The Online Payday Loan Premium*

Filipe Correia, University of Georgia[†]

Peter Han, UIUC[‡]

Jialan Wang, UIUC and NBER[§]

PRELIMINARY AND INCOMPLETE

January 2022

Abstract

Using data from a subprime credit bureau with nationwide coverage in the United States, we investigate the potential for online technology to lower fixed costs and increase lending efficiency in the expensive payday loan market. We find that prices for online loans are about 100% APR higher than storefront loans. This premium is not explained by loan or customer characteristics, differences in pricing models, or traditional measures of credit risk. At least part of the online payday loan premium seems to be due to default rates are that are double for that for storefront loans. Customers with both types of loans are much more likely to default on online loans.

Keywords: Online Lending, Payday Loans, Fintech

*We are grateful to Julia Fonseca, George Pennacchi, Heitor Almeida, Rustom Irani, and Sergio Vicente, Charles Kahn, Alexei Tchistyi, Tatyana Deryugina, Timothy Johnson, Mathias Kronlund, Elizaveta Sizova and participants at the Gies College of Business brownbag, the 28th Finance Forum, and the 7th International Young Scholars Conference for valuable suggestions. Renhao Jiang, Peichen Li, Yunrong Zhou, Hejia Liu, Kevin Tzeng, Kashish Mehta, Raja Gajula, Matthew Boyd, Yiheng Yang, and Kunal Bagali provided valuable research assistance.

[†]Terry College of Business, B339 Amos Hall, Athens, GA 30605. filipe.correia@uga.edu

[‡]Gies College of Business, 1206 South 6th Street, Champaign, IL 61820. weitong2@illinois.edu

[§]Gies College of Business, 1206 South 6th Street, Champaign, IL 61820. jialanw@illinois.edu

I Introduction

Payday loans have been a controversial credit product since gaining popularity in the 1990s, in part due to high prices that generally range 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, it also faces high fixed costs of operation and high default rates (Ernst & Young, 2009). Given the rise of online lending in many consumer credit markets, financial technology has the potential to lower prices and increase efficiency by reducing fixed costs and improving default prediction.

This paper presents some of the first evidence on price differences between online and storefront payday loans in the United States, and shows that instead of reducing prices, online payday lenders charge a significant premium compared with storefront lenders. In two independent samples covering 2013 through 2019 including customers both with and without traditional credit reports, online payday loans were about 136% APR more expensive on average than storefront loans, despite online borrowers reporting higher income, greater home ownership, and similar credit scores. This premium remains similar in magnitude when controlling for observable loan and customer characteristics, including traditional credit scores.

A significant difference in default risk conditional on consumer and loan characteristics seems to explain at least part of the online payday loan premium. Online loans are twice as likely to default as storefront loans, and this gap remains relatively constant at every level of consumer income and credit score. Customers who have both types of loans are significantly more likely to default on online loans.

II Background on the Payday Loan Market

Payday loans are a controversial source of short term credit among low- to middle-income Americans. Between 2015 and 2019, about 2 percent of households reporting 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 paydates (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.[1] The online payday industry grew from a small share of loans to 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). In 2011, the Consumer Financial Protection Bureau became the industry's first federal regulator. The CFPB issued rules governing the payday industry in 2017 which were largely rescinded in

[1] See Mann and Hawkins (2007) for more information on the "rent-a-bank" model.

2020, so regulation still largely falls on the states.[2]

As of 2015, traditional storefront lending was effectively banned in about 15 states. State payday laws are complex, and jurisdiction over online lending remains contested in the courts, although many state and federal regulators have moved to enforce laws that restrict online loans in states that also regulate storefront lending (King and Standaert, 2013; of America, 2010). In addition to regulatory considerations, other features that differ between the online and storefront payday loan markets include lead generators, which are intermediaries connecting consumers and lenders, payment and collection mechanisms that involve Automated Clearing House (ACH) transactions and bank account access instead of post-dated checks and in-person payment, and the online advertising market (Trusts, 2014).

III Data

Our data on storefront and online payday loans come from Clarity, an alternative credit bureau and subsidiary of Experian, one of the three major credit reporting agencies. Previous research using Clarity data includes Fonseca (2021), Di Maggio, Ma and Williams (2020), Miller and Soo (2020), and Miller and Soo (2021). Blattner and Nelson (2020) use similar data from FactorTrust, another alternative credit bureau. Clarity specializes in collecting 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 nonprime consumers in the United States. Like other credit reporting agencies, Clarity relies on voluntary 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 one of the best sources of nationwide subprime credit activity.

The Clarity database contains information on a variety of subprime credit products

[2]See https://www.consumerfinance.gov/payday-rule/

including payday, rent-to-own, installment, auto, and auto title loans. We focus only 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 the the type of loan applied for and basic self-reported demographics including zipcode 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 may not be verified prior to 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 represents about 1.5% of the full database. The sample of borrowers includes those who apply for payday loans as well as other products, and only a subset of applications result in originated loans, which we use in our main analysis. Out of 1 million unique borrowers who submitted an inquiry for any type of subprime credit, 366,327 inquired for either an online or storefront payday loan, and 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 who are 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 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, so 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:[3]

$$\begin{aligned} APR &= \frac{365}{LoanMaturity} \times \frac{Repayment - LoanAmount}{LoanAmount} \\ Cost\_per\_100 &= 100 \times \frac{Repayment - LoanAmount}{LoanAmount} \end{aligned} \tag{1}$$

Because payday loans have fairly simple and standardized structures, we feel that 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 that are 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 we show in the next section, risk-based pricing is very limited in both the online and storefront payday markets, so we do not think our results are driven by this potential source of measurement error. In order 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, zipcode, and terciles of loan and borrower characteristics. We try to match defaulted loans to non-defaulted ones in cells of decreasing granularity until all loans are matched (e.g. zipcode is matched first, and if no available priced loans are matched by zipcode, then state-level matches are used). In the analysis below, we will 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 analysis to reduce the effect

[3] See DeYoung, Phillips et al. (2006) and DeYoung and Phillips (2006)

of outliers.

