Digital Credit Delivery Speed and Repayment Rates[§]

Alfredo Burlando[¶] Michael A. Kuhn[‖] and Silvia Prina[**]

Abstract

Online banking and automated credit scoring technologies are speeding up the delivery of credit. While "digital credit" broadens market access and reduces frictions, it usually features high interest rates, stiff penalties for delinquency, and high default rates. We study the impact of digital credit delivery speed on loan outcomes using data from a lender in Mexico. With a regression-discontinuity design, we estimate the causal impact of delivery speed. Doubling the delivery time from ten to twenty hours reduces the default rate by 20%. Repayment timing data show that the effect is not driven by borrowers simply returning their loans.

JEL Classifications: D14, D18, G51, O16

Keywords: Digital credit, waiting periods, financial access

[§]We are grateful to Carlos Chiapa, Bilge Erten, Benjamin Hansen, Glen Waddell, Dilip Mookherjee, Bob Triest, Judy Chevalier, Kevin Lang, Diego Ubfal, and seminar participants at Bocconi University and University of Padua for helpful comments. We thank our digital lender partner for providing us with the data and answering our many questions. The authors thank the Digital Credit Observatory for generous research support. Any errors are our own. The dataset analyzed in the current study is not publicly available because it is proprietary to our lender partner, but is available from the corresponding author on reasonable request. All code is available on request.

[¶]University of Oregon and CEGA. Department of Economics, 1285 University of Oregon, Eugene, OR 97403. E-mail: burlando@uoregon.edu

[‖]University of Oregon and CEGA. Department of Economics, 1285 University of Oregon, Eugene, OR 97403. E-mail: mkuhn@uoregon.edu

[**]Northeastern University, Department of Economics, 310A Lake Hall, 360 Huntington Avenue, Boston, MA 02115, United States. E-mail: s.prina@northeastern.edu.

1 Introduction

The digital credit market has recently emerged as source of fast, automated, remotely-provided, short-term loans for millions of people in low- and middle-income countries (Francis et al., 2017). Data harvesting and analytics have enabled digital credit providers to assess consumer credit-worthiness and ability to repay without requiring any collateral to secure loans (Björkegren and Grissen, 2018). Given its characteristics, digital credit has the potential to help households cope with unexpected shocks and reduce liquidity constraints for investments (e.g., Karlan and Zinman, 2010; Morse, 2011); Bharadwaj et al. (2019) find that digital credit in Kenya has improved household resilience to negative shocks. Furthermore, the fast speed of loan provision allows borrowers to act on time-sensitive opportunities to a much greater degree than in the past.

The quick provision of digital credit, however, begets a host of consumer protection issues (Izaguirre et al., 2018; Lang et al., 2019). It can exacerbate self-control problems for borrowers, causing over-indebtedness and default (Skiba and Tobacman, 2019), making it harder to pay bills (Melzer, 2011), and reducing access to future loans. Evidence from the credit card market shows that less-sophisticated borrowers may be susceptible to over-borrowing, penalties, and back-loading repayments, suffering large welfare losses as a result (Meier and Sprenger, 2010; Heidhues and Kőszegi, 2010). This may be especially true for digital credit, where anecdotal evidence shows that borrowers do not fully understand the terms of their loans (e.g., Mazer and Fiorillo, 2015; McKee et al., 2015) and may use them to finance unproductive, time-sensitive investment and consumption opportunities like gambling (Malingha, 2019).

Rapid delivery speed is one of the hallmarks of this new, large, and rapidly expanding market. Yet, we do not know how this feature of the product will affect borrowers. This paper's goal is to begin answering that question. We provide evidence tying the speed of delivery of digital loans to the likelihood that loans are repaid. We rely on quasi-experimental variation in the delay between a successful loan application and a loan disbursement to

measure the causal impact of a longer wait on repayment.

Our data consist of loan records from the full set of approved clients from a digital lender operating in Mexico over a seven-month period in 2018-2019. The data include both the loan application timestamp and disbursement timestamp. We take advantage of the fact that the company disbursed loans in batches, and these batched disbursements occurred only two, three, or four times during the day. Hence, processing times vary from loan to loan: the first few applications added in a new batch had to wait longer than those added last. Our empirical strategy identifies those discontinuous changes in processing times that are created each time an existing batch is disbursed and a new one is opened. Cruciably, disbursement times are ex-ante unknown to borrowers, and they change day-to-day. Thus, there is no concern that clients can time their applications for faster service. However, unlike the standard regression discontinuity (RD) setup, we also do not observe the precise moment a batch is closed; we construct proxies for these cutoff times using a machine-learning technique applied to our disbursement and application submission time data.

On average for all borrowers, loans submitted just after one of these proxied cutoffs face an additional delay of 9.81 hours, roughly doubling the total amount of time it takes to get a loan. We find that the delay induced by missing a batch cutoff increases repayment by 6 percentage points, corresponding to roughly a 8% increase relative to similar loans that did not experience the extra delay. Our point estimates are large in magnitude; they translate to a 20% reduction in the likelihood of loan default. IV estimates suggest that each hour of disbursement delay causes a 0.4 percentage points increase in the repayment probability.

While our data offer only a very limited ability to test for the mechanisms behind this results, we can use repayment timing data to answer a key question: is the increase in repayment rates due to borrowers who are subject to a longer delay simply "returning" then loan –paying it back immediately (which the lender allows).[1] The answer is no: the effect of

[1]This could be either because the impetus for the loan passed, or because the borrower obtained credit elsewhere in the meantime.

a longer delay on repayment does not emerge until well into the term of the loan. Notably, most of the effect emerges close to or following due dates. In conjunction with heterogeneous treatment effects that favor married, higher-income, and higher-credit-score borrowers, we take this as evidence that borrower *willingness* to repay may be a key pathway for the effect, perhaps in addition to the ability to repay. While loan utilization is not observed, both pathways are consistent with the extra delays changing how loans are used by the consumer.

Our results are related to recent studies in economics showing that waiting periods –without any choice restrictions– can affect behavior (Imas et al., 2016; DeJarnette, 2018; Brownback et al., 2019). Waiting periods are already used in settings where myopia and impulsivity are perceived to be particularly harmful. For example, many U.S. states require waiting periods prior to the purchase of firearms (Koenig et al., 2016; Edwards et al., 2018). They are also implemented in negotiations (Brooks, 2015) and conflict resolution (Burgess, 2004). Our study also relates to the more traditional literature on behavioral biases in consumer financial choice. