The Online Payday Loan Premium*

Filipe Correia[†] Peter Han[‡] Jialan Wang[§]

March, 2024

Abstract

Using data from a subprime credit bureau with nationwide coverage in the United States, we test whether technology lowers the cost of credit in the expensive payday loan market. Contrary to the idea that technology lowers consumer prices, we find that online loans are about 100% APR higher than storefront loans. This premium is not explained by observable loan or customer characteristics including credit scores and traditional credit risk measures, and we do not find evidence of risk-based pricing on observables. Our evidence is consistent with asymmetric information and differences in fixed costs across the online and storefront markets.

Keywords: Online Lending, Payday Loans, Fintech

JEL Codes: D18, G23, G52

*We are grateful to Heitor Almeida, Scott Baker, Laura Blattner, Tatyana Deryugina, Julia Fonseca, Diego Garcia, Jitka Hilliard, Anson Ho, Rustom Irani, Timothy Johnson, Charles Kahn, Mathias Kronlund, Tom Miller, Holger Mueller, Chao Liu, Nikhil Paradkar, George Pennacchi, Andrea Presbitero, Elizaveta Sizova, Alexei Tchistyi, Sergio Vicente, Malcolm Wardlaw, Constantine Yannelis, and participants at the Gies College of Business and Terry College of Business brownbag, the 28th Finance Forum, and the 7th International Young Scholars Conference, the Edinburgh Economics of FinTech Conference, the HEC Montreal Workshop on Household Finance, and the GA Tech - Atlanta Fed Household Finance Conference for valuable suggestions. Renhao Jiang, Peichen Li, Yunrong Zhou, Hejia Liu, Kevin Tzeng, Kashish Mehta, Raja Gajula, Matthew Boyd, Yiheng Yang, Kunal Bagali, and Avishi Biljani provided valuable research assistance.

†University of Georgia, Terry College of Business, B339 Amos Hall, 620 S Lumpkin St, Athens, GA 30602. filipe.correia@uga.edu

‡University of Illinois at Urbana-Champaign, Gies College of Business, 1206 South 6th Street, Champaign, IL 61820. weitong2@illinois.edu

§University of Illinois at Urbana-Champaign & NBER. Gies College of Business, 1206 South 6th Street, Champaign, IL 61820. jialanw@gmail.com

1 Introduction

Payday loans gained popularity in the 1990s and have been a controversial credit product ever since, partly due to high prices, generally ranging between 300% to 400% APR (PEW, 2012). While the industry is believed to have relatively low barriers to entry and modest levels of concentration and profit margins (Huckstep, 2007), it also faces high fixed costs of operation and high default rates (Ernst & Young, 2009). Upon the rise of online lending in many consumer credit markets, financial technology brought the potential to lower prices and increase efficiency by reducing fixed costs and improving default prediction. Alternatively, the business ecosystem of online payday loans has costly lead generators, targeted advertising, affiliates, and lead-scoring agents. Thus, it could easily result in higher prices. These two economic forces, lower existing costs versus new ecosystem costs, lead to a natural question: are online payday loan prices higher or lower?

This paper contributes to this debate by documenting and decomposing price differences between online and storefront payday loans in the United States. We utilize a novel dataset from Clarity, a firm specializing in data collection from the subprime lending market and lead generation services, including the payday loan market. We observe loan applications and loan records. We observe loan amounts, repayments, maturities for each loan, and information on whether the loan originated at a store or online. We also have access to an array of information about the borrower, such as their income, age, ZIP code and state of residence, housing status, and homeownership status. Our dataset allows for a direct comparison between online and storefront loans, with the same data collected about the two business models.

We document the important empirical fact that online loans are \$4 per \$100 borrowed, or 100 percentage points APR more expensive. This premium is robust to controls for ex-ante predictors of risk, time and location fixed effects, and is consistent across different samples. Although realized default rates are higher online, controlling for ex-ante credit risk metrics absorbs this difference. The premium is larger when consumer fixed effects are included, with the same borrower having a larger propensity to default online.

We find no evidence of risk-based pricing at the loan level. Predictors of default observed by the lender do not strongly correlate with loan pricing. We do not detect any systematic pattern in prices or quantities across the credit score spectrum. However, borrower pools exhibiting larger default rates experience higher loan prices. This result suggests each lender sets similar prices for various consumers, regardless of observable characteristics, such as income. Lenders, however, price the credit risk of the pool of borrowers they face.

If payday loans were priced on an individual basis, we would argue that in a similar fashion to Jagtiani and Lemieux (2018), online payday lenders could use alternative sources of information and advanced algorithms to improve portfolio performance at a relatively lower cost. This would lower loan prices relative to storefront loans. Lead generators provide lead-scoring services and use other sources of information to enrich the lead. However, these services are, costly and feed a whole information supply chain, adding to the burden of the borrower endpoint.

We rationalize our findings by adapting the Stiglitz and Weiss (1981) framework of credit rationing under information asymmetry. Information asymmetry is likely to generate credit rationing in the storefront market. As a result of credit rationing, a riskier pool of borrowers will endogenously sort into the online market. Our empirical tests verify the presence of information asymmetry in the payday loan market. Information asymmetry is larger in the online market, partly justifying differences in realized default rates, and loan pricing across the two markets.

We contrast states that implemented statewide payday loan databases with states that did not. Statewide databases allow all lenders to share and access the same information about borrowers. This decreases information asymmetry in the payday loan market and equalizes the information set between online and storefront lenders. We find that, in states with a database, the premium is about 50% smaller than that of states without a database.

States with a payday loan database provide a switch from lower to higher loan amounts online, consistent with a credit rationing attenuation. We show that the online premium is larger in markets with a higher storefront market share. In those markets, the online borrower pool contains the riskiest borrowers. With a higher online market share, online lenders capture safer borrowers and the risk of the online pool decreases.

Lastly, we decompose the online payday loan premium into three components. One is the differences in risk of the borrower pool faced by each business model. Secondly, business models have different cost structures. Third, business models can differ in losses induced by credit events or recovery rates. We find evidence that cost structures are likely the most important determinant of the online payday loan premium.

Our paper directly contributes to the broader literature aiming at understanding the subprime credit market (e.g., Adams, Einav and Levin (2009)), namely the payday loan market (e.g., Nunez, Shaberg, Hendra, Servon, Addo and Mapillero-Colomina (2016)). Montezemolo (2013) reviews the legal framework of payday loans in the United States, and documents abusive practices from the lenders' angle. Wang and Burke (2022) find a significant drop in payday lending volume following mandatory state disclosure. Within the payday lending literature, our paper speaks to the interplay between pricing and business model innovation. We show that despite enjoying some leeway for regulatory arbitrage, online lending does not lead to lower loan prices.

Existing works on the payday lending market focus on welfare implications[1]. We speak to this literature by documenting the existence of a premium on online payday loans, which have become very prevalent in the last decade. Only recently, the debate on payday loan welfare started devoting attention to payday loan pricing. Saldain (2023) finds that borrowing limits and interest rate caps reduce household welfare. Relatedly, Sridhar (2023) shows

[1]See Agarwal, Skiba and Tobacman (2009), Melzer (2011), Bertrand and Morse (2011), Morse (2011), Carrell and Zinman (2014),Bhutta (2014), Bhutta, Skiba and Tobacman (2015), Carter and Skimmyhorn (2017), Dobridge (2018), Skiba and Tobacman (2019), Gathergood, Guttman-Kenney and Hunt (2019), Li and Sun (2021).

that existing rate caps are inefficiently regulated and that separating initial fees from rollover fees can be welfare-improving. Allcott, Kim, Taubinsky and Zinman (2022) estimate that while a pure payday loan ban likely generates welfare losses, limiting rollovers can increase welfare. Our paper is the first to allow for different business models within the payday loan industry and examine differences in loan pricing.

Our paper is related to the literature on costly search and price dispersion in the consumer credit market. Stango and Zinman (2016) find that consumer shopping behaviors generate economically significant dispersion in equilibrium price in credit card rates. Alexandrov and Koulayev (2018) document price dispersion in mortgage markets and a lack of rate shopping by mortgage borrowers. Argyle, Nadauld and Palmer (2023) show that search frictions in credit markets contribute to price dispersion and affect loan sizes using evidence from the auto loan market. To our knowledge, our paper is amongst the first to document price dispersion in the payday lending market and attempt to explain it.

We also contribute to the growing literature on how financial technology shapes consumer credit markets. So far, much attention has been paid to peer-to-peer lending (Morse, 2015; Li, Li, Yao and Wen, 2019), online mortgages (Fuster, Plosser, Schnabl and Vickery, 2019; Di Maggio and Yao, 2021; Chava, Ganduri, Paradkar and Zhang, 2021). Bartlett, Morse, Stanton and Wallace (2022) show that fintech fails to mitigate consumer racial/ethnic discrimination in lending. Buchak, Matvos, Piskorski and Seru (2018) attribute the rise of the originate-to-distribute logic in credit to technological advancement and regulatory frictions. Li, Liao, Wang and Xiang (2018) study consumption responses to online cash loans and find consumption increases in the digital ecosystem. Payday loans and their online dimension have not yet been addressed in fintech research. We show how fundamental differences between the in-store and online business models lead to a price differential, where online payday loans are more expensive.

Lastly, we contribute to the recent advances in the effects of information asymmetry in consumer credit markets with fintech participants. DeFusco, Tang and Yannelis (2022) back out welfare costs of information asymmetry, exogenously perturbing interest rates of a large fintech lender, in a randomized experimental setting. We add to this literature by showing how the coexistence between online and storefront lending markets with different degrees of information asymmetry and homogeneous goods generates a pooling equilibrium different from that expected from a perfectly competitive market.

The rest of the paper is organized as follows: Section 2 provides institutional background on payday lending and online payday lending; Section 3 describes the dataset, explains the data processing procedure, and puts the data in context; Section 4 documents the lack of risk-based pricing and provides evidence of the online payday loan premium; Section 5 documents asymmetric information and explains the pooling equilibrium; Section 6 decomposes the premium in differences in business model, and differences in credit risk; Section 7 concludes.

2 Background on the Payday Loan Market

Payday loans are a common source of short-term credit among low- to middle-income Americans. Between 2015 and 2019, about 2 percent of households reported using at least one payday loan per year, with higher shares among lower-income groups and higher shares that had ever used a payday loan (Kutzbach, Lloro, Weinstein and Chu, 2020). They are typically between \$300 and \$500 in principal and are structured as a single balloon payment of the amount borrowed and fees, timed to coincide with the borrower's next payday. Fees generally average between \$10 and \$20 per hundred dollars borrowed and typically do not vary with loan duration. A flat \$15 per hundred fee annualizes to nearly 400% APR for a 14-day loan corresponding to biweekly pay dates (Consumer Financial Protection Bureau, 2013).

The storefront payday industry expanded through the 1990s and early 2000s, driven in part by the loosening of state usury laws and partnership structures between payday lenders and banks to "import" regulations across state lines, a practice ended by the FDIC in the mid-2000s.[2] The online payday industry grew from a small share of loans to a significant market share over the 2010s, reaching a steady state of between 35% to 45% of the overall payday market between 2013 and 2019, with overall loan volumes including storefront and online declining from $46 billion to $25 billion annually during this period (Hecht, 2014, 2018; Graham and Golden, 2019).

The payday industry has attracted controversy and regulatory scrutiny due to high annualized costs and the high frequency of repeat borrowing. In recent years, state regulators have imposed restrictions including loan size caps, fee caps, limits on roll-over activity, cooling-off periods, and outright bans, among other measures (Kaufman, 2013). The Consumer Financial Protection Bureau became the industry's first federal regulator in 2011 . The CFPB issued rules governing the payday industry in 2017 which were largely rescinded in 2020, so regulation still largely falls on the states[3] (Kirsch, Mayer and Silber, 2014).

As of 2015, 15 states had banned traditional storefront lending. State payday laws are complex, and jurisdiction over online lending remains contested in the court system, despite many state and federal regulators enforcing laws that restrict online loans in states that also regulate storefront lending (King and Standaert, 2013; Consumer Federation of America, 2010). In addition to regulatory considerations, other features that differ between the online and storefront payday loan markets include payment and collection mechanisms that involve Automated Clearing House (ACH) transactions and bank account access instead of dated checks and in-person payment, and the online advertising market (PEW, 2014).

Online payday loans have structural differences from the storefront business model in that the former requires two additional layers of financial intermediaries, which add to the lending costs. In the first step, online payday lending starts with online lead generation. A lead is an opportunity to do business, proxied by an expression of interest by a potential customer. Leads can be tracked and traded. A merchant that sells leads to potential lenders

[2]See Mann and Hawkins (2007) for more information on the "rent-a-bank" model.
[3]See https://www.consumerfinance.gov/payday-rule/

is a lead generator. As is common in digital marketing, lead generation is heavily targeted and follows consumers in every action they perform online. According to a Wall Street Journal piece[4], 75% of online payday loan volume is sourced from lead generators, at the same time that lead generators have a large space to conduct fraudulent and abusive behavior, such as selling personal financial information without consent.

In the second step, lead aggregators do auctions to sell their lead portfolios. These auctions can be conducted via ping trees. Ping trees can be used when leads can be sold more than once to different buyers and when leads are exclusive and only sold once. They are electronic queues lenders pay to define their priority in receiving a lead. As a lead aggregator explains it on their website, ping tree placement can range from $2 to upwards of $120. Online lenders set their price points, and as leads come in, they are shown to the lenders willing to pay the highest price per lead. Lead information helps lenders decide whether to purchase the lead. Lower ping tree bids result in a lower tree position and a higher chance of lower lead quality. Higher returns from lending are promised in exchange for a higher bid for a privileged pin tree position.

3 Data

Our storefront and online payday loan data are sourced from Clarity, an alternative credit bureau and a subsidiary of Experian, one of the three major credit reporting agencies. Previous research using Clarity data includes Di Maggio, Ma and Williams (2020), Miller and Soo (2020), Miller and Soo (2021), and Fonseca (2023). Blattner and Nelson (2021) use similar data from FactorTrust, another alternative credit bureau. Clarity compiles application, origination, and repayment information for subprime loans to help lenders make underwriting decisions. Its database includes about 63 million borrowers and over 70% of non-prime consumers in the United States. Like other credit reporting agencies, Clarity relies on vol-

[4]See 'Middlemen for Payday Lenders Under Fire', The Wall Street Journal, April 7, 2014. Link: https://www.wsj.com/articles/SB10001424052702304819004579487983000120324

untary reporting of inquiries, originations, and performance by its network of lenders and data furnishers, which may not reflect the full universe of subprime loans or the universe of information from all participating lenders. Nonetheless, it is among the best sources of nationwide subprime credit activity.

The Clarity database contains information on various subprime credit products including payday, rent-to-own, installment, auto, and auto title loans. Our focus is on storefront and online payday loans in this study, which represent about 32% of inquiries and 47% of tradelines in the full database. For each inquiry, Clarity reports information about the type of loan applied for and some self-reported demographics including ZIP code and state of residence, monthly income, age, housing status, months at the same address, and paycheck frequency. While some lenders may employ income and identity verification and fraud detection mechanisms, the information reported in inquiries is self-reported by borrowers and might not be verified before submission to Clarity. For originated tradelines, we observe loan type, highest credit, scheduled and actual payment amounts, payment dates, and delinquency status.

We use two samples provided by Clarity in our analysis. The first one, known hereafter as the "standalone" or "random Clarity" sample, consists of 1 million consumers randomly drawn from Clarity's database from 2013 to 2017. According to the data provider, Clarity's full database consisted of about 63 million consumers as of 2020, so our sample contains about 1.5% of the database. The sample of borrowers includes those who apply for payday loans and other products, and only a subset of applications result in originated loans, which we use in our main analysis. In 1 million unique borrowers who submitted an inquiry for any subprime credit tradeline, 366,327 applied for an online or a storefront payday loan. Of those, 65,733 originated a payday loan, comprising our final sample.

The second sample, known hereafter as the 'credit visible' sample, consists of payday borrowers matched to a random 1% sample of all consumers in the traditional Experian credit report database as of 2018. All payday loans originated by 35,550 unique borrowers between 2013 and 2019 are included in this sample. The random Clarity and credit visible samples are drawn independently.

Because payday loan fees typically do not vary with duration, they are generally marketed to customers in terms of cost per $100 borrowed. However, lenders are also required to disclose prices in APR terms. Laws on interest rate caps are also commonly written using interest rates or APR terms. Therefore, we examine both measures of loan prices. We do not observe prices directly in the Clarity data, and infer them based on observed loan maturity, highest credit amount, and repayment amount:[5]

$$\begin{aligned} APR &= \frac{365}{LoanMaturity} \times \frac{Repayment - LoanAmount}{LoanAmount} \\ Cost\ per\ 100 &= 100 \times \frac{Repayment - LoanAmount}{LoanAmount} \end{aligned} \tag{1}$$

Because payday loans have fairly simple and standardized structures, these basic formulas accurately capture realized prices for most loans. However, one caveat is that scheduled payment amounts are missing in much of the data, so we need to use realized payments instead. This means that prices will not be accurately captured for loans not repaid in full (e.g. prices would be inferred to be zero for loans that are fully defaulted on). While defaults represent a small fraction of loans, if defaulted loans are systematically priced differently from repaid loans, our method would lead to measurement error that could be correlated with our variables of interest.