IV Descriptive Results

In this section, we report summary statistics for our sample and provide descriptive evidence on pricing in the storefront and online payday loan markets.

IV.A Summary statistics and external validity

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 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, while 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).

Turning back to Table 1, the random Clarity sample in Panel A 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 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 both the historical evolution of Clarity's client base and the greater use of reporting and verification systems by online lenders to mitigate fraud risk. Thus, while our data do not cover the majority of the payday loan market, it covers a significant fraction and as described below, basic loan characteristics are broadly consistent with those from previous studies.

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 Charitable Trusts 2014, Wang and Burke 2021). While loan and customer characteristics are fairly comparable to previous studies using storefront payday data (e.g. Skiba and Tobacman 2008, Wang and Burke 2021), 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 measure default as any loan that was not paid in full, as reported to Clarity by lenders. We do not attempt to distinguish between delinquency and default, track the ultimate recovery rate of defaulted loans, or account for reporting error or reporting lags (e.g. lenders failing to report defaults to Clarity). Nonetheless, 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) reports a 4% charge-off rate and Wang and Burke (2021) 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.

The 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. Online loans are significantly smaller in size, 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 (554 vs. 561). 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, 2020). The higher income, lower age, and higher default risk associated with online payday loans are consistent with previous survey evidence (Trusts, 2014).

Based on the pricing formulas and imputation algorithm described in Section III, we find average APRs of 385% in the random Clarity sample (Panel A) and 401% in the credit visible sample (Panel B). Average cost per $100 is $17.0 in the random Clarity sample and $17.5 in the credit visible sample. Despite the assumptions needed to calculate prices in the Clarity data, these estimates are very 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, which is one of the most expensive lending markets. Wang and Burke (2021) 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, prices for storefront loans range from 297% to 305% APR and $12.4 to $13.4 per $100 in our Clarity samples.

Price statistics for online payday loans are rarer, and we are unaware of previous academic studies of 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 434% - 441% and cost per $100 of $19.2 - $19.5 in the Clarity samples. Despite Clarity's substantial market share of online payday loans, it is possible that less-compliant or more predatory lenders who may charge higher prices and engage in other unfriendly practices toward consumers are 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.

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. 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 that is not possible 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 and the fact that relatively few loans realize default even in higher-default groups, allowing us to reliably use non-defaulted loans to impute prices for defaulted loans. Overall, we feel that 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.

IV.B Descriptive results

To further explore pricing differences between the online and storefront payday markets, Figure 2 plots 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, suggesting that there are few systematic pricing differences that depend on whether a customer has a traditional credit report.

Consistent with pricing schedules that are based on integer values of cost per $100, the distributions in graphs (b) and (d) exhibit several modes for both online and storefront loans, with common support across these distributions but higher price points 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.

Some potential mechanisms for why prices are higher for online loans are that the customer base is inherently different, that online lenders use different pricing functions, and that customers exhibit different default rates depending on lender type. To shed initial light on these mechanisms, Figures 3 through 6 present unconditional binscatters of how prices and default rates change depending on loan duration, customer income, months at address, and Vantage score.

Figure 3 shows binscatters of prices and default risk by loan duration, which is 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 very flat and does not vary monotonically with loan duration, although 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. Again, APRs are higher for online loans conditional on loan duration, although this disparity is greater in absolute terms for lower loan durations. Default risk is slightly negatively correlated with loan duration only for online loans, but there is generally little relationship.

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 extremely flat for both online and storefront loans by 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. This provides initial evidence that online lenders do not seem to employ more sophisticated risk-based pricing algorithms despite their greater use of credit reporting agencies such as Clarity, greater price dispersion and more continuous pricing functions, and the high overall levels of credit risk that could make underwriting technology particularly valuable in this market. Another potential demographic that could be a driver of credit risk is the number of months a consumer has lived at their current address, shown in Figure 5, but we find limited evidence of a relationship with either default or prices, possibly due to noise in this self-reported measure.

Finally, Figure 6 presents binscatters of prices and default risk by Vantage score, one of the most widely-used consumer credit scores that 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 even conditional on taking out subprime credit, Vantage score is significantly predictive of default risk for both storefront and online payday loans. However, as with income, online payday loans have significantly default credit risk at every level of Vantage. Despite its strong correlation with credit risk, both APR and cost per $100 are only weakly correlated with Vantage score in both the online and storefront markets, again consistent with a general lack of risk-based pricing.

V Regression Analysis

In this section, we further disentangle the drivers of the online payday loan premium using regression analysis. To test whether the unconditional pricing differences are driven by differences in loan or customer characteristics, we implement the following regression model:

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

where Y_{ist} is a price or default outcome for a given loan from customer i living in state or zipcode s originated at time t. All regressions include fixed effects α_{is} for either state, zipcode, or customer; fixed effects α_t for day of week, day of month, month of year, and calendar year; and controls X_{ist} for 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 also include a dummy variable for online loans with the coefficient of interest β. Standard errors are clustered at the state level for all specifications.

Table 2 presents our main results. The table includes three columns for each of the three key outcome variables: APR, cost per $100, and default rate. The three different specifications per outcome variable include either state fixed effects, zipcode 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, with and without the inclusion of Vantage score. As shown in column (1), when state fixed effects are included, the APR premium is between 93% to 98% across samples and models. Surprisingly, the premium increases when zipcode or customer fixed effects are included instead of state fixed effects. The coefficient on the online loan dummy ranges from 104% to 110% APR when including zipcode fixed effects, and from 131% to 141% when including customer fixed effects. As shown in columns (4) through (6), the online payday loan premium is between \$3.5 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 136-137% APR and \$5.8 to \$7.1 cost per \$100, which are within the range of the regression estimates. Overall, these results show that the online payday loan premium is not driven by differences in observable loan or customer characteristics between the two loan types.

Finally, as shown in columns (7) through (9), default risk is between 2.6% and 8.4% 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% 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 loan compared with a storefront one, so differences in default risk are not purely driven by time-invariant customer characteristics.

Consistent with the lack of risk-based pricing, the inclusion of controls for vantage score makes almost no difference in the estimated price premia, as shown in Panel C. Interestingly, although Vantage score is correlated with default risk, its inclusion has no effect on 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 sample size and precision, 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. While other factors such as differences in fixed costs, the lead generation system, and advertising and customer acquisition costs are likely to drive some of these differences, and may work in conflicting directions, we show that the much higher default rates of online loans – which are even present within the same customer – are likely to be part of the explanation.

The higher default rates for online loans remain a puzzle we will explore in further research, since online borrowers appear to have higher income and other characteristics consistent with lower or similar levels of credit risk. The lack of in-person interaction and higher potential for fraud and identity theft could be part of the cause of higher default rates and higher prices. The higher prices themselves could also select for customers with higher unobservable risk or cause these loans to be lower on the repayment hierarchy for a given customer. The underlying causes of higher default rates are also complicated by differences in collection mechanisms, which are generally thought to be more aggressive for online loans since lenders have direct access to consumers' bank accounts through the ACH network (Bureau, 2016).

VI Conclusion

This paper presents novel evidence on the 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 are more expensive by around 100% APR even conditioning on loan and customer characteristics. Although neither storefront nor online payday loans seem to employ significant degrees of risk-based pricing, available measures of default risk also don't explain 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. Thus, while differences in consumer or loan characteristics do not seem to explain the online loan premium, inherent differences in default risk is likely to explain at least part of this price disparity.

References

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

Blattner, Laura and Scott Nelson, "How Costly is Noise? Data and Disparities in the US Mortgage Market," Technical Report, Working Paper. 2020.

Bureau, Consumer Financial Protection, "Online Payday Loan Payments," Technical Report, Consumer Financial Protection Bureau 2016.

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

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

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

Fonseca, Julia, "Access to credit and financial health: Evaluating the impact of debt collection," 2021.

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.

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.

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

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

Maggio, Marco Di, Angela T Ma, and Emily Williams, "In the Red: Overdrafts, Payday Lending and the Underbanked," Technical Report, National Bureau of Economic Research 2020.

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

Miller, Sarah and Cindy K Soo, "Does increasing access to formal credit reduce payday borrowing?," Technical Report, National Bureau of Economic Research 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.

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

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

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.

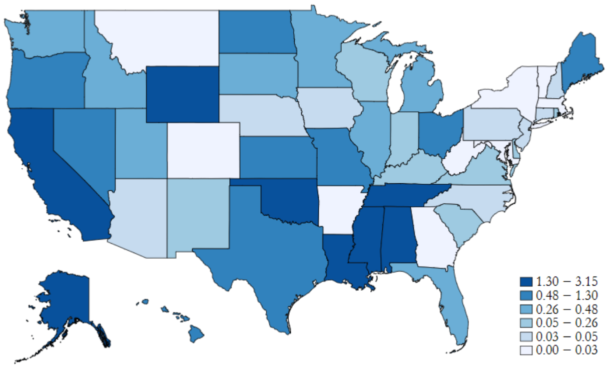
Trusts, Pew Charitable, "Payday Lending in America: Who Borrows, Where They Borrow, and Why," *Washington, DC*, 2012.

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

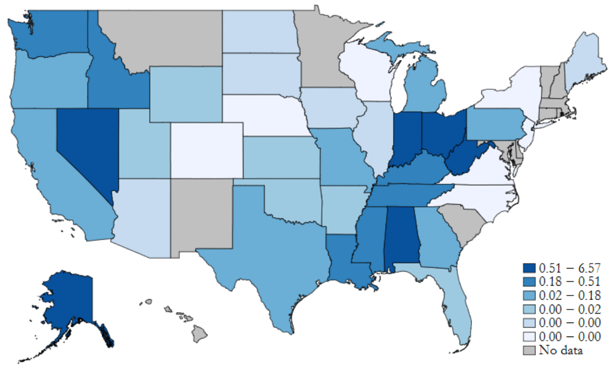
Wang, Jialan and Kathleen Burke, "The effects of disclosure and enforcement on payday lending in Texas," *Journal of Financial Economics*, 2021.

Figure 1: Geographic Distribution of Loans per Capita

Panel A: Random Clarity sample

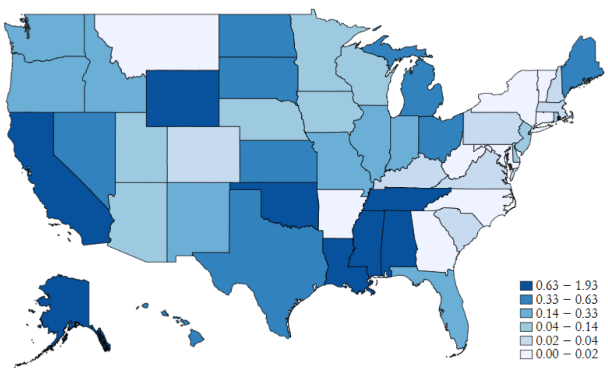


(a) Online

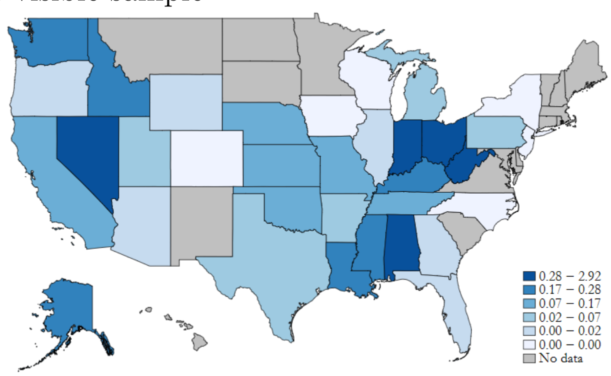


(b) Storefront

Panel B: Credit visible sample



(c) Online

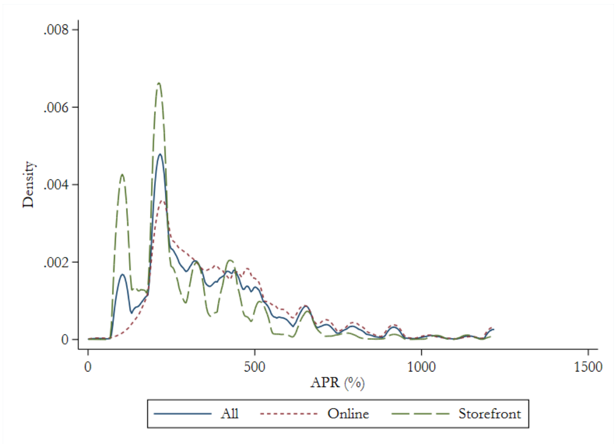


(d) Storefront

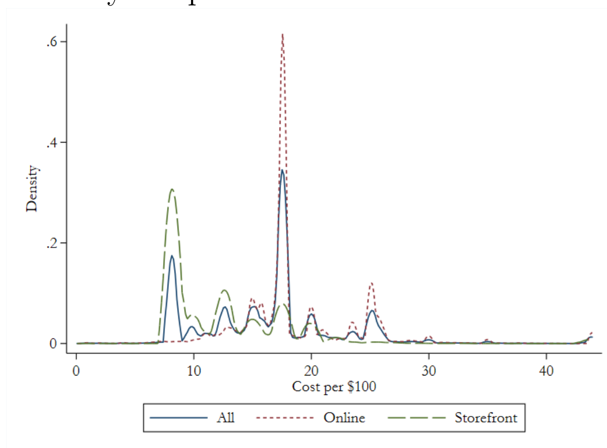
Note: In this figure we map the geographical distribution of payday loans by state. Graphs show total numbers of online and storefront loans per 1000 people in the state population based on the 2010 Census. The random Clarity sample presented in Panel A consists of a random sample of 1 million unique 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

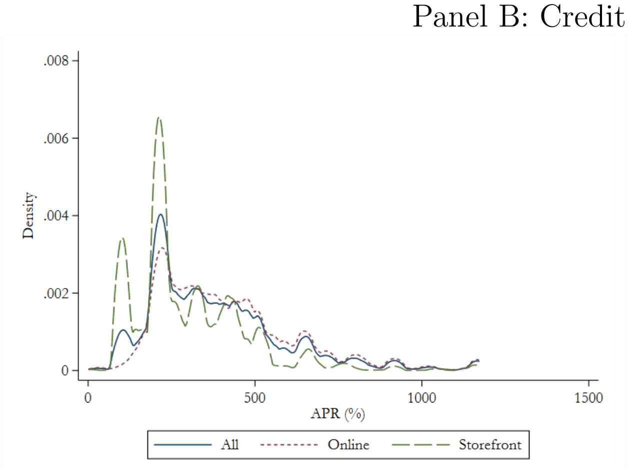


(a) APR

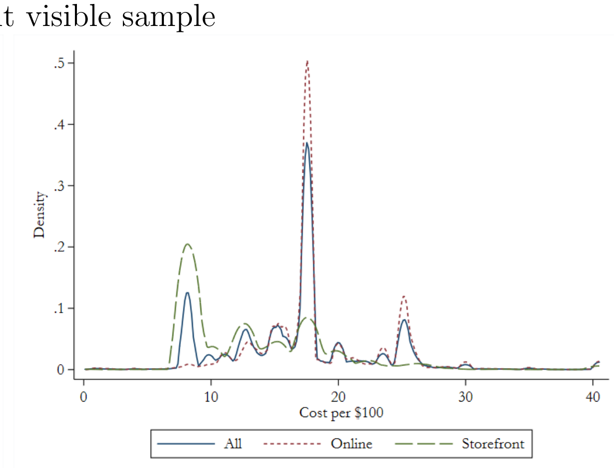


(b) Cost per $100

Panel B: Credit visible sample



(c) APR

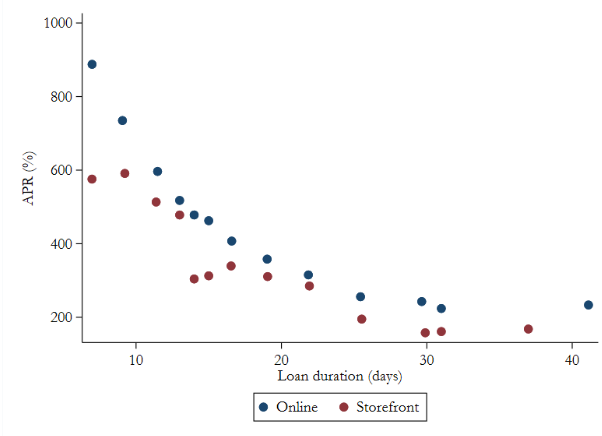


(d) Cost per $100

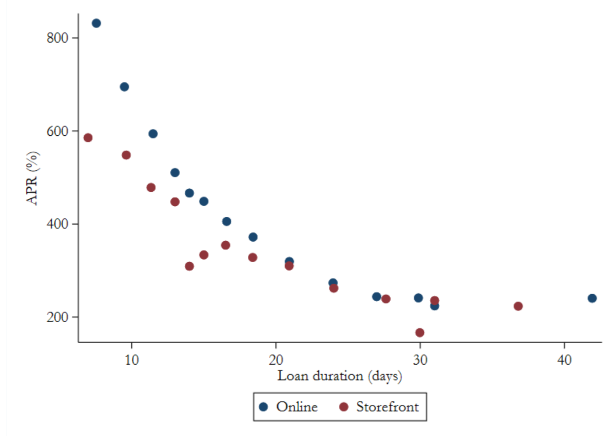
Note: This figure shows the distributions of prices for online and storefront payday loans. The random Clarity sample presented in Panel A consists of a random sample of 1 million unique borrowers that submitted loan inquiries in Clarity's full database between 2013 and 17. Only originated payday loans 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

