. Kigabo, C. Minoiu, A. F. Presbitero, and A. F. Silva (2023). Serving the underserved: microcredit as a pathway to commercial banks. *Review of Economics and Statistics 105*(4), 780–797.

Berg, T., V. Burg, A. Gombović, and M. Puri (2020). On the rise of fintechs: Credit scoring using digital footprints. *Review of Financial Studies 33*(7), 2845–2897.

Berg, T., M. Puri, and J. Rocholl (2020). Loan officer incentives, internal rating models, and default rates. *Review of Finance 24*(3), 529–578.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics 130*(3), 453–483.

Chen, B. S., S. G. Hanson, and J. 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, C., L. Pelizzon, and A. Thakor (2022). P2P lenders versus banks: Cream skimming or bottom fishing? *Review of Corporate Finance Studies 11*(2), 213–262.

Di Maggio, M., D. Ratnadiwakara, and D. Carmichael (2021). Invisible primes: Fintech lending with alternative data. Working paper, Harvard Business School.

Di Maggio, M. and V. Yao (2021). Fintech borrowers: Lax screening or cream-skimming? *Review of Financial Studies 34*(10), 4565–4618.

Dolson, E. and J. 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, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2022). Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance 77*(1), 5–47.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019). The role of technology in mortgage lending. *Review of Financial Studies 32*(5), 1854–1899.

Hau, H., Y. Huang, H. Shan, and Z. Sheng (2019). How fintech enters China's credit market. *AEA Papers and Proceedings 109*, 60–64.

Jagtiani, J. and C. Lemieux (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management 48*(4), 1009–1029.

Jiang, E. X. (2023). Financing competitors: Shadow banks' funding and mortgage market competition. *Review of Financial Studies 36*(10), 3861–3905.

Scott, J. (2020). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.

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

US Census (2019). US median household income was $63,179 in 2018, not significantly different from 2017. *Rothbaum, Jonathan and Edwards, Ashley*. https://www.census.gov/library/stories/2019/09/us-median-household-income-not-significantly-different-from-2017.html.