However, as shown in the next section, risk-based pricing is very limited in the online and storefront payday markets. Therefore, we do not think this potential source of measurement error drives our results. To impute prices for defaulted loans, we employ a waterfall methodology to match defaulted loans to the median price of similar non-defaulted loans within cells by origination month, loan type, state, ZIP code, and terciles of loan and

[5] See DeYoung and Phillips (2006a,b)

borrower characteristics. We try to match defaulted loans to non-defaulted ones in cells of decreasing granularity until all loans are matched (e.g. ZIP code is matched first, and if no available priced loans are matched by ZIP code, then state-level matches are used). In the analysis below, we show the results for both the full sample and the 'non-imputed' sample of loans where we measure prices from equation (1) instead of via matching. We winsorize APR and cost per \$100 at the 99th percentile in all analyses to reduce the effect of outliers.

3.1 Summary Statistics and External Validity

In this section, we report summary statistics for our sample. Table 1 presents summary statistics for the random Clarity sample in Panel A and the credit visible sample of loans matched to Experian consumer credit records in Panel B. Despite differences in the sample periods and the existence of traditional credit reports between the two samples, the descriptive statistics are extremely similar across our two samples. Figure 1 shows the geographical distribution of loans by state in both of our samples, comparing the online and storefront markets. Online loans are significantly present in all fifty states. Storefront loans are absent in some sparsely populated states and those where state laws are likely to effectively prohibit traditional payday lending during our sample period (e.g. Montana, New Mexico, and much of New England).

The random Clarity sample in Panel A of Table 1 consists of 336,690 loans from more than 65,733 borrowers, 65% of which are online. The credit visible sample in Panel B includes 188,913 loans and 35,550 borrowers with a 70% online share. By scaling our random Clarity sample by the size of the full Clarity universe and comparing to industry payday market size estimates, we calculate that Clarity represents 8% of the storefront market and 23% of the online payday market as of 2017, with coverage of the total payday market growing from 4% to 15% of originated loan volume between 2015 and 2017 (Hecht, 2014, 2018; Graham and Golden, 2019). The larger market share of online versus storefront loans likely reflects the historical evolution of Clarity's client base and online lenders' greater use of reporting and verification systems to mitigate fraud risk.

The characteristics of loans and consumers in our samples are consistent with those from previous literature and policy reports, with average loan amounts of \$365 to \$370 across our two samples and average maturities of 19 to 20 days corresponding to a combination of consumers with weekly, biweekly, and monthly pay dates (Skiba and Tobacman, 2008, Consumer Financial Protection Bureau, 2013, PEW, 2014, Wang and Burke, 2022). While loan and customer characteristics are fairly comparable to previous studies using storefront payday data (e.g. Skiba and Tobacman 2008, Wang and Burke 2022), the average borrower income of \$2822 to \$2849 is significantly lower in the Clarity online payday data compared with \$4334 among online payday borrowers in an account aggregator sample studied by Baugh (2016), which could reflect differences in income measurement or the likely higher income of consumers included in account aggregator data.

We consider a lower bound for the default rate loans reported to Clarity as not paid in full, by lenders. The average default rate of 7% in the full Clarity samples and 4% in the storefront samples are comparable to those from previous studies using administrative data from storefront payday lenders. Skiba and Tobacman (2008) report a 4% charge-off rate and Wang and Burke (2022) report a 3% default rate. Both of these previous papers report substantially higher delinquency rates than default rates, suggesting that the default rate we measure in Clarity likely corresponds to ultimate charge-offs and not to temporary delinquency. We do not attempt to distinguish between delinquency and default, track the ultimate recovery rate of defaulted loans, or account for reporting errors or reporting lags (e.g. lenders failing to report defaults to Clarity).

Average default rates for online loans are 8% to 9% in our samples, about double that of storefront loans. This contrasts with general demographics associated with lower credit risk for online loans and borrowers and with a slightly lower propensity for late payments on online loans (31%) versus storefront (35%). This discrepancy is consistent with technological advances in the online sphere. ACH processing via routing and account numbers can be done to avoid delays and cashing checks can generate payments posterior to due dates.

Online loans are significantly smaller, and online borrowers report significantly higher income and home ownership and slightly longer months at address compared with storefront borrowers in both Panels A and B. The main exception to this pattern is that online borrowers in the credit visible sample are more than twice as likely to be unscoreable compared with storefront borrowers (21% vs. 10%), and have lower Vantage scores conditional on being scoreable (500 vs. 535). Even though we classify all borrowers with a traditional Experian credit report as part of our 'credit visible' sample, some nonetheless lack valid Vantage scores, which likely reflects borrowers with thin or potentially incomplete or incorrect credit files (Blattner and Nelson, 2021). The higher income, lower age, and higher default risk associated with online payday loans are consistent with previous survey evidence (PEW, 2014).

Based on the pricing formulas and imputation algorithm described in Section 3, we find average APRs of 373% in the random Clarity sample (Panel A) and 382% in the credit visible sample (Panel B). The average cost per $100 is $16.6 in the random Clarity sample and $16.9 in the credit visible sample. Despite the assumptions needed to calculate prices in the Clarity data, these estimates are consistent with those from previous studies that use prices directly observed in administrative data from storefront payday lenders. Skiba and Tobacman (2008) report a cost per $100 of $17.9 using a sample from Texas, one of the most expensive lending markets. Wang and Burke (2022) report an average cost per $100 of $12 in a multi-state sample and $20 in Texas, corresponding to APRs of 281% and 508%. By comparison, average prices for storefront loans range from 295% to 300% APR and $12.4 to $13.3 per $100 in our Clarity samples.

Price statistics for online payday loans are rarer, and we are unaware of previous academic studies on this topic. In a survey of lender websites, Consumer Federation of America (2011) reports an average APR of 652% and cost per $100 of $25, which are substantially higher than the average APRs of 416% - 417% and cost per $100 of $18.5 - $18.9 in the Clarity samples. Despite Clarity's substantial market share of online payday loans, less-compliant or more predatory lenders who may charge higher prices and engage in other unfriendly practices toward consumers may be less likely to report to Clarity, causing a disparity relative to the sample of lender websites. Nonetheless, the significant price disparity between storefront and online loans has been widely described in industry and policy reports, so we believe the difference of well over 100% APR between these two loan types reflects true underlying heterogeneity, even if its magnitude in the full universe of payday loans is unknown. To the best our knowledge, this paper is the attempt to quantify the price difference between online and storefront payday loans.

While we use the pricing formulas in equation (1) to measure prices for most loans, these formulas rely on realized payments, which would not accurately measure prices for defaulted loans. The fourth column of Table 1 shows statistics for the non-imputed loan sample. By construction, the default rate in this sample is zero, and the proportion of loans with late payments is lower. The average loan amount is also slightly lower and repayment amount is significantly higher, but other characteristics, including prices, are similar between the imputed and non-imputed samples.

Using imputed prices allows us to investigate the important role of default rates in loan pricing which is impossible in the non-imputed sample. To support the validity of this analysis, we show that other results are remarkably similar across the imputed and non-imputed loan samples due to the lack of risk-based pricing at the individual loan level. Few loans realize default even in higher-default groups, allowing us to use non-defaulted loans to impute prices for defaulted loans. Overall, the summary statistics in Table 1 establish a basic level of external validity for our analysis and show that the pricing differences across online and payday loans are not purely driven by sample selection or measurement error.

Our payday loan sample statistics are well-aligned with Liu, Lu and Xiong (2022), who describe online loans as smaller, higher interest, earlier repayment, and more repeated borrowing. Despite not studying payday loans, they suggest that online lenders specialize in serving short-term liquidity needs, rather than providing long-term financing solutions.

4 Online Payday Loan Premium

Online payday loans are contracted at a higher price than storefront payday loans. We validate the existence this spread in graphical analysis along consumer credit risk predictors, and in a regression analysis, where we control for a rich set of confounding effects.

4.1 Risk-based Pricing in Payday Loan Markets

To understand the differences between online and storefront prices, we start by visually exploring the distribution of prices across the two business models. Figure 2 plots the kernel densities for APR in graphs (a) and (c) and cost per $100 borrowed in graphs (b) and (d) for the two Clarity samples. As with other descriptive statistics, the distributions are almost identical between the random Clarity and credit visible samples. This suggests that payday loan pricing is not substantially different for borrowers with and without existing credit reports. This is consistent with the segmentation between traditional and alternative credit markets.

Consistent with pricing schedules based on integer values of cost per $100, the distributions in graphs (b) and (d) exhibit several local modes for both online and storefront loans. Even though both distributions have common support, higher price points are more common for online loans. The pricing function is more continuous for online loans, where cost per $100 is an exact integer in 17% of observations compared with 32% for storefront. The most common integer values of cost per $100 are $15, $17, $20, and $25 for online and $8, $10, $15, and $20 for storefront loans. The interaction of discrete price points for cost per $100 and common pay frequencies leads to a multi-modal distribution of APRs, especially for storefront loans.

Potential reasons for higher prices online could be higher default incidence than at the storefront or at least differences in default predictors such as income or credit score, across the two business models. To shed initial light on these mechanisms, Figures 3 through 5 present unconditional binscatter plots of how prices and default rates change depending on loan duration, customer income, and Vantage score.

Figure 3 shows binscatter plots of prices and default risk by loan duration, typically driven by borrowers' pay frequency. Confirming our discussion above and typical practices in the industry that present uniform prices across different pay frequencies, Panel B shows that cost per $100 is flat and does not vary monotonically with loan duration. Despite that, costs are uniformly higher for online loans at all levels of loan duration. Uniform pricing by cost per $100 mechanically causes APRs to be strongly negatively correlated with loan duration, as shown in Panel A. APRs are higher for online loans conditional on loan duration. This disparity is greater in absolute terms for lower loan duration. Default risk is slightly negatively correlated with loan duration only for online loans, but there is generally a limited relationship between these two dimensions, outside accurately controlled setting as in Hertzberg, Liberman and Paravisini (2018).

Next, Figure 4 shows the relationships between prices and default risk by self-reported income. As with loan duration, Panel B shows that cost per $100 is flat for both online and storefront loans, across levels of borrower income. As shown in Panel A, APR is also flat across borrower income for online loans but is significantly positively related to income for storefront loans, which is driven by a strong negative correlation between income and loan duration. The lack of price differentiation by income contrasts with a significant negative relationship between income and default risk, shown in Panel C, which is stronger for online loans.

The greater use of credit reporting agencies such as Clarity, greater price dispersion,more continuous pricing functions, and the high overall levels of credit risk could make underwriting technology particularly valuable in this market. However, we do not observe this sophistication resulting in risk-based pricing at the loan level or driven by default predictors. Another potential demographic that drive credit risk is housing stability, as measured by the number of months a consumer has lived at their current address, shown in Appendix Figure A1, but we find limited evidence of a relationship with either default or prices, possibly due to noise in this self-reported measure.

While monthly income predicts default, credit scores encompass other dimensions of consumer credit risk. Beer, Ionescu and Li (2018) show evidence of a limited correlation between self-reported income and Vantage scores. Figure 5 presents binscatter plots of prices and default risk by Vantage score, one of the most widely-used consumer credit scores. Vantage advertises a particular ability to predict default risk for subprime and near-prime consumers who are not scoreable by other widely-used models such as FICO. In all three subfigures, consumers that are in the credit visible sample but without a valid Vantage score in the year the loan was originated are pooled and shown in the leftmost data point on the x-axis (marked as a Vantage score of 300) for comparison with scoreable consumers. The figure shows that the Vantage score predicts default risk for storefront and online payday loans, even conditional on taking out subprime credit. However, as with income, online payday loans have significantly greater credit risk at every level of Vantage. Despite its strong correlation with credit risk, APR and cost per $100 are only weakly correlated with Vantage score in the online and storefront markets, consistent with a general lack of risk-based pricing.

Lastly, we bring together loan prices and loan amounts, and correlate them with credit score, in Figure 6. We do not find any systematic pattern in pricing across the credit score and loan amount spectrum. As in Dobbie and Skiba (2013), loan amounts can be seen either as a proxy for income-based eligibility or adverse selection. Highest prices take place for large loan amounts, but we cannot assert a monotonic trend between quantities and prices.

4.2 Measuring the Online Payday Loan Premium

The previously documented relationships between loan pricing, default, and credit risk predictors support the existence of an online payday loan premium, across the income, maturity, and credit score spectrum, not explained by differences in default. The results are consistent with each lender setting their optimal cost per $100 borrowed, not distinguishing among borrowers. Establishing empirical evidence of this practice is very important to support our choice of the theoretical framework used to rationalize in section 5 to rationalize our findings.

In this section, we use regression analysis, to control for variables observable to the lender and to the econometrician. Only conditioning on those observables, we statistically test for the existence and pin down the magnitude of the online payday loan premium. We test for the existence of conditional pricing differences between the online and the storefront business models using the following specification:

$$Y_{ist} = \alpha_{is} + \alpha_t + \beta Online + X_{ist} + \epsilon_{ist} \tag{2}$$

where Y_{ist} is a price or default outcome for a given loan from customer i living in state or ZIP code s originated at time t. All regressions include borrower-location fixed effects α_{is} at borrower-ZIP code or borrower-state level, depending on the specification. α_t are time fixed effects for day of week, day of month, month of year, and calendar year. X_{ist} is a vector of controls that includes deciles of loan duration, loan size, age, and income, categorical variables for housing status and pay frequency, and number of inquiries per week as a measure of time-varying credit demand. For variables that include missing values, we include a separate category for missing values to maximize sample size. The regressions include a dummy variable for online loans with the coefficient of interest β. It estimates differences in means of the outcome, conditional on all the controls and fixed effects employed. Standard errors are clustered at the state level, as loans originated in the same state are subject to similar legal and market circumstances.

Table 2 presents our results. The table includes three columns for each outcome: APR, cost per $100, and a dummy variable indicating whether the loan defaults. The three different specifications per outcome variable include either state fixed effects, ZIP code fixed effects, or customer fixed effects. Panel A shows results for the random Clarity sample, which covers loans originated between 2013 and 2017. Panel B shows results for the credit visible sample, which covers 2013 through 2019, and Panel C includes deciles of vantage score as an additional control in the credit visible sample.

The online payday loan premium is very similar across the two samples, and robust to the inclusion of Vantage score as a control. As shown in column (1), when state fixed effects are included, the APR premium is between 92 to 98 p.p. across samples and models. Surprisingly, the premium increases when ZIP code or customer fixed effects are included instead of state fixed effects. The coefficient on the online loan dummy ranges from 103 to 110 p.p. APR when including ZIP code fixed effects, and from 127 to 141 p.p. when including customer fixed effects. As shown in columns (4) through (6), the online payday loan premium is between $3.4 and $6.5 when expressed in terms of cost per $100. For comparison, the descriptive statistics in Table 1 showed an unconditional online payday loan premium of 117-121 p.p. APR and $5.2 to $6.5 per $100 borrowed, which are within the range of the regression estimates. Overall, these results show that the online payday loan premium is not explained by differences in observable loan or customer characteristics between the two loan types.

As shown in columns (7) through (9), default risk is between 2.6 p.p. and 8.4 p.p. higher for online loans, although the online coefficient is imprecisely estimated in some models and samples using the linear probability specification. These estimates are within the range of the unconditional difference of 4 to 5 p.p. in default probability from Table 1. The large increase in both prices and default risk when including customer fixed effects reflects the fact that only 2-3% of consumers have both online and storefront loans, and these customers on average face both higher prices and higher default risk. Nonetheless, the results show that even customers with both types of loans are more likely to default on an online payday than on a storefront loan. Thus, differences in default risk are not explained by time-invariant customer characteristics.

Consistent with the lack of risk-based pricing, the inclusion of controls for Vantage score does not produce detectable differences in the estimated price premia, as shown in Panel C. Interestingly, despite Vantage score being tightly correlated with default risk, its inclusiondoes not explain the gap in default risk between online and storefront loans. Thus, the findings in Panel C further confirm that consumer characteristics explain neither the price premium nor the default gap for online payday loans.

To maximize precision and sample size, we use all available loans in both Clarity samples in the regression analysis. However, there were very few storefront payday loans in the Clarity data in 2013, so we replicate the analysis dropping 2013 in Appendix Table A1, which shows similar results to our main sample. We also replicate the analysis on the subsample of non-defaulted loans, where we calculate prices directly using equation (1) instead of imputing them from matching non-defaulted loans. These results are shown in Appendix Table A2. While this sample by definition has a default rate of zero, the estimated price premia are similar to those using the full sample that includes imputed and non-imputed prices.

Overall, the results in this section show that the online payday loan premium is not driven by differences in consumer or loan characteristics or differences in pricing models between online and storefront loans. Other factors such as differences in fixed costs, the lead generation system, and advertising and customer acquisition costs likely drive some of these differences. So far we have not tested whether higher default rates of online loans – which are even present within the same customer – are likely to be part of the explanation. To do so, we run the specification in equation (2), but with two major modifications: (i) we collapse our data at the market level, where a market is defined alternatively as a state or a ZIP code in a given week, separately for online and storefront loans and (ii) we control for both the lagged average propensity to default and for the lagged average propensity for a loan to have a late payment, in that market.