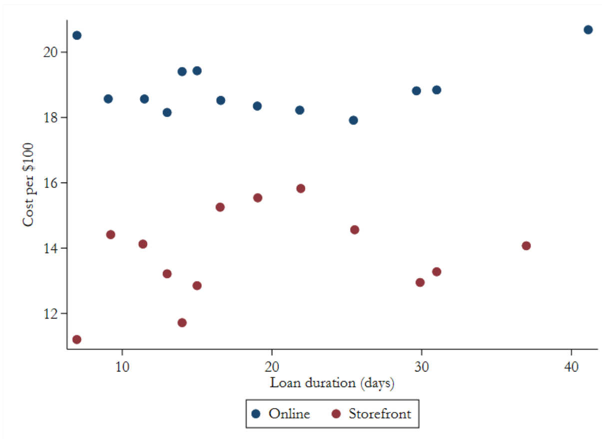


(a) Random Clarity sample

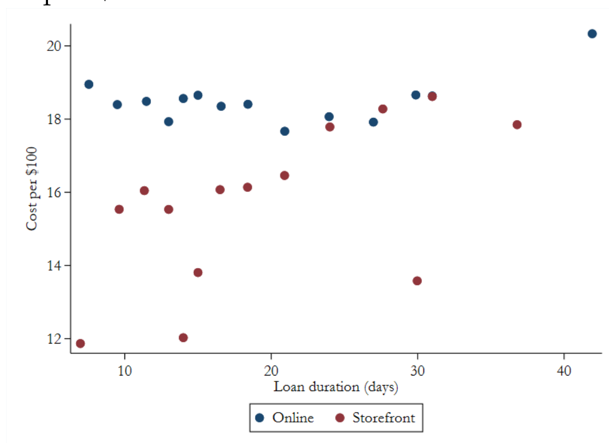


(b) Credit visible sample

Panel B: Cost per $100

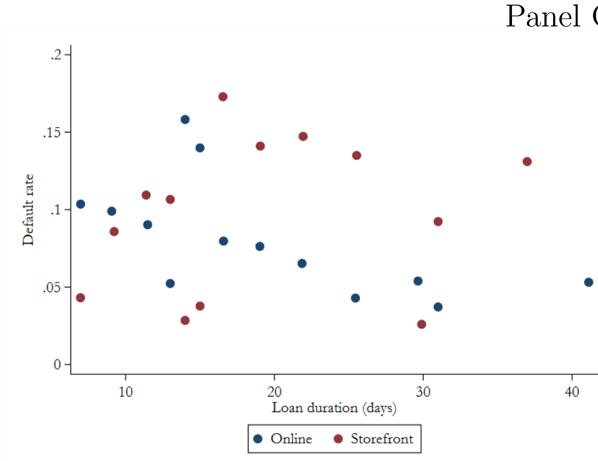


(c) Random Clarity sample

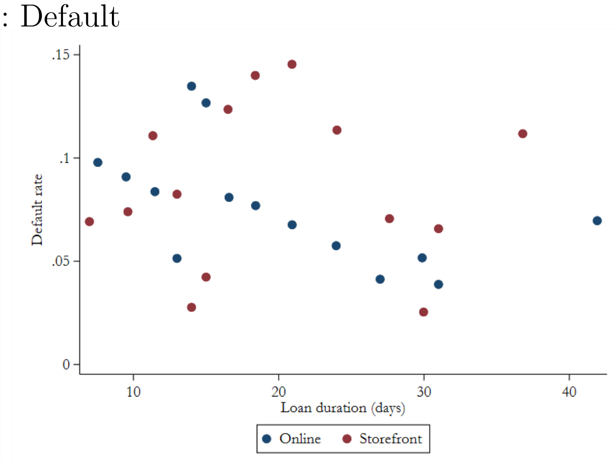


(d) Credit visible sample

Panel C: Default



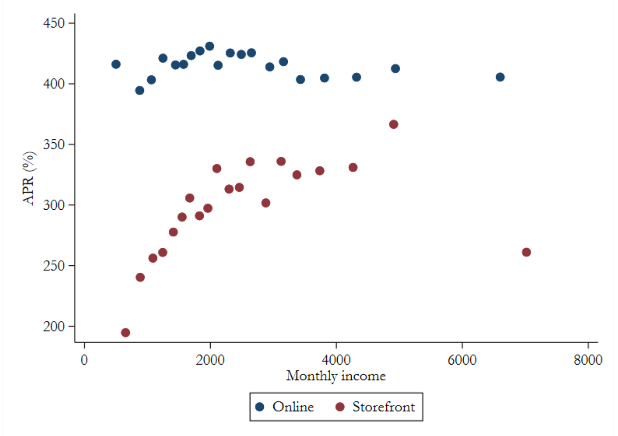
(e) Random Clarity sample



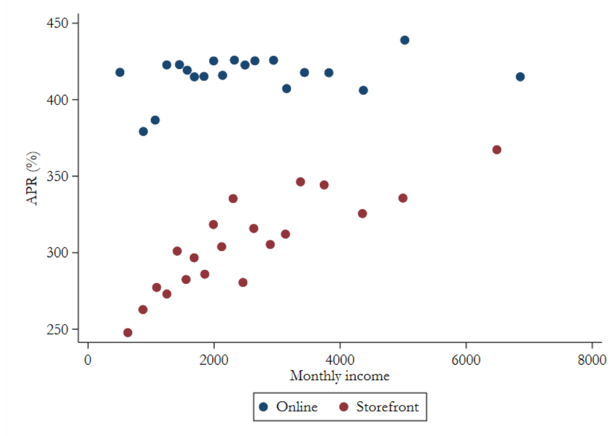
(f) Credit visible sample

Figure 4: Prices and Default Rates by Borrower Income

Panel A: APR

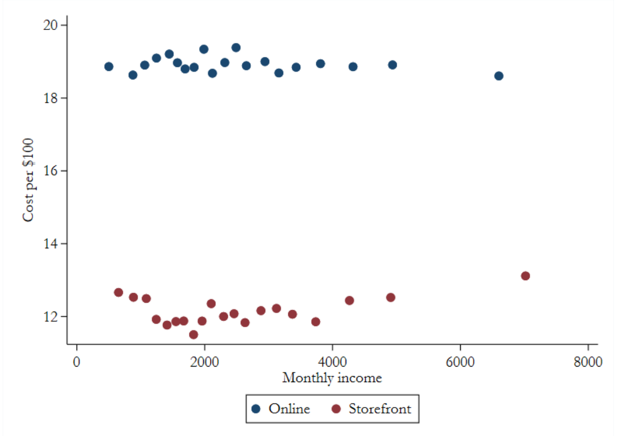


(a) Random Clarity sample

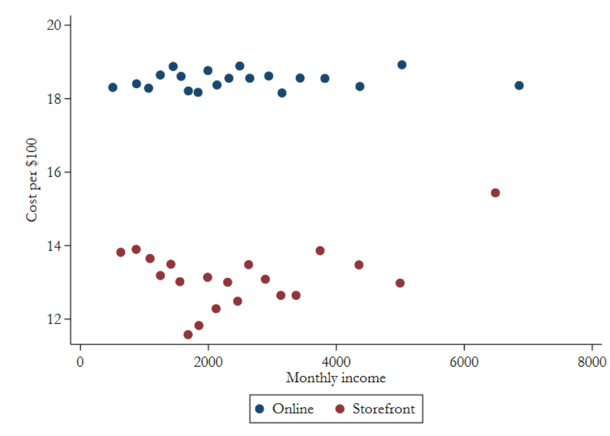


(b) Credit visible sample

Panel B: Cost per $100

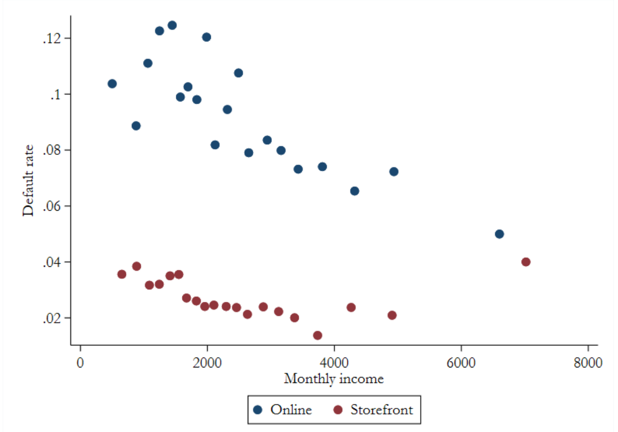


(c) Random Clarity sample

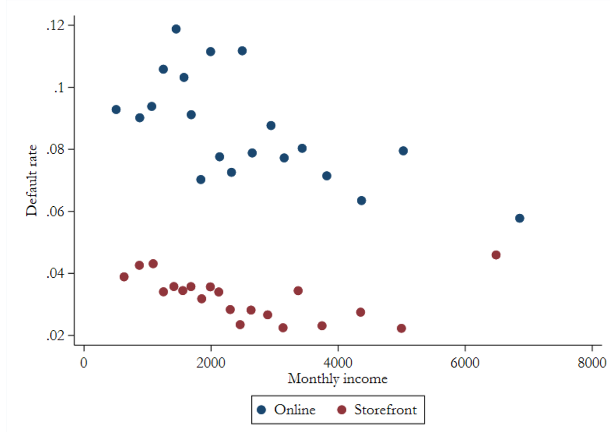


(d) Credit visible sample

Panel C: Default



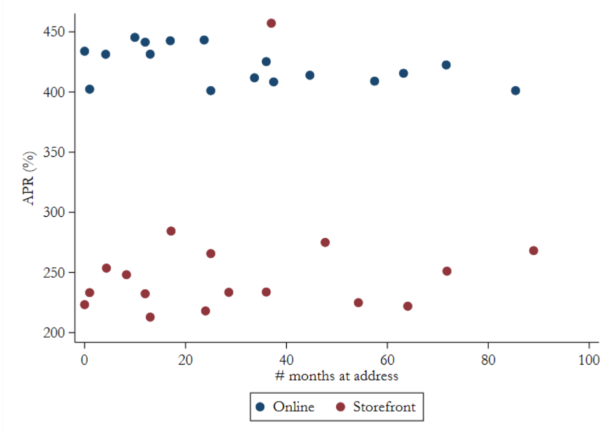
(e) Random Clarity sample



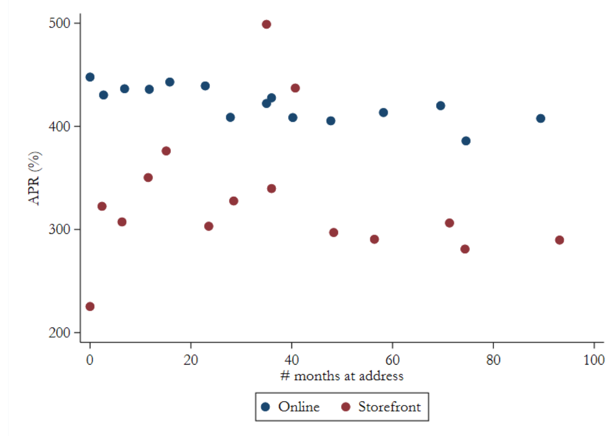
(f) Credit visible sample

Figure 5: Prices and Default Rates by Months at Address

Panel A: APR

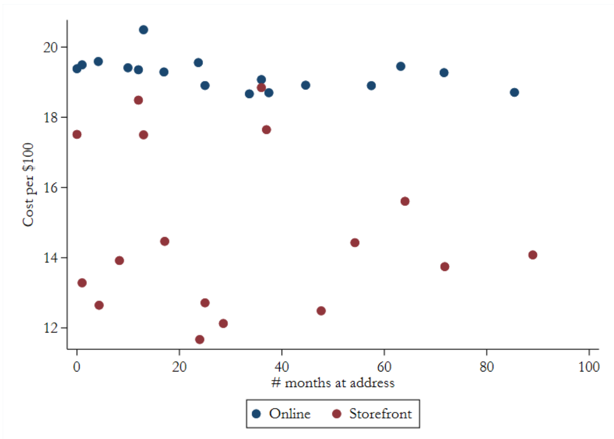


(a) Random Clarity sample

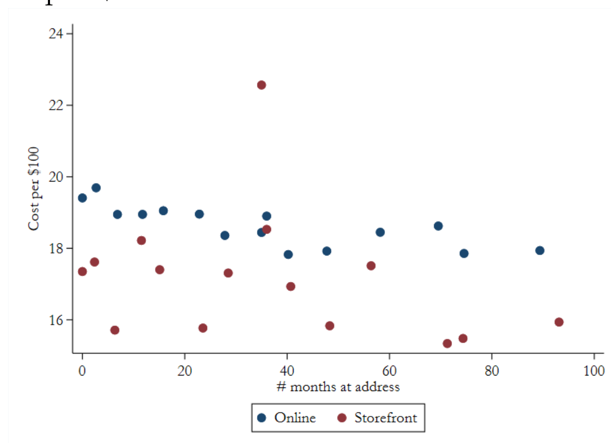


(b) Credit visible sample