In Table 3, we report the results. As shown in column (1) across the three samples, the average storefront APR at the ZIP code level is around 300%, with estimates for the premium being close to 100 p.p., even when controlling for past default and incidence of late payments on loans in that ZIP code. Surprisingly, we observe in column (2) that the average storefront APR at the state level is much higher, around 360%. Given the ZIP code average, this much higher average is caused by some outlier loans with abnormally large APRs, happening all across the nation. This leaves a smaller room for a spread between storefront and online to be statistically detected, yielding a lack of statistical significance to the estimate from the random Clarity sample in Panel A, and a statistically significant estimate at 10% in Panel B. However, in our most conservative specification that accounts for the average Vantage score in a given market for a given week, the spread is sizeable (64.7 p.p.) and statistically significant. Furthermore, outlier APRs are normally caused by short-maturity loans, with the same flat cost per \$100, as discussed before. Hence, in columns (3) and (4) we use the cost per \$100, that is not contaminated by outlier maturities. The online payday premium is detectable and very stable across all samples, and all market definitions, ranging from around \$3-4 per \$100 borrowed at the state market denomination to \$6-7 per \$100 borrowed.

Together with the graphical analysis in subsection 4.1, our results suggest an online payday loan premium of about \$4 per \$100 borrowed, or 100 p.p. APR, not explained by ex-ante predictors of default, nor by market-level realized default.

5 Information Asymmetry and Payday Loan Prices

The absence of face-to-face interactions and the heightened risk of fraud and identity theft in online transactions can lead to increased prices and a greater likelihood of credit risk. This scenario could arise due to several factors, including the potential for selecting customers with higher unobservable risk or placing online loans lower in the repayment hierarchy for a given customer. Moreover, the complexities underlying higher default rates are further compounded by variations in collection mechanisms. It is widely believed that collection methods tend to be more assertive for online loans, given that lenders have direct access to consumers' bank accounts through the ACH network (CFPB, 2016). Our finding that online loans are less likely to result in a late payment corroborates this fact.

Previous literature has empirically examined how information asymmetry in payday loan markets affects the relationship between quantities of credit (loan amounts) and default (Dobbie and Skiba, 2013). In this section, we bring information asymmetry to rationalize prices in the payday loan market, how they relate to default, and to explain the online payday loan premium. To our knowledge, this paper is the first to conduct such an exercise.

5.1 Conceptual Framework and Testable Hypotheses

Thus far we have documented that: (i) online payday loans are more expensive than storefront payday loans; (ii) ex-ante predictors of default are not priced at the loan-level; (iii) default rates are higher online than at the storefront.

Selecting a conceptual framework to analyze payday loan prices and default presents a non-trivial decision. A conventional approach that directly links credit risk to credit prices would be at odds with our findings, which indicate that lenders do not engage in risk-based pricing.

An alternative would be to use an industrial organization framework to consider competition features within and between markets for homogeneous goods or services. However, it would not be empirically testable in our setting, since we do not observe lender identifiers, and are therefore unable to directly measure lender-level market shares, market power concentration, or elements of their cost structures.

To incorporate information asymmetry in our setting, and still be able to speak to how the industry organizes itself and to credit prices, we build on the framework from the seminal work of Stiglitz and Weiss (1981).

In their setting, information is asymmetric, and thus lenders cannot price each individual loan. Therefore, lenders choose a rate to charge every customer. The expected profit of each loan depends on the chosen rate, and on its riskiness. In turn, riskiness itself depends on the chosen rate, as an excessively high rate makes the repayment too large and defaulting becomes optimal or the consumer experiences hardship. Borrowers are heterogeneous in risk, and each borrower is only willing to borrow up to a certain price. So, as the rate increases, the pool of borrowers gets riskier. The assumptions that need to hold and how they apply to the payday loan market are described in the Appendix.

The main result of this framework is credit rationing at storefront payday lenders, who need to choose a one-size-fits-all fee per $100 borrowed, and customers self-selecting, likely adversely, to enter their stores and inquire for a loan. Instead of choosing the rate at which supply equals demand the storefront lender chooses the rate for which expected profit is maximized, under asymmetric information. This rate leaves some loan demand unattended in the storefront market. Customers not served by storefront lenders, then resort to online payday lenders.

Arnold and Riley (2009) note that under mild assumptions within the Stiglitz-Weiss setting, if there is credit rationing at a lower rate, then a second rate is also optimal. That optimal rate is higher and clears the market, being paid by higher-risk borrowers. In the same fashion, Cassar and Wydick (2012) incorporate present bias rather than undertaking risky projects as the borrower source of risk under asymmetric information, as is likely the case in the payday loan market. They assert that a new lending technology would facilitate an equilibrium where high-risk borrowers borrow at higher rates, and low-risk borrowers borrow at a low rate.

Altogether, and put in the context of the storefront-online dichotomy, it would have to be the case that online lenders would simultaneously (i) face higher asymmetry of information,

(ii) maximize their profit at a higher rate and (iii) lend to a riskier borrower pool. These are our testable hypotheses. Intuitively, the loss of soft information that can be collected in person at the store can be claimed as aggravated asymmetric information in online payday loan markets, especially if we consider the increased risk of identity theft and fraud.

The existence of the online payday loan premium establishes that online lenders likely maximize their profits at higher rates. This can be due to their superior ability to collect repayment through the ACH network, as shown by the lower occurrence of late payments, mitigating some of the moral hazard stemming from practicing higher rates. Another explanation is the ability to better diversify their portfolio, in quantities originated and geographically, due to the online nature of their lending activity.

Online prices are empirically verified to be higher, and online borrowers exhibit higher default rates. The only hypothesis lacking a formal test so far is that of establishing that the online market has a higher degree of information asymmetry. We do so in the next subsection.

5.2 Validating Testable Hypotheses

We hypothesize that there is information asymmetry in payday loan markets. We explicitly remain agnostic as to whether it manifests through adverse selection, moral hazard, or both. As is common in cases where asymmetric information is claimed, in this section, we test for correlations between the outcome of interest (loan prices) and default. A positive correlation indicates information asymmetry, either because riskier borrowers adversely select loans with higher rates, or because too high of a rate might make it optimal to default on a loan. Our goal is to test whether asymmetric information happens to a greater extent in the online market than in the storefront market.

To establish a comparable set of loan prices and defaults, we bin our loans with two different schemes. In the first scheme, we bin loans according to age $\times$ income quintiles, totaling 25 bins for online and 25 bins for storefront market. This is done in both the standalone and credit visible samples. In the second scheme, we bin loans according to credit score quintiles, only possible in the credit visible sample. This procedure ensures we always compare similar loans between markets, and within the same risk class.

Our first exercise is to ensure the online premium still holds within bins (ruling out the alternative explanation that the online premium is driven by a different composition of borrowers in the two markets). In Appendix Table A3, we report the results of running the specification in eq. 2, but including bin fixed effects. Therefore the coefficient on the online dummy is identified by average differences between the storefront and online prices within the same borrower bin. We observe an online premium of 91-124 p.p. APR, or $4-5 per $100 borrowed, which are very consistent magnitudes with the previous tables. All in all, our binning scheme does not eliminate any of the online payday loan premiums.

With this binning scheme, we test for asymmetric information, and to what extent it is more prevalent online. A direct test of asymmetric information relates loan pricing with default, and tests for a correlation. Furthermore, we test this hypothesis as suggested by Chiappori and Salanie (2000) and implemented by Dobbie and Skiba (2013). If similar borrowers obtain different prices on their loans, and higher prices are associated with higher default rates, there is evidence of asymmetric information in this market. Additionally, we interact the price variables with the online dummy to test for differences between the two business models. The specification is the following:

$$Default_{isbt} = \alpha_s + \alpha_b + \alpha_t + \beta_P Price_{ibst} + \beta_O Online_{ibst} + \beta_{O,P} Online_{ibst} \times Price_{ibst} + X_{ibst} + \epsilon_{ibst} \tag{3}$$

where the coefficient β_P tests for the correlation between loan prices and default in storefront markets, and the coefficient $\beta_{O,P}$ tests to what extent that correlation is higher online than at the storefront.

Table 4 reports the results. In Columns (1)-(6), we find a positive correlation between loan prices and default. Again, we do not claim a direction: it could be that within the same bin, borrowers more likely to default choose higher rates, or that a higher repayment makes borrowers more likely to default. This association is more accentuated in online loans as reported by the interaction terms, most of them statistically significant at 10%, across different samples, outcomes, and binning schemes. Importantly, due to the lack of individual-level risk-based-pricing, we can have high confidence that the correlation shown in Table 4 is not a simple reflection of the risk premium imposed by the lenders on riskier borrowers.

We take the evidence in Table 4 as support for asymmetric information in payday loan markets and that such asymmetry is higher online. In practice, payday lenders do not perform hard credit pulls from credit bureaus of prime credit products. This is particularly acute in the online market, where lenders do not observe beyond noisy self-reported information until far later in the process, making these results reasonably expected.

Graphically, we explore this relation separately for online and storefront loans using binscatter plots of each loans' APR relative to their bins' average APR and default. What we aim to compare is the slope of the relation. The steeper it is, the more robust is the evidence of asymmetric information in credit markets. We report our findings in the credit visible sample in Figure 7. On the two left panels, we show the binscatter plot for storefront loans. On the two right panels we report the same graph for online loans. Panels (a)-(b) report correlations between APR (relative to bin average) and default. Panels (c)-(d) report correlations between cost per $100 borrowed (relative to bin average) and default. The results are consistent with Table 4. There is an upward-sloping relationship between APR and credit risk within the same bin, and it is stronger in the online market, suggesting a higher degree of asymmetric information online. These results are robust to sampling, as shown Appendix Figure A2, where we plot the same analysis using the random Clarity sample.

To complement our results, we use an alternative test, as proposed by Chiappori and

Salanie (2000). We directly estimate the correlation between price and credit risk via bivariate probit estimation. The advantage of this approach is the absence of a stance regarding the dependent or independent variable. The disadvantage is that it requires two binary variables. We use an indicator variable of ex-post loan default as the endogenous variable for credit risk. We use an indicator of whether the loan price lies above the average price of the borrowers' bin as the endogenous variable for loan price.

We report results for the estimate for correlation between loan price and default in Table 5. In Columns (1) and (3) we report correlations between prices and defaults in the storefront market, and in Columns (2) and (4) we report correlations for the online market. Each of panels A, B, and C reports estimations on the random Clarity sample, on the credit visible sample, and the credit visible sample controlling for vantage score, respectively. Across all samples and all outcomes, it is consistently true that there is a positive correlation between credit prices and default. That correlation is higher in the online payday loan market.

Lastly, if our interpretation of the pooling market equilibrium is correct, we can make two testable claims. First, a smaller online market share, leads to an online borrower pool extremely risky, compared to the larger, much safer storefront borrower pool. Second, as a result, the online premium should be larger when the online market share is lower. In Figure 8, we plot binscatter plots of default rates of each pool (online vs storefront) in each market (state-year), in panels (a) and (c), and the online premium in each market, in panels (b) and (d), sorted by online market share. We observe exactly what our predictions indicate: a larger online market share, captures safer borrowers to the online pool, and makes the riskiness of the two pools similar, reducing the online payday loan premium.

5.3 Evidence from Statewide Payday Loan Databases

We have quantified the online payday loan premium and conducted empirical tests to detect asymmetric information within the payday loan market. Furthermore, we have demonstrated that the heightened degree of asymmetric information inherent in online payday loans becomes apparent when subjected to classic tests for information asymmetry.

We argue that information asymmetry serves as a rationale for the observed higher realized default rates and loan prices online, particularly in a market where default predictors are not factored into loan pricing.

The crucial question that arises is whether attenuating information asymmetry would impact the magnitude of the online payday loan premium. Most importantly, if both online and brick-and-mortar lenders had access to identical information, would such a significant disparity in their prices still exist?

To test this, we use cross-sectional variation across state laws that establish a statewide payday loan database, where lenders are mandated to report the loans they provide, and which can be accessed through the payment of a small fee. There are 13 states which have this database in place: Alabama, Delaware, Florida, Illinois, Indiana, Kentucky, Michigan, North Dakota, Oklahoma, South Carolina, Virginia, Washington, and Wisconsin, where 19% of the loans of our sample are originated. These databases have two effects: first, they diminish information asymmetry for the market as a whole, as all the lenders in a state can now observe the same information. Second, they decrease the differences in information asymmetry between online and storefront lenders.

A limitation arising from our samples ranging 2013-2017 and 2013-2019 is that we cannot directly exploit the effects of introducing a database in a given state, as the 13 states implemented databases during 2002-2012. Notwithstanding, the cross-sectional variation still allows us to test the asymmetric information explanation. To do so, we estimate the following specification:

$$Y_{ist} = \alpha_t + \alpha_s + \beta_O Online_{ist} + \beta_{DID} Online_{ist} \times Database_s + X_{ist} + \epsilon_{ist} \tag{4}$$

where Y_{ist} is the price of a loan i, both measured as APR (in p.p.) or cost per \$100 borrowed (in dollars), for a customer living in state s originating at time t. We run the specification with and without time fixed effects α_t, and loan and customer level controls X_{ist}. We cannot include state fixed effects, as the cross-sectional variation of the variable $Database_s$ is state-level. The three most important coefficients are β_O, which measures the online premium in states without a database, where asymmetric information is higher and different between storefront and online lenders; and β_{DID} the difference in online premia between states with and without databases. The size of β_{DID}, relative to β_O allows us to speak to the economic significance of information asymmetry in explaining the two-price, dual-credit risk equilibrium.

Table 6 reports the results for the estimation of equation 4 on both the random Clarity sample (Panel A) and the credit visible sample (Panel B). The online payday loan premium ranges between 102-122 APR percentage points or \$4.1-\$5.8 per \$100 dollars borrowed in states without a database. Even though accessing the database incurs a cost for the lender, legislation provisions such as Wis.Stat. § 138.14(10)(b)1. state "A licensee may not assess a customer any fee or charge for database access or usage."

In states with a database, our point estimates indicate that the online premium has about half of the magnitude of the premium in states without a database. The β_{DID} coefficient is statistically significant at the 10% level in all specifications for cost per \$100 borrowed, but is not, when using APR as the dependent variable.

As an extended test, in columns (5)-(6), we use loan amounts as the dependent variable. In states without a database, loan amounts are smaller online. However, in states with a database, loans are larger online. This is consistent with databases reducing information asymmetries online, leading to a lesser extent of credit rationing.

In our best attempt to use all the time series variation we can, we study pricing dynamics following the implementation of a database, and whether the records added to it through time have any impact on (i) loan prices on the market as a whole, and (ii) in the online payday loan premium. To do so, we document the year in which each state implemented its database, and for each loan in our sample, we compute the number of years since the implementation of the database, yielding variations along the time series dimension. We then re-estimate equation 4, but we decompose the $Database_s$ variable into dummy variables that indicate whether state s has implemented a database less than 6 years before, 7-8, 8-10, 10-12, or more than 13 years ago. This binning scheme is the most granular we can have where 3 or more states are used to identify the coefficient for each bin.

In Figure 9, we report the results for our two price outcomes, estimated in our two different samples. The *Baseline* estimates are estimates of online (green) and storefront (orange) prices, and the premium (black) in states without a database. We then plot estimated prices for each market and the premium (obtained from the estimates in Appendix Table A4) for each time bin relative to database implementation in the state.

Three expected patterns arise: (i) The equilibrium interest rates in the states without payday loan database are higher than those in the states without the database. This pattern is consistent with the estimations in Table 6. (ii) As time passes, after the introduction of payday loan database, the average equilibrium interest rates in both the online and storefront markets decrease. This pattern is consistent with the increase in market efficiency in payday loan markets via the reduction in information asymmetry when all the lenders share the same information set; and (iii) the gap between online prices and storefront prices remains statistically indistinguishable from zero as time goes by. The standard deviation of the interest rate gap is large in the first few years and drops to a much lower value as time passes. This pattern is consistent with the stabilization of the payday lending market after the changes from introducing a shared database. All three patterns are robust across samples and different definitions of price.

6 Decomposing the Online Payday Loan Premium

Asymmetric information explains the pooling equilibrium evident in the payday loan market, encompassing both online and brick-and-mortar business models. The presence of information asymmetry alone can give rise to an equilibrium where two distinct loan prices coexist. However, what remains uncertain is whether the observed premium in information asymmetry justifiably accounts the magnitude of the online premium. In this section, we delve into the significance of other factors in elucidating the price differential across various business models.

We start by formalizing the relationship between two quantities that we observe: the price and default rate in a given market, for each business model. We model cost per \$100 borrowed, as it is independent of loan maturity. Business model i prices their loans following:

$$c_{im}(\lambda_{im}) = \pi_i + \theta_i \lambda_{im} + \varepsilon_{im} \tag{5}$$

where c_{im} is the cost per \$100 of business model i in market m, and λ_{im} is the default rate experienced by business model i in market m. Note that we allow each business model to have different default rates in a given market, as that matches our empirical evidence. π_i is the gross profit per \$100 that business model i requires, even in the absence of any default. Economically, it has to cover all the costs the business model incurs to originate a loan. θ_i is the compensation business model i demands per unit of risk. This term will be a function of recovery rates, among others: a larger loss in default, leads to a higher compensation demanded for that risk. A business model i is then defined by the pair $\{\pi_i, \theta_i\}, i = O, S$.

Using a variation of the Oaxaca-Blinder decomposition, we can write the online payday loan premium as the difference between the two pricing equations:

$$c_{Om} - c_{Sm} = (\pi_O - \pi_S) + \lambda_{Sm}(\theta_O - \theta_S) + \theta_O(\lambda_{Om} - \lambda_{Sm}) + (\varepsilon_{Om} - \varepsilon_{Sm}) \tag{6}$$

Equation 6 decomposes the premium into differences in business model fundamentals, differences in credit risk, and differences in unexplained factors. Differences in business model fundamentals can be broken into different cost structures and different compensations for default events. The decomposition assumes $\mathbb{E}(\varepsilon_{Om} - \varepsilon_{Sm}) = 0$.

To estimate the contribution of each component, we define a market with three alternative binnings: state-year, income deciles by Vantage deciles by state, Vantage deciles by year by month. The latter two binnings are only feasible for the credit visible sample.

Table 7 reports the estimates for the decomposition of unconditional differences in cost per \$100 between online and storefront payday loan business models. The difference in loan prices ranges between \$3.1 and \$5.2, depending on binning schemes. Using columns solely (3)-(4) to assess the qualitative consistency of our estimates, we rely on columns (1)-(2) as robust definitions of a market, and comparable across samples.

While differences in default rates carry statistical and economic significance, they account for only about 25% of the online payday loan premium, with less than 25% of this premium attributed to discrepancies in default losses. Surprisingly, the primary component of the online payday loan premium stems from variation in cost structures between business models, explaining approximately 50% of the online premium.

This discovery aligns with the notion that online payday lenders cater to borrowers who are visibly similar to those served by brick-and-mortar lenders but face higher default rates. It also resonates with the idea that online lenders encounter greater susceptibility to fraud and recoup smaller portions of defaulted amounts. We interpret the significance of differences in cost structures as a consequence of the intricate network of affiliates, ping trees, lead generators, and aggregators. This aspect stands as a fundamental distinction between online and brick-and-mortar lending models, directly influencing cost structures, and in turn, loan prices.

7 Conclusion

This paper presents novel evidence of an online payday loan premium. Using data from a national subprime credit bureau, we show that despite the potential for online technology to lower fixed costs and increase lending efficiency, online payday loans incur a higher cost by $4 per $100 or 100 p.p. APR, even conditioning on loan and customer characteristics. Interestingly, neither brick-and-mortar nor online payday loans appear to utilize risk-based pricing, and available metrics of default risk fail to elucidate the price premium.

Default rates are about double for online payday loans compared with storefront loans, and customers with both types of loans are much more likely to default on online loans. However, predictors of default and realized default do not absorb the premium. From an economic standpoint, the patterns we observe are reconciled in an information asymmetry framework that generates credit rationing at the storefront and clears the market with riskier borrowers facing higher rates in the online payday loan market.

Our findings underscore the significance of disparities in cost structures between both business models as the predominant component of the online payday loan premium. This paper speaks to the relevance of evaluating the online and brick-and-mortar credit markets concurrently, rather than in isolation, as similar dynamics can arise in the presence of financial frictions.

This is the first paper asserting the existence of a premium in online payday loans, measuring, and decomposing it. The study of online lending business models is pivotal for understanding their pricing mechanisms and the impact of FinTech players on consumer credit, particularly in markets catering to the most vulnerable borrowers.

References

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity constraints and imperfect information in subprime lending," *American Economic Review*, 2009, *99* (1), 49–84.

Agarwal, Sumit, Paige Marta Skiba, and Jeremy Tobacman, "Payday loans and credit cards: New liquidity and credit scoring puzzles?," *American Economic Review*, 2009, *99* (2), 412–17.

Alexandrov, Alexei and Sergei Koulayev, "No shopping in the us mortgage market: Direct and strategic effects of providing information," *Consumer Financial Protection Bureau Office of Research Working Paper*, 2018, (2017-01).

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman, "Are high-interest loans predatory? theory and evidence from payday lending," *The Review of Economic Studies*, 2022, *89* (3), 1041–1084.

Argyle, Bronson, Taylor Nadauld, and Christopher Palmer, "Real effects of search frictions in consumer credit markets," *The Review of Financial Studies*, 2023, *36* (7), 2685–2720.

Arnold, Lutz G and John G Riley, "On the possibility of credit rationing in the stiglitz-weiss model," *American Economic Review*, 2009, *99* (5), 2012–21.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, "Consumer-lending discrimination in the FinTech era," *Journal of Financial Economics*, 2022, *143* (1), 30–56.

Baugh, Brian, "Payday borrowing and household outcomes: Evidence from a natural experiment," 2016.

Beer, Rachael, Felicia Ionescu, and Geng Li, "Are income and credit scores highly correlated?," *FEDS Notes*, 2018.

Bertrand, Marianne and Adair Morse, "Information disclosure, cognitive biases, and payday borrowing," *The Journal of Finance*, 2011, *66* (6), 1865–1893.

Bhutta, Neil, "Payday loans and consumer financial health," *Journal of Banking & Finance*, 2014, *47*, 230–242.

_ , Paige Marta Skiba, and Jeremy Tobacman, "Payday loan choices and consequences," *Journal of Money, Credit and Banking*, 2015, *47* (2-3), 223–260.

Blattner, Laura and Scott Nelson, "How costly is noise? Data and disparities in consumer credit," *Unpublished Working Paper, arXiv preprint arXiv:2105.07554*, 2021.

Blinder, Alan S, "Wage discrimination: reduced form and structural estimates," *Journal of Human resources*, 1973, pp. 436–455.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, "Fintech, regulatory arbitrage, and the rise of shadow banks," *Journal of Financial Economics*, 2018, *130* (3), 453–483.

Carrell, Scott and Jonathan Zinman, "In harm's way? Payday loan access and military personnel performance," *The Review of Financial Studies*, 2014, *27* (9), 2805–2840.

Carter, Susan Payne and William Skimmyhorn, "Much ado about nothing? New evidence on the effects of payday lending on military members," *Review of Economics and Statistics*, 2017, *99* (4), 606–621.

Cassar, Alessandra and Bruce Wydick, "Credit rationing with behavioral foundations: Revisiting Stiglitz and Weiss," *Unpublished Working Paper*, 2012.

Chava, Sudheer, Rohan Ganduri, Nikhil Paradkar, and Yafei Zhang, "Impact of marketplace lending on consumers' future borrowing capacities and borrowing outcomes," *Journal of Financial Economics*, 2021, *142* (3), 1186–1208.

Chiappori, Pierre-André and Bernard Salanie, "Testing for asymmetric information in insurance markets," *Journal of Political Economy*, 2000, *108* (1), 56–78.

Consumer Federation of America, "States Have Jurisdiction over Online Payday Lenders," Technical Report 2010.

_ , "CFA Survey of Online Payday Loan Websites," Technical Report 2011.

Consumer Financial Protection Bureau, "Payday Loans and Deposit Advance Products: A White Paper of Initial Data Findings," Technical Report 2013.

DeFusco, Anthony A, Huan Tang, and Constantine Yannelis, "Measuring the welfare cost of asymmetric information in consumer credit markets," *Journal of Financial Economics*, 2022, *146* (3), 821–840.

DeYoung, Robert and Ronnie J Phillips, "Payday loan pricing," *Networks Financial Institute*, 2006, pp. 09–07.

_ **and Ronnie Phillips**, "Strategic pricing of payday loans: Evidence from Colorado, 2000-2005," *Networks Financial Institute Working Paper*, 2006.

Di Maggio, Marco and Vincent Yao, "Fintech borrowers: lax screening or cream-skimming?," *The Review of Financial Studies*, 2021, *34* (10), 4565–4618.

Dobbie, Will and Paige Marta Skiba, "Information asymmetries in consumer credit markets: Evidence from payday lending," *American Economic Journal: Applied Economics*, 2013, *5* (4), 256–82.

Dobridge, Christine L, "High-cost credit and consumption smoothing," *Journal of Money, Credit and Banking*, 2018, *50* (2-3), 407–433.

Ernst & Young, "The Cost of Providing Payday Loans in a US Multiline Operator Environment," Technical Report 2009.

Fonseca, Julia, "Less mainstream credit, more payday borrowing? Evidence from debt collection restrictions," *The Journal of Finance*, 2023, *78* (1), 63–103.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, "The role of technology in mortgage lending," *The Review of Financial Studies*, 2019, *32* (5), 1854–1899.

Gathergood, John, Benedict Guttman-Kenney, and Stefan Hunt, "How do payday loans affect borrowers? Evidence from the UK market," *The Review of Financial Studies*, 2019, *32* (2), 496–523.

Graham, Karen and Elaine Golden, "Financially Underserved Market Size Study 2019," Technical Report 2019.

Hecht, John, "Alternative Financial Services: Innovating to Meet Customer Needs in an Evolving Regulatory Framework," Technical Report 2014.

_ , "Short Term Lending Update: Moving Forward with Positive Momentum," Technical Report 2018.

Hertzberg, Andrew, Andres Liberman, and Daniel Paravisini, "Screening on loan terms: evidence from maturity choice in consumer credit," *The Review of Financial Studies*, 2018, *31* (9), 3532–3567.

Huckstep, Aaron, "Payday lending: Do outrageous prices necessarily mean outrageous profits," *Fordham J. Corp. & Fin. L.*, 2007, *12*, 203.

Jaffee, Dwight M and Thomas Russell, "Imperfect information, uncertainty, and credit rationing," *The Quarterly Journal of Economics*, 1976, *90* (4), 651–666.

Jagtiani, Julapa and Cathy Lemieux, "The Roles of Alternative Data and ML in Fintech Lending: Evidence from LendingClub Consumer Platform," *Unpublished Working Paper*, 2018.

Kaufman, Alex, "Payday Lending Regulation," *Finance and Economics Discussion Series*, 01 2013, *2013*, 1–38.

King, Uriah and Diane Standaert, "Effective State and Federal Payday Lending Enforcement: Paving the Way for Broader, Stronger Protections," Technical Report, Center for Responsible Lending 2013.

Kirsch, Larry, Robert N Mayer, and Norman I Silber, "The CFPB and payday lending: New agency/old problem," *Journal of Consumer Affairs*, 2014, *48* (1), 1–16.

Kutzbach, M, A Lloro, J Weinstein, and K Chu, "How America Banks: Household Use of Banking and Financial Services, 2019 FDIC Survey," 2020.

Li, Emma, Li Liao, Zhengwei Wang, and Hongyu Xiang, "Individual Consumption Response to Expanding Credit Access: Evidence from online cash loan platform," *PBCSF-NIFR Research Paper*, 2018.

Li, Tsung-Hsien and Jan Sun, "The Payday Loan Puzzle: A Credit Scoring Explanation," *Available at SSRN 3954233*, 2021.

Li, Zhiyong, Ke Li, Xiao Yao, and Qing Wen, "Predicting prepayment and default risks of unsecured consumer loans in online lending," *Emerging Markets Finance and Trade*, 2019, *55* (1), 118–132.

Liu, Lei, Guangli Lu, and Wei Xiong, "The big tech lending model," *National Bureau of Economic Research*, 2022.

Maggio, Marco Di, Angela T Ma, and Emily Williams, "In the Red: Overdrafts, Payday Lending and the Underbanked," 2020.

Mann, Ronald J and Jim Hawkins, "Just Until Payday," *UCLA Law Review*, 2007, *54*, 855.

Melzer, Brian T, "The real costs of credit access: Evidence from the payday lending market," *The Quarterly Journal of Economics*, 2011, *126* (1), 517–555.

Miller, Sarah and Cindy K Soo, "Does increasing access to formal credit reduce payday borrowing?," 2020.

_ and _ , "Do neighborhoods affect the credit market decisions of low-income borrowers? evidence from the moving to opportunity experiment," *The Review of Financial Studies*, 2021, *34* (2), 827–863.

Montezemolo, Susanna, "Payday lending abuses and predatory practices," *The State of Lending in America & Its Impact on US Households, Center for Responsible Lending*, 2013.

Morse, Adair, "Payday lenders: Heroes or villains?," *Journal of Financial Economics*, 2011, *102* (1), 28–44.

_ , "Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending," *Annual Review of Financial Economics*, 2015, *7*, 463–482.

Nunez, Stephen, K Shaberg, Richard Hendra, Lisa Servon, Mina Addo, and A Mapillero-Colomina, "Online payday and installment loans: Who uses them and why? A demand side analysis from linked administrative, survey, and qualitative interview data," *Subprime Lending Data Exploration Project*, 2016.

Oaxaca, Ronald, "Male-female wage differentials in urban labor markets," *International economic review*, 1973, pp. 693–709.

PEW Charitable Trusts, "Payday Lending in America: Who Borrows, Where They Borrow, and Why," *Washington, DC*, 2012.

_ , "Fraud and abuse online: Harmful practices in internet payday lending," *Washington, DC*, 2014.

Saldain, Joaquin, "High-Cost Consumer Credit: Desperation, Temptation and Default," *Unpublished Working Paper*, 2023.

Skiba, Paige and Jeremy Tobacman, "Payday Loans, Uncertainty and Discounting: Explaining Patterns of Borrowing, Repayment, and Default," Vanderbilt Law and Economics Research Paper No. 08-33 08 2008.

Skiba, Paige Marta and Jeremy Tobacman, "Do payday loans cause bankruptcy?," *The Journal of Law and Economics*, 2019, *62* (3), 485–519.

Sridhar, Sharada, "Liquidity Based Contracting: A Path to Greater Efficiency in Payday Lending," *Unpublished Working Paper*, 2023.

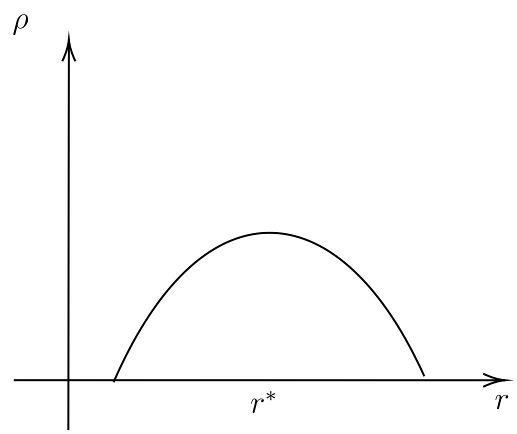
Stango, Victor and Jonathan Zinman, "Borrowing high versus borrowing higher: price dispersion and shopping behavior in the US credit card market," *The Review of Financial Studies*, 2016, *29* (4), 979–1006.

Stiglitz, Joseph E and Andrew Weiss, "Credit rationing in markets with imperfect information," *The American Economic Review*, 1981, *71* (3), 393–410.

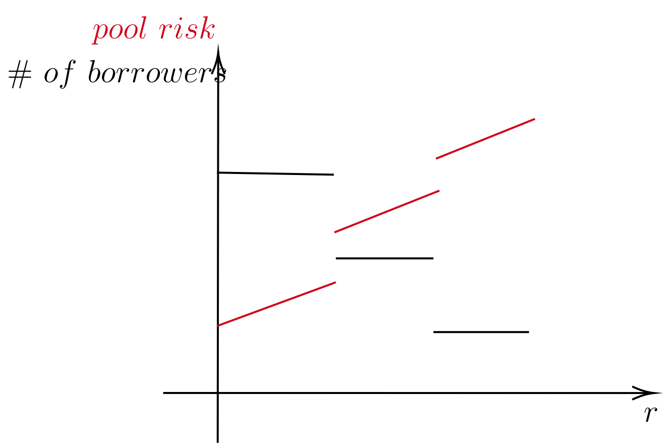
Wang, Jialan and Kathleen Burke, "The effects of disclosure and enforcement on payday lending in Texas," *Journal of Financial Economics*, 2022, *145* (2), 489–507.

A Appendix

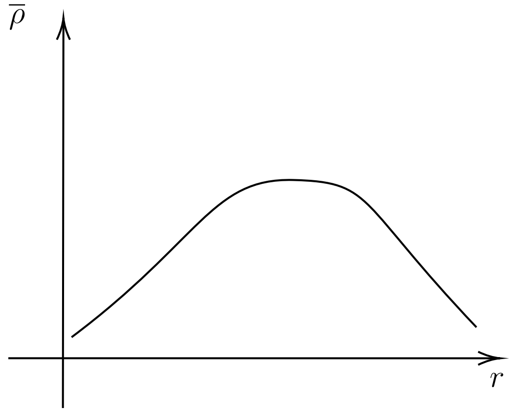
Four assumptions need to hold in the Stiglitz and Weiss (1981) framework. We address them one by one in explaining how the payday loan market is empirically likely to verify them. The first assumption is that there is a rate at which the profitability of a given loan is maximized. Beyond that rate, the lender selects risky borrowers, high repayment makes it optimal to default, or both.