Panel B: Cost per $100

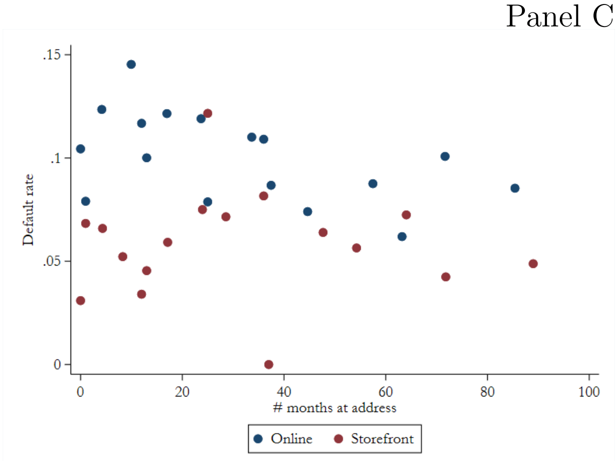


(c) Random Clarity sample

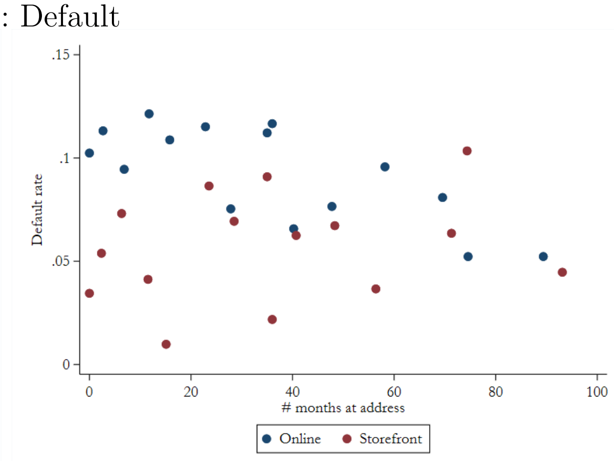


(d) Credit visible sample

Panel C: Default

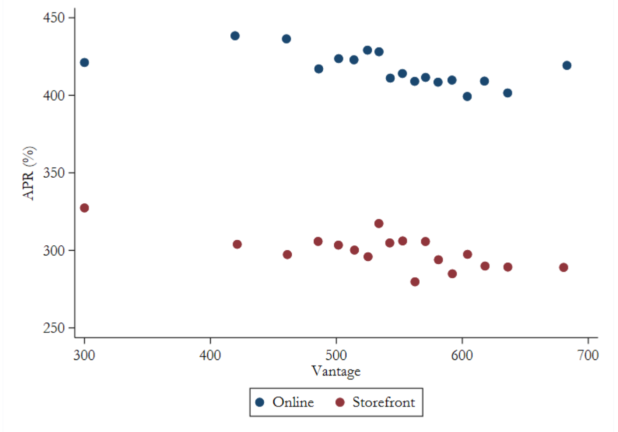


(e) Random Clarity sample

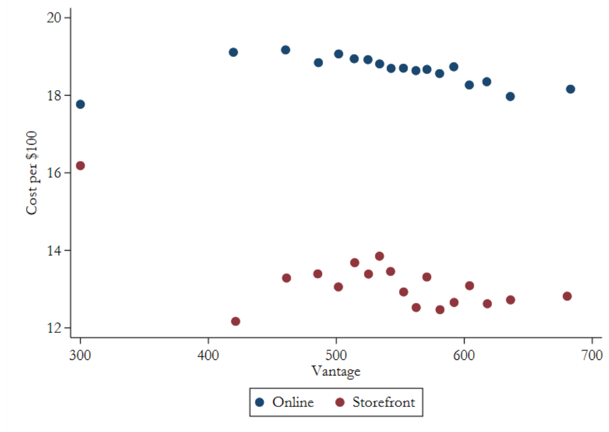


(f) Credit visible sample

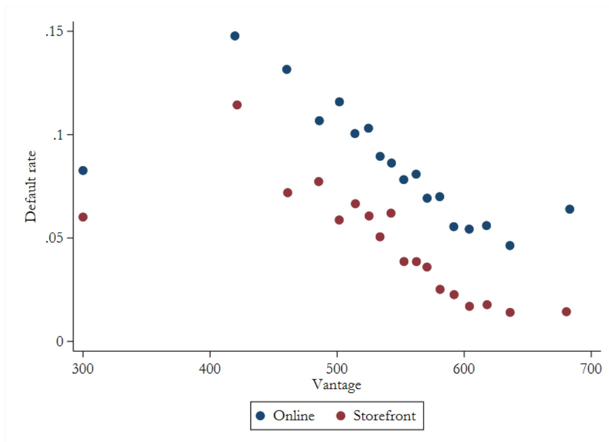
Figure 6: Prices and Default Rates by Vantage Score



(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.

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 rate	7%	0%	26%	0%	9%	4%
APR	385%	322%	329%	368%	434%	297%
Cost per $100 ($)	17.0	17.3	11.6	16.2	19.5	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 rate	7%	0%	26%	0%	8%	4%
APR	401%	336%	1240%	379%	441%	305%
Cost per $100 ($)	17.5	17.5	22.6	16.6	19.2	13.4
Vantage score	556	555	61	559	554	561
Unscoreable	18%	0%	38%	18%	21%	10%
Online Loan	70%	100%	46%	68%	100%	0%
Self-Reported Information						
Owns Home	19%	0%	39%	17%	24%	7%
Age	42.0	41.0	13.5	42.3	40.1	46.4
Months at Address	30.8	24.0	24.5	30.4	31.2	27.2
Net Monthly Income	2558	2204	1520	2553	2822	1920
# 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]
R2	0.534	0.615	0.775	0.582	0.663	0.812	0.138	0.172	0.448