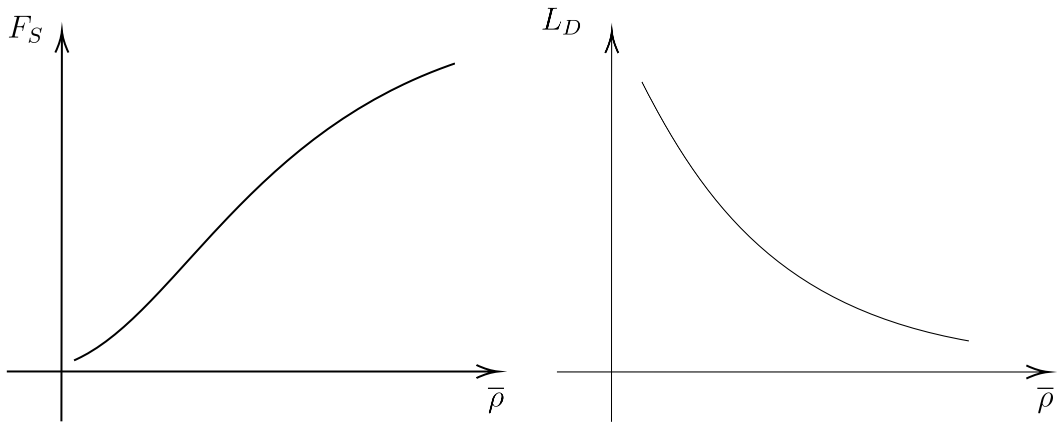
This is likely to hold in the payday loan market. As empirical evidence, the absence of risk-based pricing, even at the storefront, supports the idea of a market where information is asymmetric, therefore generating this pattern. The second assumption is that borrowers are heterogeneous in risk, and given their risk, there is a rate at which each borrower exits the borrower pool, leaving the remaining pool of borrowers with higher average riskiness.



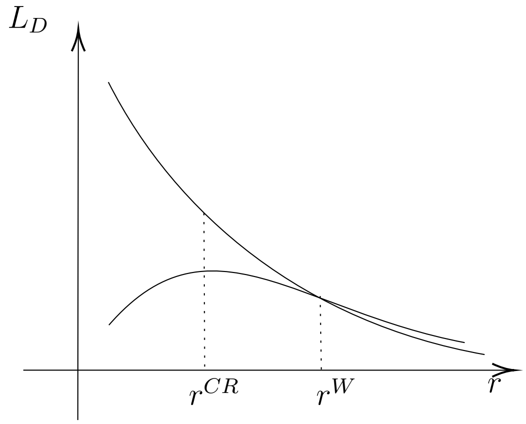
Altogether, the two first assumptions, in the simplest context of this model, generate a hump-shaped expected profit function of the interest rate charged, with a well defined optimal interest rate, for which the expected profit of the loan portfolio is maximized.



The two last assumptions are conventional. The third assumption is that lenders compete for loanable funds, and that under a competitive environment, ultimately they will offer their expected return as compensation to the investor, who is willing to supply more loanable funds for a larger the return, leading to a zero net profit situation. The supply of loanable funds is therefore increasing in profitability. Payday loans are not funded by deposits, like bank loans. They have to compete for loanable funds, and being a high credit risk product, the competition is necessarily through the return they are able to offer. Lastly, the fourth assumption is that demand for loans is decreasing in interest rate. Like any usual demand function, also likely to hold in payday loan markets.



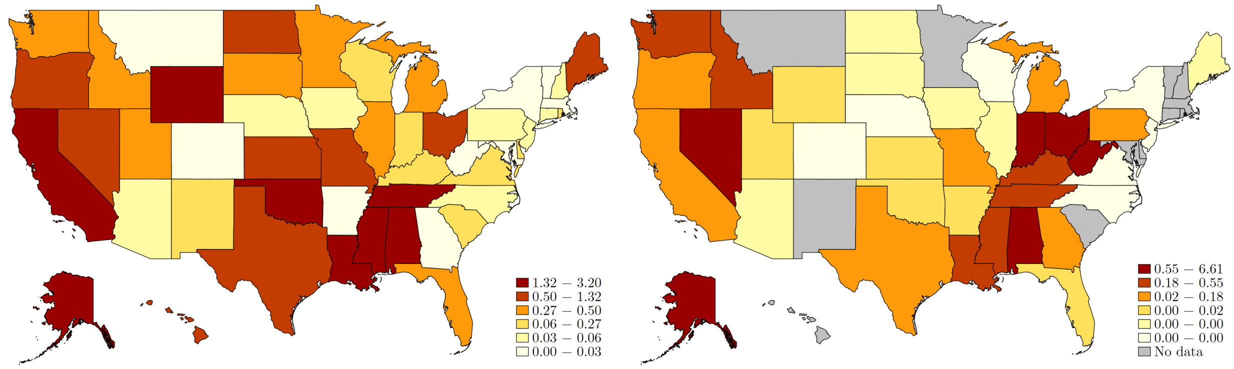
The equilibrium determination happens with tension between the usual Walrasian interest rate, at which the supply of credit equals the demand of credit, and a credit rationing scenario as in Jaffee and Russell (1976), where the lender ultimately chooses an interest rate that maximizes their profit, even if it implies supplying a quantity of credit inferior to the demanded quantity.



L_D
r^{CR}
r^W
r

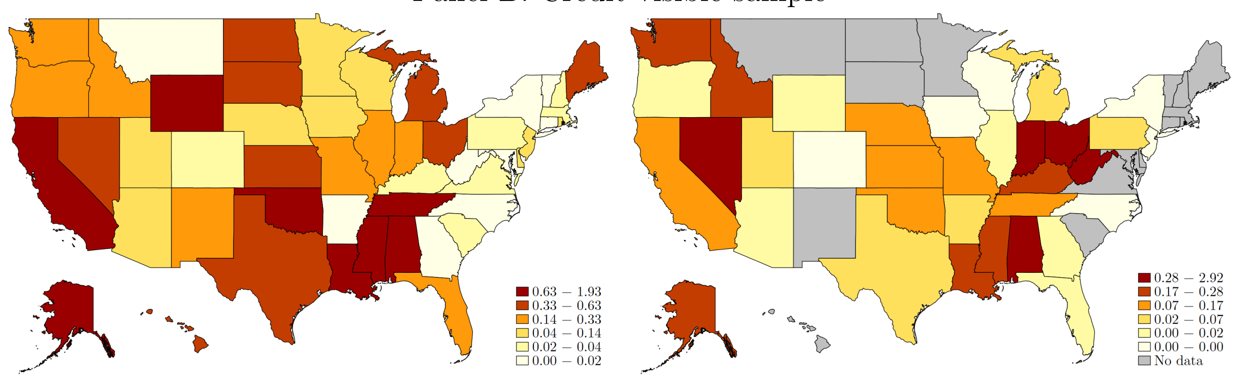
Figure 1: Geographic Distribution of Loans per Capita

Panel A: Random Clarity sample



(a) Online (b) Storefront

Panel B: Credit visible sample

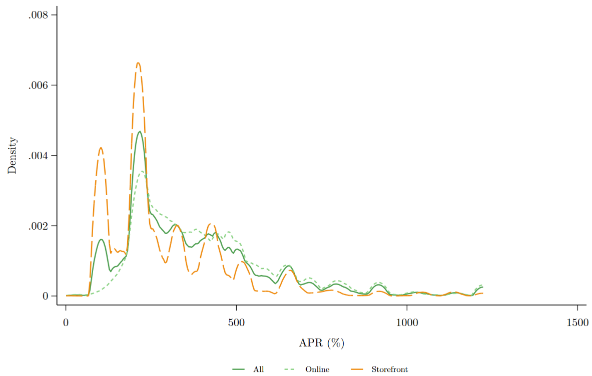


(c) Online (d) Storefront

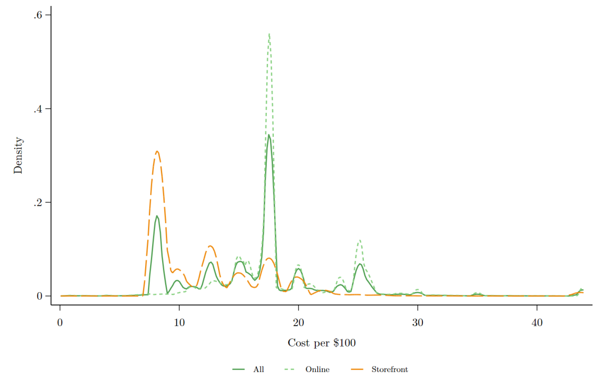
Note: In this figure we map the geographical distribution of payday loans by state. Graphs show total numbers of online and storefront loans per 1000 people in the state population based on the 2010 Census. The random Clarity sample presented in Panel A consists of a random sample of 1 million unique borrowers that submitted loan inquiries in Clarity's full database between 2013 and 17. Only originated payday loans from this sample of consumers are included in the analysis sample. The credit visible sample shown in Panel B consists of payday borrowers that are matched to a random 1% sample of all consumers in the Experian credit bureau database in 2018. All loans originated by matched borrowers between 2013 and 2019 are included in this sample.

Figure 2: Online and Storefront Price distributions

Panel A: Random Clarity sample

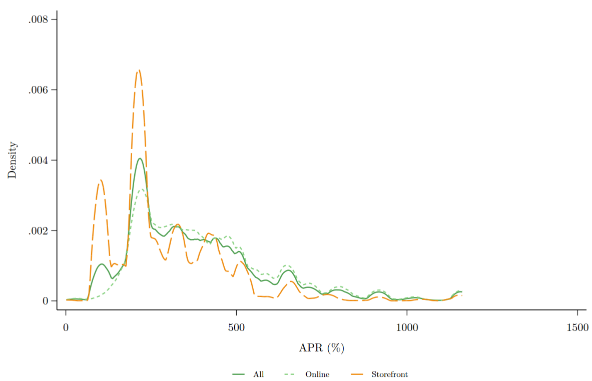


(a) APR

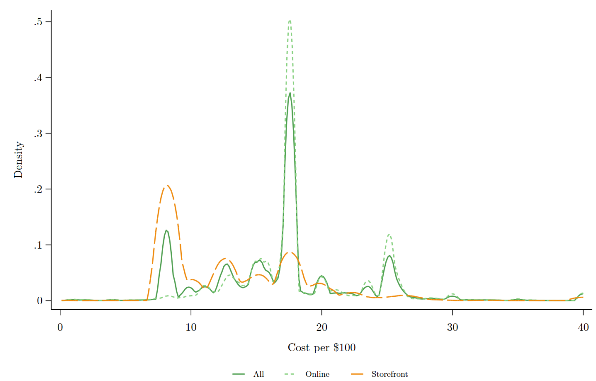


(b) Cost per $100

Panel B: Credit visible sample



(c) APR

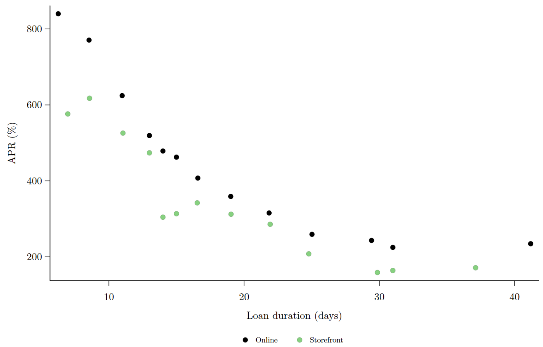


(d) Cost per $100

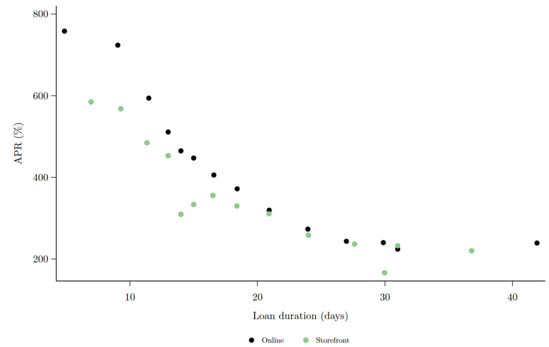
Note: This figure shows the distributions of prices for online and storefront payday loans. The random Clarity sample presented in Panel A consists of a random sample of 1 million unique borrowers that submitted loan inquiries in Clarity's full database between 2013 and 17. Only originated payday loans from this sample of consumers are included in the analysis sample. The credit visible sample shown in Panel B consists of payday borrowers that are matched to a random 1% sample of all consumers in the Experian credit bureau database in 2018. All loans originated by matched borrowers between 2013 and 2019 are included in this sample.

Figure 3: Prices and Default Rates by Loan Duration

Panel A: APR

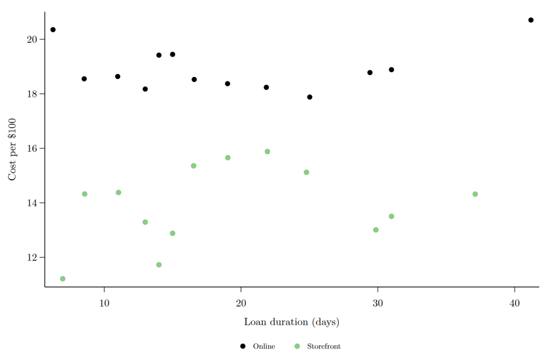


(a) Random Clarity sample

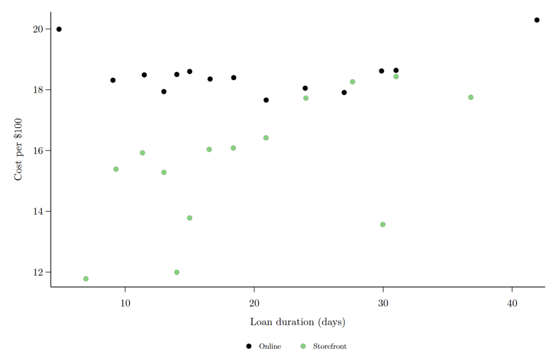


(b) Credit visible sample

Panel B: Cost per $100

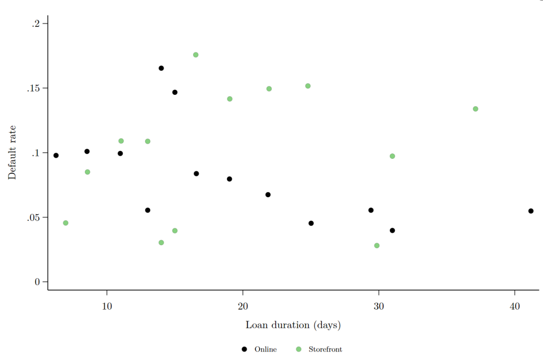


(c) Random Clarity sample

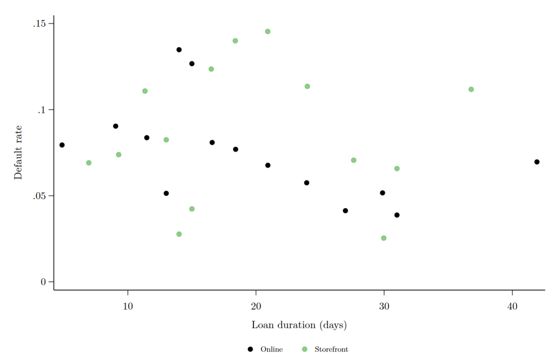


(d) Credit visible sample

Panel C: Default



(e) Random Clarity sample

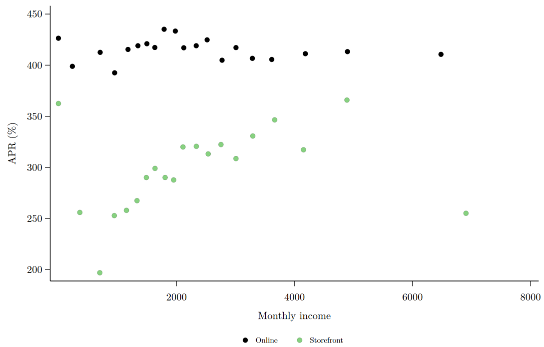


(f) Credit visible sample

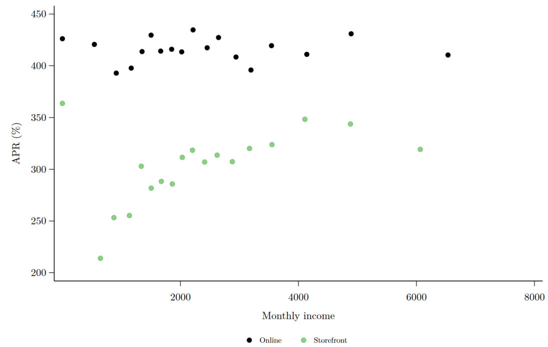
Note: The figure presents binscatter plots of APR, cost per $100 borrowed, and default rates in the credit visible sample in graphs (b), (d), and (f), and in the standalone sample in graphs (a), (c), and (e). The x-axis represents the loan maturity.

Figure 4: Prices and Default Rates by Borrower Income

Panel A: APR

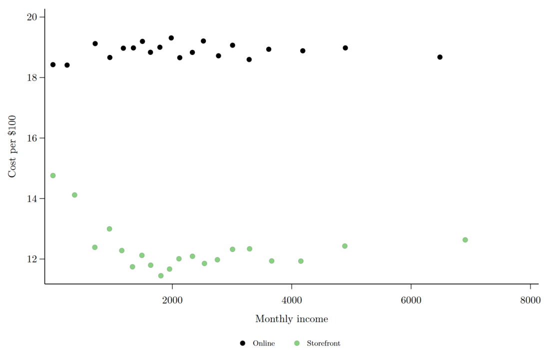


(a) Random Clarity sample

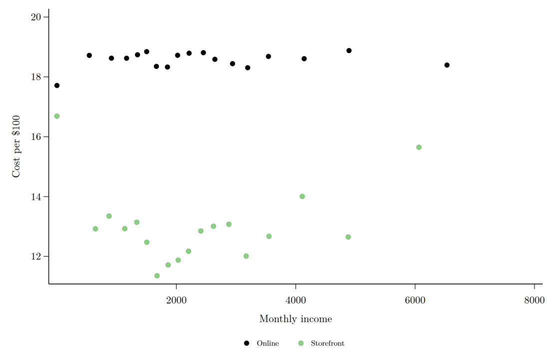


(b) Credit visible sample

Panel B: Cost per $100

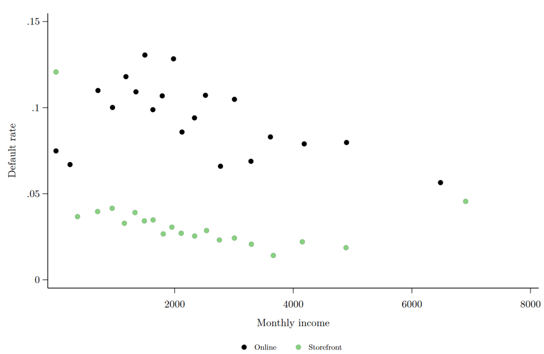


(c) Random Clarity sample

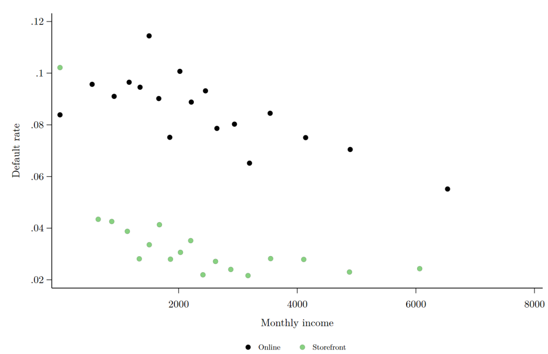


(d) Credit visible sample

Panel C: Default



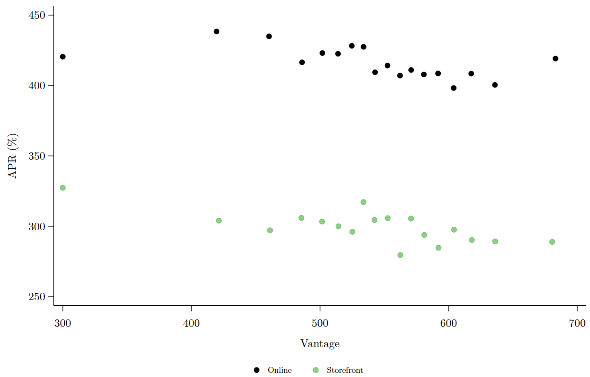
(e) Random Clarity sample



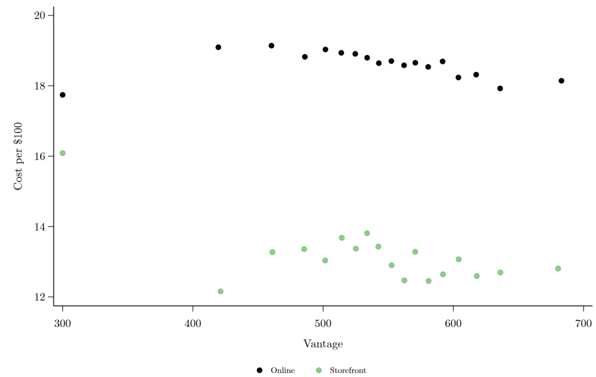
(f) Credit visible sample

Note: The figure presents binscatter plots of APR, cost per $100 borrowed, and default rates in the credit visible sample in graphs (b), (d), and (f), and in the standalone sample in graphs (a), (c), and (e). The x-axis represents the borrower's monthly income.

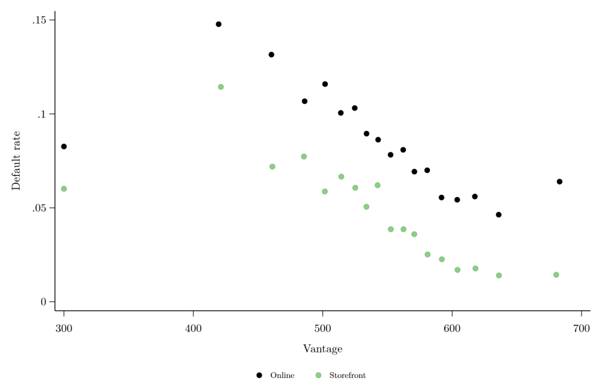
Figure 5: Prices and Default Rates by Vantage Score



(a) APR



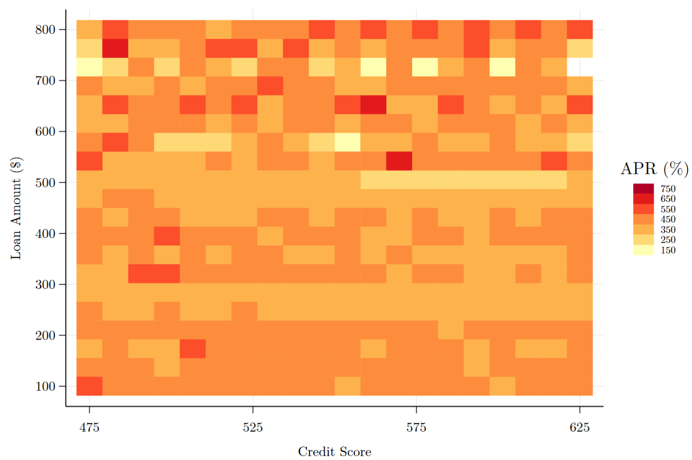
(b) Cost per $100



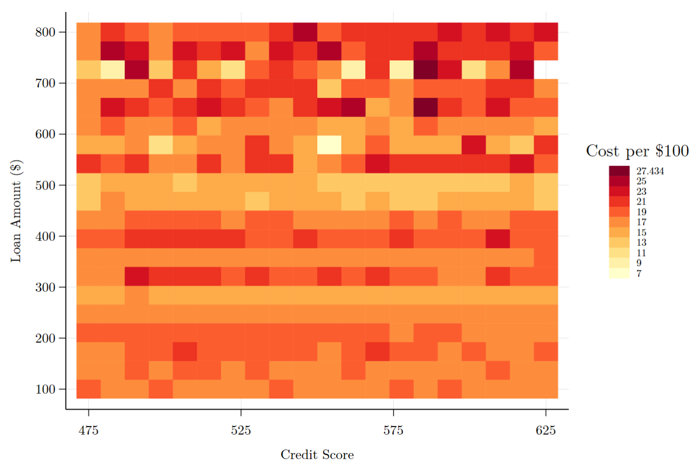
(c) Default rate

Note: The figure presents binscatters of APR, cost per $100 borrowed, and default rates in the credit visible sample. Missing Vantage scores are set to 300 and outcomes for these borrowers are shown as the leftmost data point in each graph.

Figure 6: Absence of Risk-Based Pricing



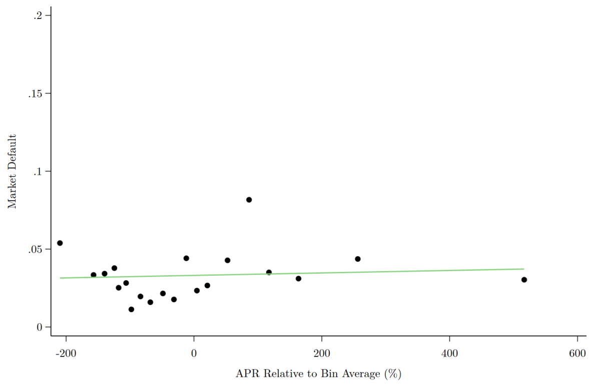
(a) APR (%)



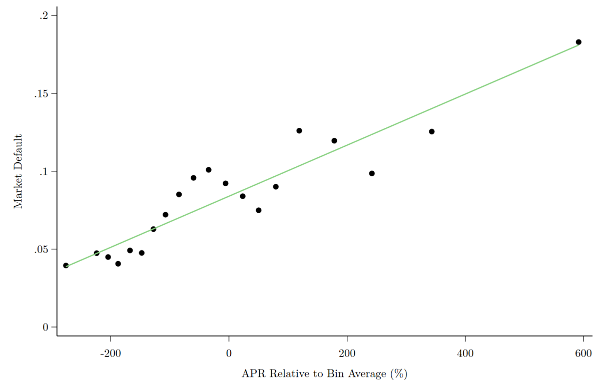
(b) Cost per $100

Note: We plot loan prices (in color scale) versus loan amounts and credit scores. In panel (a), loan price is measured as APR, and in panel (b) loan price is measured as cost per $100.

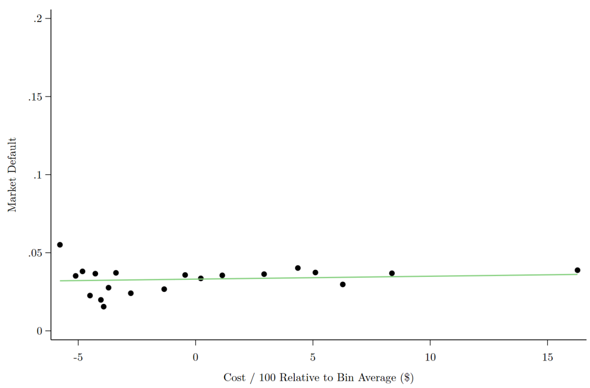
Figure 7: Prices and Default



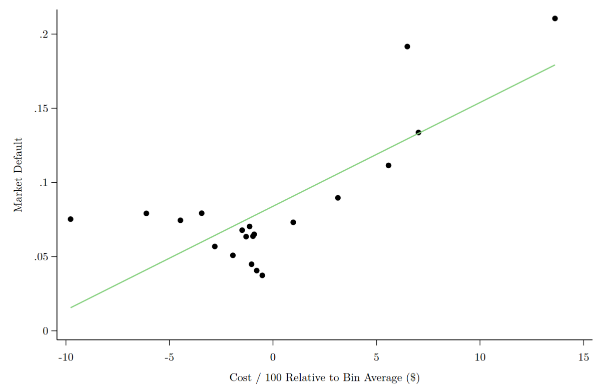
(a) APR vs Default, Storefront



(b) APR vs Default, Online



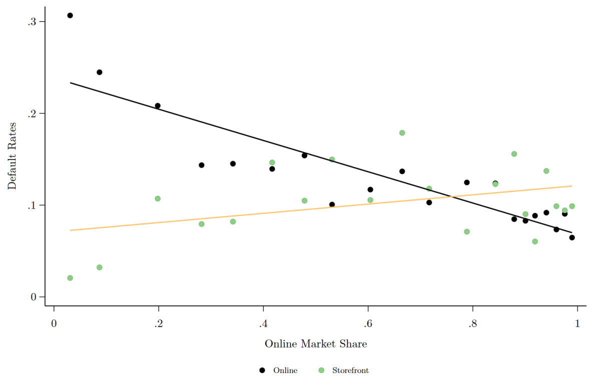
(c) Cost per $100 vs Default, Storefront



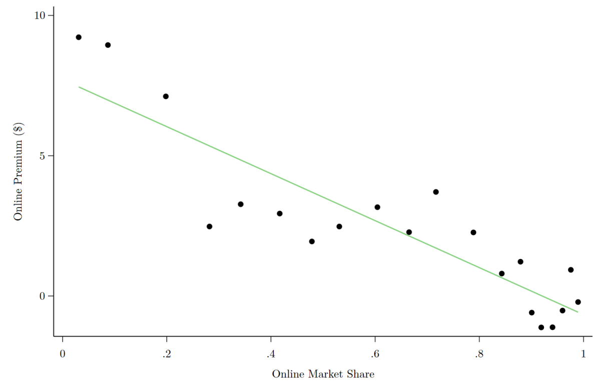
(d) Cost per $100 vs Default, Online

Note: The figure presents binscatter plots of default by bins of price relative to the bin average price in the credit visible sample. Prices are expressed as APR in panels (a) and (b) and cost per $100 borrower in panels (c) and (d). Bins are age-income bins.

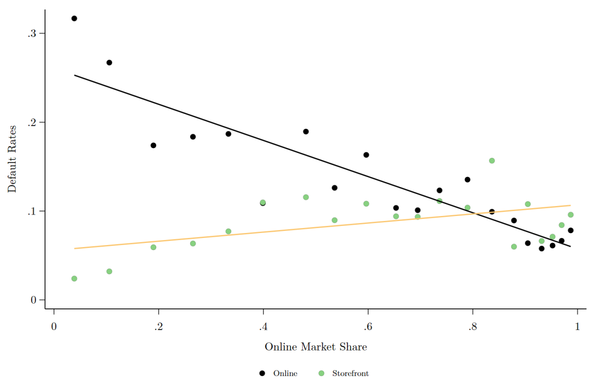
Figure 8: Online Market Shares, Pool Risk, and the Online Premium



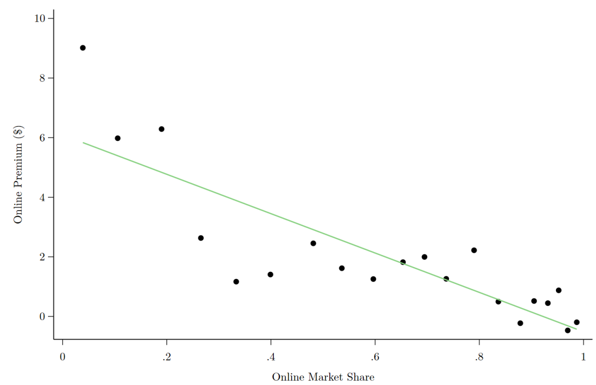
(a) Pool risk, Random Clarity Sample



(b) Online Premium, Random Clarity Sample



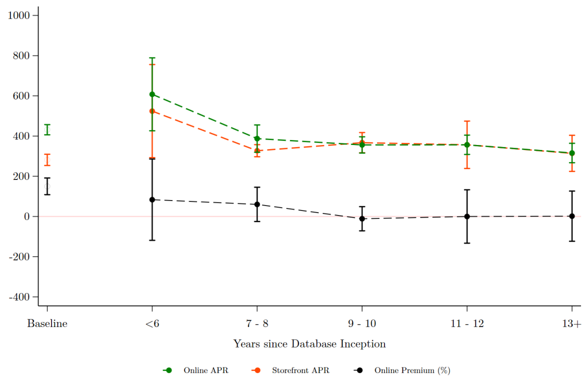
(c) Pool risk, Cerit Visible Sample



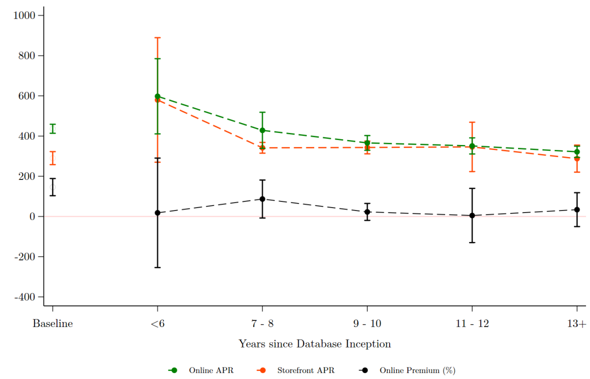
(d) Online Premium, Credit Visible Sample

Note: The figure presents binscatter plots of default rates online and storefront, in panels (a) and (c), and the online premium, in panels (b) and (d), along markets with different online market shares. Panels (a) and (b) are built from the random Clarity sample, and panels (c) and (d) are built from the credit visible sample.

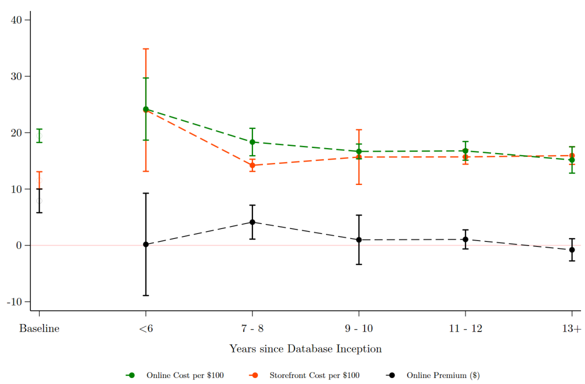
Figure 9: Statewide Payday Loan Database



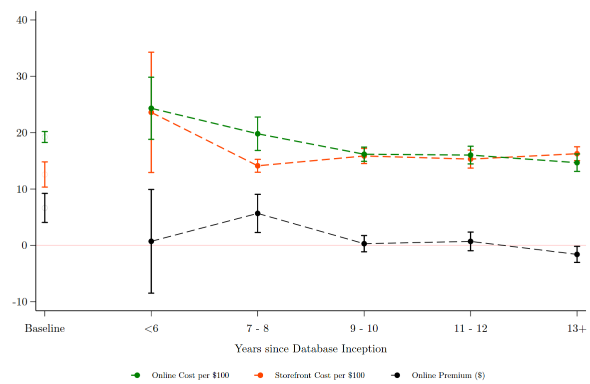
(a) APR, Random Clarity Sample



(b) APR, Credit Visible Sample



(c) Cost per $100, Random Clarity Sample



(d) Cost per $100, Credit Visible Sample

Note: The figure presents coefficients for loan-level regressions of prices and 95% confidence intervals, in APR and cost per $100 borrowed, on a dummy variable that indicates whether the loan is an online loan, and dummy variables for each bin of time since implementation of a statewide payday loan database, and their interaction. Regressions contain loan and customer-level controls, and time fixed effects. Robust standard errors are clustered at the state level.