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	93.2	104.4	130.6	3.51	3.84	5.10	0.040	0.036	0.055
	(25.6)	(33.8)	(55.7)	(1.48)	(1.78)	(2.41)	(0.019)	(0.025)	(0.032)
	[0.001]	[0.003]	[0.023]	[0.022]	[0.036]	[0.039]	[0.038]	[0.155]	[0.089]
R2	0.531	0.604	0.743	0.541	0.621	0.756	0.120	0.166	0.319
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	93.2	104.4	130.6	3.52	3.85	5.10	0.040	0.036	0.055
	(25.6)	(33.8)	(55.6)	(1.48)	(1.78)	(2.41)	(0.019)	(0.025)	(0.032)
	[0.001]	[0.003]	[0.023]	[0.022]	[0.036]	[0.039]	[0.037]	[0.151]	[0.089]
R2	0.532	0.604	0.743	0.541	0.621	0.756	0.124	0.170	0.319
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 and default 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, zipcode, 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 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:		293			12.3			0.041	
Online dummy	90.9	104.9	130.9	4.04	4.39	6.15	0.026	0.019	0.062
	(36.8)	(42.6)	(55.3)	(2.17)	(2.50)	(3.12)	(0.026)	(0.031)	(0.036)
	[0.017]	[0.017]	[0.022]	[0.068]	[0.085]	[0.054]	[0.326]	[0.533]	[0.092]
R2	0.547	0.645	0.805	0.635	0.710	0.844	0.099	0.160	0.463
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:		299			13.2			0.044	