Table 1: **Summary Statistics**

Panel A: Random Clarity Sample (2013-2017)						
Subsample:	All			Non-imputed	Online	Storefront
	Mean	Median	SD	Mean	Mean	Mean
Loan Characteristics						
Loan Amount ($)	365	260	265	344	316	456
Repayment Amount ($)	369	300	285	397	303	490
Loan Maturity (days)	20	15	9	19	20	19
Default	7%	0%	26%	0%	9%	4%
Late Payment	32%	0%	47%	26%	31%	35%
APR	373%	322%	217%	363%	416%	295%
Cost per $100 ($)	16.6	17.3	6.3	16.1	18.9	12.4
Online Loan	65%	100%	48%	62%	100%	0%
Self-Reported Information						
Owns Home	15%	0%	35%	13%	19%	5%
Age	42.5	41.0	14.0	43.1	39.9	47.3
Months at Address	29.1	24.0	23.9	28.6	29.5	25.6
Net Monthly Income	2545	2200	1490	2533	2849	1970
# of Loans		336,690		272,220	217,596	119,094
# of Unique Borrowers		65,733		46,010	49,877	17,484

Panel B: Credit Visible Sample (2013-2019)						
Subsample:	All			Non-imputed	Online	Storefront
	Mean	Median	SD	Mean	Mean	Mean
Loan Characteristics						
Loan Amount ($)	370	255	284	342	332	460
Repayment Amount ($)	372	300	299	396	320	494
Loan Maturity (days)	19	15	9	19	19	19
Default	7%	0%	26%	0%	8%	4%
Late Payment	32%	0%	47%	26%	32%	34%
APR	382%	336%	208%	372%	417%	300%
Cost per $100 ($)	16.9	17.5	5.8	16.4	18.5	13.3
Vantage score	510	538	113	512	500	535
Unscoreable	18%	0%	38%	18%	21%	10%
Online Loan	70%	100%	46%	68%	100%	0%
Self-Reported Information						
Owns Home	16%	0%	36%	14%	20%	6%
Age	42.1	41.0	13.5	42.4	40.1	46.6
Months at Address	30.5	24.0	24.2	30.0	31.0	26.0
Net Monthly Income	2578	2244	1514	2576	2844	1956
# of Loans		188,913		149,458	132,520	56,393
# of Unique Borrowers		35,550		24,654	27,473	9,097

Note: Table contains summary statistics for two samples of online and storefront payday loans from Clarity. The random Clarity sample presented in Panel A consists of a random sample of 1 million unique borrowers that submitted loan inquiries in Clarity's full database between 2013-17. Only inquiries resulting in originated payday loans are included in the analysis sample. The credit visible sample shown in Panel B consists of payday borrowers who are matched to a random 1% sample of all consumers in the Experian credit bureau database in 2018. All inquiries and loans originated by matched borrowers between 2013 and 2019 are included in this sample. The two samples are drawn independently. Each panel shows statistics for the full set of loans, 'non-imputed' loans where prices are calculated directly from loan-level terms instead of imputed based on loans with similar characteristics (see text for details), online payday loans, and storefront payday loans.

Table 2: **Online Payday Loan Premium**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Random Clarity Sample (2013-2017)								
Outcome:	APR			Cost / 100			Default		
Storefront mean:	295			12.4			0.041		
Online dummy	98.4	110.3	141.2	4.35	4.65	6.52	0.030	0.026	0.084
	(36.0)	(41.7)	(54.3)	(2.20)	(2.52)	(3.07)	(0.026)	(0.032)	(0.038)
	[0.009]	[0.011]	[0.012]	[0.053]	[0.071]	[0.039]	[0.253]	[0.422]	[0.032]
R^2	0.534	0.615	0.775	0.582	0.663	0.812	0.139	0.202	0.511
N	336,690	332,937	305,775	336,690	332,937	305,775	336,690	332,937	305,775
	Panel B: Credit Visible Sample (2013-2019)								
Outcome:	APR			Cost / 100			Default		
Storefront mean:	300			13.3			0.044		
Online dummy	91.9	103.1	127.1	3.36	3.70	4.91	0.036	0.034	0.060
	(25.8)	(33.7)	(54.2)	(1.48)	(1.78)	(2.38)	(0.020)	(0.027)	(0.028)
	[0.001]	[0.004]	[0.023]	[0.027]	[0.042]	[0.045]	[0.080]	[0.202]	[0.041]
R^2	0.531	0.605	0.743	0.543	0.623	0.757	0.121	0.167	0.320
N	188,913	186,687	171,518	188,913	186,687	171,518	188,913	186,687	171,518
	Panel C: Credit Visible Sample with Vantage								
Online dummy	91.9	103.1	127.2	3.37	3.71	4.91	0.036	0.035	0.060
	(25.8)	(33.7)	(54.2)	(1.48)	(1.78)	(2.38)	(0.020)	(0.027)	(0.029)
	[0.001]	[0.004]	[0.023]	[0.027]	[0.042]	[0.044]	[0.077]	[0.199]	[0.041]
R^2	0.531	0.605	0.743	0.544	0.623	0.757	0.125	0.171	0.320
N	188,913	186,687	171,518	188,913	186,687	171,518	188,913	186,687	171,518
State FE	Yes	No	No	Yes	No	No	Yes	No	No
Zip FE	No	Yes	No	No	Yes	No	No	Yes	No
Consumer FE	No	No	Yes	No	No	Yes	No	No	Yes

Note: The table presents coefficient estimates of the online loan dummy from regressions of payday loan prices, default probability, and late payment probability for the random Clarity sample in Panel A and the credit visible sample in Panels B and C. All regressions include fixed effects for either state, ZIP code, or customer; fixed effects for day of week, day of month, month of year, and calendar year; and controls for deciles of loan duration, loan size, age, and income, categorical variables for housing status and pay frequency, and number of inquiries per week. Panel C additionally includes controls for decile of Vantage score. Robust standard errors clustered at the state level are in parentheses, and p-values are in brackets.

Table 3: **Online Payday Loan Premium - Market Level**

	(1)	(2)	(3)	(4)
	Panel A: Random Clarity Sample (2013-2017)			
Outcome:	APR		Cost / 100	
Storefront mean:	304	360	12.7	16.3
Online dummy	120.2	31.5	6.6	2.7
	(7.3)	(29.4)	(0.3)	(1.2)
	[0.000]	[0.290]	[0.000]	[0.027]
Market Border	ZIP Code	State	ZIP Code	State
R2	0.621	0.543	0.720	0.512
N	220,140	10,574	220,140	10,574
	Panel B: Credit Visible Sample (2013-2019)			
Outcome:	APR		Cost / 100	
Storefront mean:	306	357	13.7	16.6
Online dummy	126.7	62.2	6.1	3.5
	(8.7)	(30.9)	(0.4)	(1.3)
	[0.000]	[0.049]	[0.000]	[0.008]
Market Border	ZIP Code	State	ZIP Code	State
R2	0.619	0.546	0.723	0.551
N	120,152	11,623	120,152	11,623
	Panel C: Credit Visible Sample with Vantage			
Outcome:	APR		Cost / 100	
Online dummy	126.7	62.2	6.1	3.5
	(8.7)	(30.9)	(0.4)	(1.3)
	[0.000]	[0.049]	[0.000]	[0.008]
Market Border	ZIP Code	State	ZIP Code	State
R2	0.619	0.546	0.723	0.551
N	120,152	11,623	120,152	11,623

Note: The table presents coefficient estimates of the online loan dummy from regressions of payday loan prices for the random Clarity sample in Panel A and the credit visible sample in Panels B and C. The unit of observation is market-week, and markets are defined as ZIP codes in columns (1) and (3) and states in columns (2) and (4). All regressions include fixed effects for either state or ZIP code and week fixed effects and controls for averages of loan duration, loan size, age, income, number of inquiries per week, and lagged default rates and lagged percentage of late payments in the market. Panel C controls for average Vantage score of applicants in that market in that week. Robust standard errors clustered at the market level are in parentheses, and p-values are in brackets.

Table 4: **Asymmetric Information - APR and Default**

	(1)	(2)	(3)	(4)	(5)	(6)
			Propensities (p.p.)			
Outcome:			Default			
Online Dummy x APR	0.005	0.003	0.000			
	(0.001)	(0.001)	(0.000)			
	[0.000]	[0.000]	[0.001]			
APR	0.003	0.003	0.000			
	(0.000)	(0.001)	(0.000)			
	[0.000]	[0.000]	[0.000]			
Online Dummy x Cost per $100				0.007	0.061	0.079
				(0.028)	(0.035)	(0.034)
				[0.793]	[0.079]	[0.021]
Cost per $100				0.334	0.185	0.162
				(0.019)	(0.024)	(0.024)
				[0.000]	[0.000]	[0.000]
Online Dummy	5.950	5.360	0.054	6.730	5.060	4.760
	(0.286)	(0.380)	(0.004)	(0.566)	(0.692)	(0.684)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R2	0.094	0.075	0.078	0.095	0.075	0.077
N	304,404	157,400	157,400	304,404	157,400	157,400
Sample	Stand-Alone	Credit Visible	Credit Visible	Stand-Alone	Credit Visible	Credit Visible
Binning Quintiles	Age-Income	Age-Income	Credit Score	Age-Income	Age-Income	Credit Score
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table presents coefficient estimates of correlations between payday loan prices and default for the random Clarity sample in Columns (1) and (4) and the credit visible sample in Columns (2)-(3) and (5)-(6). The unit of observation is loan, and bins are defined by quintiles of age and income in columns (1)-(2) and (4)-(5) , and credit score in columns (3) and (6). All regressions include fixed effects for bin and week fixed effects and controls for averages of loan duration, loan size, age, income, number of inquiries per week. Robust standard errors clustered at the market level are in parentheses, and p-values are in brackets.

Table 5: **Asymmetric Information - Bivariate Probit**

	(1)	(2)	(3)	(4)
	Panel A: Random Clarity Sample (2013-2017)			
	APR > Bin Avg. vs. Default		Cost per $100 > Bin Avg. vs. Default	
Rho (ρ)	0.053	0.222	0.057	0.297
	(0.010)	(0.005)	(0.010)	(0.005)
	[0.000]	[0.000]	[0.000]	[0.000]
N	105,002	199,402	105,002	199,402
Sample	Storefront	Online	Storefront	Online
	Panel B: Credit Visible Sample (2013-2019)			
Rho (ρ)	0.095	0.191	0.029	0.271
	(0.014)	(0.007)	(0.014)	(0.007)
	[0.000]	[0.000]	[0.039]	[0.000]
N	46,937	110,463	46,937	110,463
Sample	Storefront	Online	Storefront	Online
	Panel C: Credit Visible Sample with Vantage			
Rho (ρ)	0.154	0.191	0.012	0.267
	(0.014)	(0.007)	(0.014)	(0.007)
	[0.000]	[0.000]	[0.400]	[0.000]
N	46,937	110,463	46,937	110,463
Sample	Storefront	Online	Storefront	Online

Note: The table presents bivariate Probit coefficient estimates of correlations between payday loan prices and default for the random Clarity sample in Panel A, in Panel B and C the credit visible sample, being Panel C controlled for credit score. In columns (1) and (2) we present correlation estimates for the bivariate distribution of default and a dummy if the loan has an APR larger than the average APR within its bin. In Columns (3) and (4), instead of APR, we use cost per $100 borrowed. Standard errors are in parentheses, and p-values are in brackets.

Table 6: **Statewide Payday Loan Database - Difference in Differences**

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Random Clarity Sample (2013-2017)					
Outcome:	APR (p.p.)		Cost per \$100 (\$)		Loan Amounts (\$)	
Online x Database	- 46.6	- 49.0	- 4.9	- 3.6	55.6	107.3
	(50.6)	(38.3)	(1.6)	(1.5)	(29.6)	(42.8)
	[0.357]	[0.202]	[0.002]	[0.016]	[0.062]	[0.013]
Online	122.4	109.1	5.8	5.2	- 17.2	- 50.4
	(36.6)	(26.5)	(1.5)	(1.4)	(25.0)	(25.5)
	[0.001]	[0.000]	[0.000]	[0.000]	[0.493]	[0.049]
Controls	No	Yes	No	Yes	No	Yes
FE	No	Yes	No	Yes	No	Yes
R2	0.199	0.520	0.545	0.573	0.166	0.261
N	336,690	336,690	336,690	336,690	336,690	336,690
	Panel B: Credit Visible Sample (2013-2019)					
Outcome:	APR (p.p.)		Cost per \$100 (\$)		Loan Amounts (\$)	
Online x Database	- 39.2	- 34.1	- 4.1	- 2.5	79.4	134.9
	(36.7)	(31.1)	(1.3)	(1.4)	(28.1)	(37.2)
	[0.287]	[0.274]	[0.001]	[0.072]	[0.005]	[0.000]
Online	120.7	101.5	4.8	4.1	- 27.3	- 61.3
	(27.3)	(20.3)	(1.2)	(1.1)	(22.3)	(20.4)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.222]	[0.003]
Controls	No	Yes	No	Yes	No	Yes
FE	No	Yes	No	Yes	No	Yes
R2	0.207	0.518	0.486	0.530	0.153	0.225
N	188,913	188,913	188,913	188,913	188,913	188,913

Note: The table presents coefficient estimates of loan-level regressions of prices, in APR and cost per \$100 borrowed, and loan amounts, on a dummy variable that indicates whether the loan is an online loan, and a dummy variable that indicates whether the state where the tradeline originated has a statewide payday loan database, and their interaction. Columns (2), (4), and (6) contain loan and customer-level controls, and time fixed effects. The samples used are in Panel A and the random Clarity sample, in Panel B the credit visible sample. Robust standard errors clustered at the state-year level are in parentheses, and p-values are in brackets.

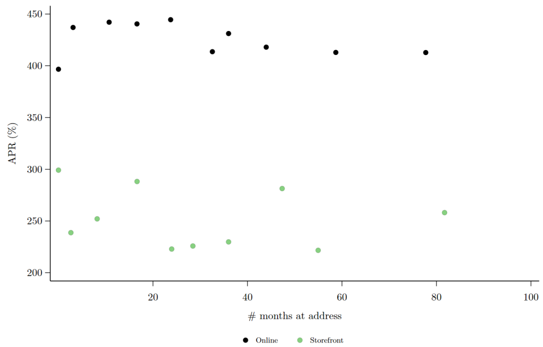
Table 7: **Oaxaca-Blinder Decomposition**

	(1)	(2)	(3)	(4)
c_O	21.3	19.7	20.8	19.1
c_S	16.4	16.5	17.1	13.9
$c_O - c_S$	4.8	3.1	3.7	5.2
	(0.2)	(0.2)	(0.2)	(0.2)
	[0.000]	[0.000]	[0.000]	[0.000]
$\theta_O(\lambda_O - \lambda_S)$	1.4	0.8	1.0	1.0
	(0.1)	(0.1)	(0.1)	(0.1)
	[0.000]	[0.000]	[0.000]	[0.000]
$\lambda S(\theta_O - \theta_S)$	1.1	0.6	0.8	0.5
	(0.1)	(0.1)	(0.1)	(0.1)
	[0.000]	[0.000]	[0.000]	[0.000]
$(\pi_O - \pi_S)$	2.3	1.7	1.9	3.8
	(0.3)	(0.2)	(0.2)	(0.2)
	[0.000]	[0.000]	[0.000]	[0.000]
Online # Bins	2,336	2,990	3,686	660
Storefront # Bins	1,017	1,336	1,263	565
Sample	Random Clarity	Credit Visible	Credit Visible	Credit Visible
Bins	State-Month	State-Month	State-Income-Vantage	Vantage-Month

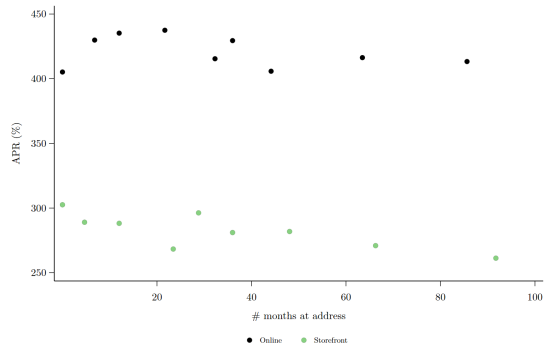
Note: This table contains estimates for the Oaxaca-Blinder decomposition of the unconditional differences between online and storefront payday loan prices. We decompose them in differences in default rates $\theta_O(\lambda_O - \lambda_S)$, differences in default losses $\lambda S(\theta_O - \theta_S)$, and differences in cost structures $(\pi_O - \pi_S)$. In columns (1) and (2) we use state-month bins in the random Clarity, and the credit visible sample, respectively. Columns (3) and (4) allow for binnings using deciles of Vantage score, using the credit visible sample.

Figure A1: Prices and Default Rates by Months at Address

Panel A: APR

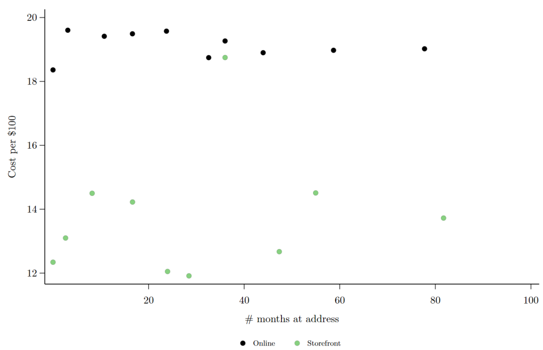


(a) Random Clarity sample

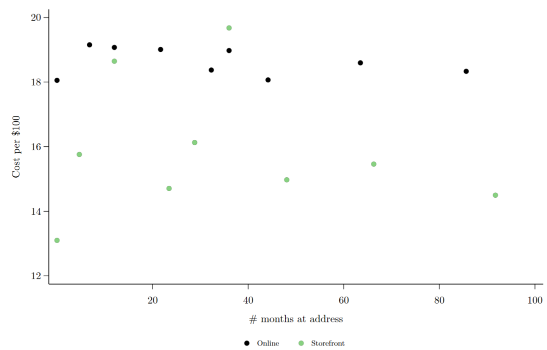


(b) Credit visible sample

Panel B: Cost per $100

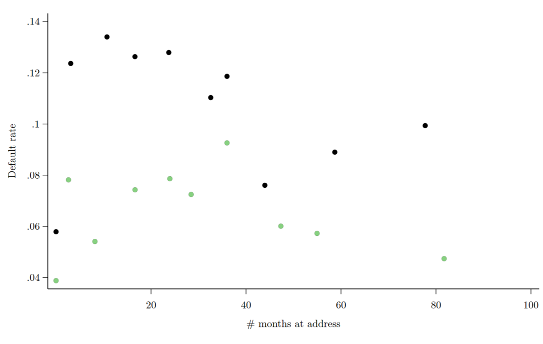


(c) Random Clarity sample

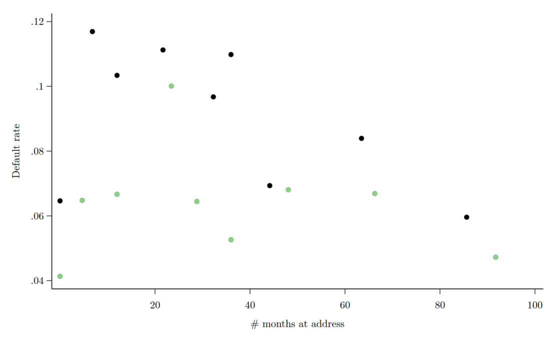


(d) Credit visible sample

Panel C: Default



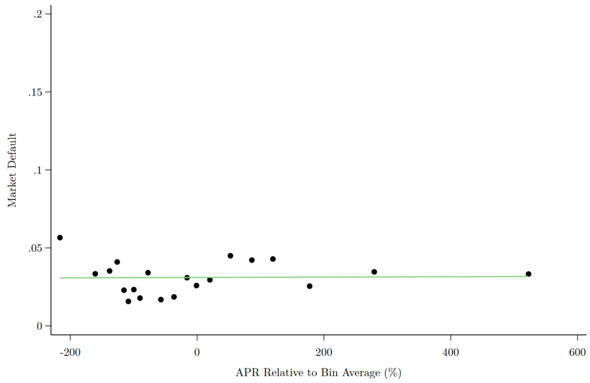
(e) Random Clarity sample



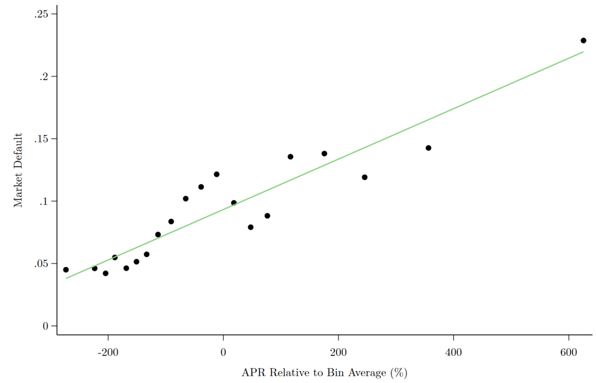
(f) Credit visible sample

Note: The figure presents binscatter plots of APR, cost per $100 borrowed, and default rates in the credit visible sample in graphs (b), (d), and (f), and in the standalone sample in graphs (a), (c), and (e). The x-axis contains the number of months living at the same address for the loan applicant, at the application date.

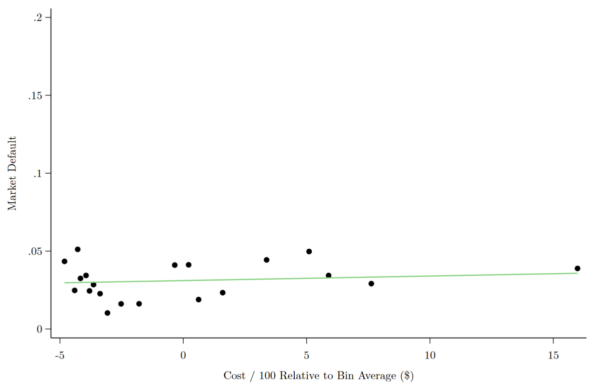
Figure A2: Prices and Default - Random Clarity Sample



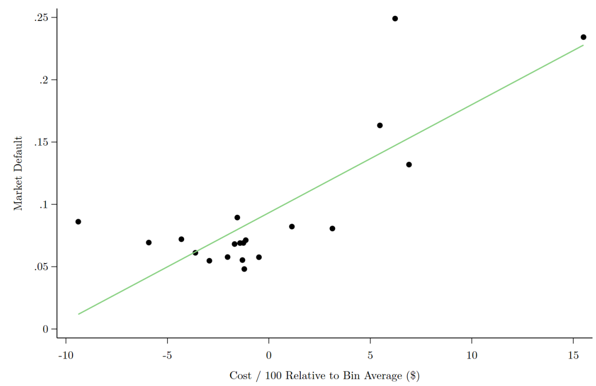
(a) APR vs Default, Storefront



(b) APR vs Default, Online



(c) Cost per $100 vs Default, Storefront



(d) Cost per $100 vs Default, Online

Note: The figure presents binscatter plots of default by bins of price relative to the bin average price in the stand-alone sample. Prices are expressed as APR in panels (a) and (b) and cost per $100 borrower in panels (c) and (d). Bins are age-income bins.

Table A1: **Online Payday Loan Premium: Excluding 2013**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Random Clarity Sample (2014-2017)								
Outcome:		APR			Cost / 100			Default	
Storefront mean:		294			12.4			0.041	
Online dummy	92.2	106.4	133.0	4.13	4.48	6.17	0.026	0.020	0.062
	(37.0)	(42.7)	(56.0)	(2.14)	(2.47)	(3.10)	(0.026)	(0.031)	(0.036)
	[0.016]	[0.016]	[0.021]	[0.060]	[0.076]	[0.052]	[0.320]	[0.528]	[0.091]
R2	0.538	0.627	0.779	0.613	0.698	0.825	0.099	0.129	0.398
N	312,298	309,186	286,945	312,298	309,186	286,945	312,298	309,186	286,945
	Panel B: Credit Visible Sample (2014-2019)								
Outcome:		APR			Cost / 100			Default	
Storefront mean:		300			13.3			0.044	
Online dummy	85.3	96.2	116.4	3.18	3.49	4.56	0.034	0.029	0.042
	(25.4)	(32.3)	(53.5)	(1.42)	(1.68)	(2.30)	(0.020)	(0.025)	(0.027)
	[0.001]	[0.005]	[0.035]	[0.029]	[0.044]	[0.053]	[0.094]	[0.265]	[0.131]
R2	0.533	0.608	0.746	0.556	0.635	0.764	0.096	0.140	0.287
N	178,953	177,197	163,898	178,953	177,197	163,898	178,953	177,197	163,898
	Panel C: Credit Visible Sample with Vantage								
Online dummy	85.4	96.2	116.5	3.19	3.50	4.57	0.034	0.029	0.042
	(25.4)	(32.3)	(53.5)	(1.42)	(1.68)	(2.30)	(0.020)	(0.025)	(0.027)
	[0.001]	[0.005]	[0.034]	[0.029]	[0.044]	[0.053]	[0.089]	[0.259]	[0.132]
R2	0.533	0.608	0.746	0.557	0.635	0.764	0.100	0.143	0.287
N	178,953	177,197	163,898	178,953	177,197	163,898	178,953	177,197	163,898
State FE	Yes	No	No	Yes	No	No	Yes	No	No
Zip FE	No	Yes	No	No	Yes	No	No	Yes	No
Consumer FE	No	No	Yes	No	No	Yes	No	No	Yes

Note: The table presents coefficient estimates of the online loan dummy from regressions of payday loan prices and default probability for the random Clarity sample in Panel A and the credit visible sample in Panels B and C, excluding loans made in 2013. All regressions include fixed effects for either state, ZIP code, or customer; fixed effects for day of week, day of month, month of year, and calendar year; and controls for deciles of loan duration, loan size, age, and income, categorical variables for housing status and pay frequency, and number of inquiries per week. Panel C additionally includes controls for decile of Vantage score. Robust standard errors clustered at the state level are in parentheses, and p-values are in brackets.

Table A2: **Online Payday Loan Premium: Non-Imputed Sample**

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Random Clarity Sample (2013-2017)					
Outcome:		APR			Cost / 100	
Storefront mean:		298			12.4	
Online dummy	98.7	111.0	119.2	3.45	3.80	4.80
	(33.7)	(40.4)	(47.9)	(1.88)	(2.27)	(2.94)
	[0.005]	[0.008]	[0.016]	[0.073]	[0.101]	[0.109]
R2	0.595	0.686	0.823	0.641	0.734	0.870
N	272,220	269,436	252,430	272,220	269,436	252,430
	Panel B: Credit Visible Sample (2013-2019)					
Outcome:		APR			Cost / 100	
Storefront mean:		300			13.0	
Online dummy	87.3	99.8	114.1	2.70	3.13	3.85
	(23.7)	(30.9)	(57.8)	(1.30)	(1.70)	(2.34)
	[0.001]	[0.002]	[0.054]	[0.043]	[0.072]	[0.107]
R2	0.593	0.671	0.801	0.602	0.692	0.837
N	149,458	148,104	138,710	149,458	148,104	138,710
	Panel C: Credit Visible Sample with Vantage					
Online dummy	87.3	99.9	114.1	2.70	3.13	3.85
	(23.7)	(30.9)	(57.8)	(1.30)	(1.70)	(2.34)
	[0.001]	[0.002]	[0.054]	[0.043]	[0.072]	[0.107]
R2	0.593	0.672	0.801	0.603	0.692	0.837
N	149,458	148,104	138,710	149,458	148,104	138,710
State FE	Yes	No	No	Yes	No	No
Zip FE	No	Yes	No	No	Yes	No
Consumer FE	No	No	Yes	No	No	Yes

Note: The table presents coefficient estimates of the online loan dummy from regressions of payday loan prices and default probability for the random Clarity sample in Panel A and the credit visible sample in Panels B and C, excluding defaulted loans and those with missing information in the pricing formula in equation (1). All regressions include fixed effects for either state, ZIP code, or customer; fixed effects for day of week, day of month, month of year, and calendar year; and controls for deciles of loan duration, loan size, age, and income, categorical variables for housing status and pay frequency, and number of inquiries per week. Panel C additionally includes controls for decile of Vantage score. Robust standard errors clustered at the state level are in parentheses, and p-values are in brackets.

Table A3: **Online Payday Loan Premium - Bin Level**

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:		APR			Cost / 100	
Online Dummy	91.3	99.4	123.6	4.8	4.0	4.0
	(1.4)	(1.9)	(1.6)	(0.0)	(0.1)	(0.0)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Binning Quintiles	Age-Income	Age-Income	Credit Score	Age-Income	Age-Income	Credit Score
Sample	Stand-Alone	Credit Visible	Credit Visible	Stand-Alone	Credit Visible	Credit Visible
R^2	0.291	0.297	0.273	0.571	0.54	0.54
N	304,404	157,400	157,400	304,404	157,400	157,400
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table presents coefficient estimates of the online loan dummy from regressions of payday loan prices for the random Clarity sample in Columns (1) and (4) and the credit visible sample in Columns (2)-(3) and (5)-(6). The unit of observation is loan, and bins are defined by quintiles of age and income in columns (1)-(2) and (4)-(5), and credit score in columns (3) and (6). All regressions include fixed effects for bin and week fixed effects and controls for averages of loan duration, loan size, age, income, number of inquiries per week, and lagged default rates and lagged percentage of late payments in the bin. Robust standard errors clustered at the market level are in parentheses, and p-values are in brackets.

Table A4: **Statewide Payday Loan Database - Time Series**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Random Clarity Sample (2013-2017)				Panel B: Credit Visible Sample (2013-2019)			
Outcome:	APR		Cost / 100		APR		Cost / 100	
Online × Database 6 or less years ago	- 66.4	- 115.3	- 7.7	- 8.6	- 127.8	- 167.3	- 5.9	- 6.7
	(105.3)	(112.0)	(4.8)	(5.1)	(140.6)	(143.8)	(4.9)	(5.3)
	[0.529]	[0.304]	[0.106]	[0.095]	[0.364]	[0.245]	[0.225]	[0.207]
Online × Database 7 - 8 years ago	- 89.7	- 67.6	- 3.8	- 3.8	- 59.2	- 45.7	- 1.0	- 0.9
	(48.4)	(38.7)	(1.9)	(1.9)	(52.9)	(37.6)	(2.2)	(1.8)
	[0.065]	[0.082]	[0.045]	[0.053]	[0.264]	[0.225]	[0.654]	[0.614]
Online × Database 9 - 10 years ago	- 161.2	- 179.1	- 6.9	- 7.7	- 123.3	- 134.5	- 6.3	- 5.9
	(37.3)	(28.3)	(2.5)	(2.3)	(30.6)	(27.7)	(1.5)	(1.3)
	[0.000]	[0.000]	[0.006]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]
Online × Database 11 - 12 years ago	- 150.0	- 118.3	- 6.8	- 5.6	- 141.1	- 111.0	- 5.9	- 4.8
	(71.0)	(55.0)	(1.4)	(1.3)	(72.1)	(60.3)	(1.6)	(1.4)
	[0.036]	[0.033]	[0.000]	[0.000]	[0.051]	[0.066]	[0.000]	[0.001]
Online × Database 13 or more years ago	- 148.3	- 136.9	- 8.7	- 6.7	- 112.0	- 95.1	- 8.2	- 6.0
	(67.1)	(59.7)	(1.5)	(1.5)	(48.2)	(46.5)	(1.5)	(1.5)
	[0.028]	[0.023]	[0.000]	[0.000]	[0.021]	[0.042]	[0.000]	[0.000]
Online	149.9	129.1	7.9	7.4	146.0	118.2	6.7	5.7
	(21.2)	(20.8)	(1.1)	(1.1)	(21.8)	(19.3)	(1.3)	(1.1)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Database 6 or less years ago	242.5	228.1	12.5	11.4	289.4	267.8	11.0	9.7
	(119.0)	(127.6)	(5.6)	(5.8)	(158.9)	(159.5)	(5.6)	(6.0)
	[0.043]	[0.075]	[0.027]	[0.052]	[0.069]	[0.094]	[0.048]	[0.107]
Database 7 - 8 years ago	45.6	27.0	2.7	2.0	51.4	39.9	1.5	0.7
	(20.9)	(22.2)	(1.0)	(1.1)	(21.4)	(21.2)	(1.3)	(0.9)
	[0.030]	[0.226]	[0.006]	[0.074]	[0.017]	[0.061]	[0.228]	[0.453]
Database 9 - 10 years ago	85.5	65.5	4.1	3.8	53.0	25.7	3.3	1.9
	(29.3)	(29.5)	(2.6)	(2.4)	(23.0)	(21.7)	(1.3)	(1.2)
	[0.004]	[0.027]	[0.112]	[0.122]	[0.022]	[0.238]	[0.014]	[0.103]
Database 11 - 12 years ago	74.9	62.8	4.2	3.0	55.8	40.2	2.7	1.7
	(61.8)	(48.2)	(1.0)	(0.9)	(64.8)	(54.4)	(1.4)	(1.1)
	[0.226]	[0.193]	[0.000]	[0.002]	[0.390]	[0.460]	[0.052]	[0.109]
Database 13 or more years ago	32.5	59.9	4.4	3.1	- 2.5	7.7	3.7	1.7
	(48.0)	(39.3)	(1.1)	(0.7)	(38.1)	(36.4)	(1.3)	(1.0)
	[0.499]	[0.129]	[0.000]	[0.000]	[0.948]	[0.832]	[0.005]	[0.090]
Controls	No	Yes	No	Yes	No	Yes	No	Yes
FE	No	Yes	No	Yes	No	Yes	No	Yes
R2	0.111	0.465	0.330	0.401	0.111	0.454	0.256	0.346
N	336,690	336,690	336,690	336,690	188,913	188,913	188,913	188,913

Note: The table presents coefficient estimates of loan-level regressions of prices, in APR and cost per $100 borrowed, on a dummy variable that indicates whether the loan is an online loan, and dummy variables for each bin of time since implementation of a statewide payday loan database, and their interaction. Columns (2), (4), (6), and (8) contain loan and customer level controls, and time fixed effects. The samples used are in Panel A and the random Clarity sample, in Panel B the credit visible sample. Robust standard errors clustered at the state-year level are in parentheses, and p-values are in brackets.