Online dummy	83.0	93.8	114.3	3.11	3.43	4.50	0.034	0.029	0.042
	(24.9)	(31.8)	(52.6)	(1.42)	(1.69)	(2.32)	(0.020)	(0.025)	(0.027)
	[0.002]	[0.005]	[0.035]	[0.033]	[0.048]	[0.058]	[0.093]	[0.263]	[0.128]
R2	0.540	0.627	0.774	0.574	0.656	0.791	0.096	0.173	0.361
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	83.1	93.8	114.4	3.11	3.43	4.50	0.035	0.029	0.042
	(24.9)	(31.8)	(52.6)	(1.42)	(1.69)	(2.32)	(0.020)	(0.025)	(0.027)
	[0.002]	[0.005]	[0.035]	[0.033]	[0.048]	[0.058]	[0.088]	[0.257]	[0.129]
R2	0.540	0.627	0.774	0.574	0.656	0.791	0.100	0.176	0.361
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, zipcode, 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	85.5	98.3	113.5	2.76	3.22	3.94
	(24.3)	(32.1)	(59.6)	(1.34)	(1.74)	(2.39)
	[0.001]	[0.004]	[0.064]	[0.045]	[0.071]	[0.106]
R2	0.592	0.671	0.801	0.600	0.691	0.837
N	149,458	148,104	138,710	149,458	148,104	138,710
	Panel C: Credit Visible Sample with Vantage					
Online dummy	85.5	98.3	113.5	2.76	3.22	3.94
	(24.3)	(32.1)	(59.6)	(1.34)	(1.74)	(2.39)
	[0.001]	[0.004]	[0.063]	[0.045]	[0.071]	[0.106]
R2	0.592	0.671	0.801	0.600	0.691	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, zipcode, 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.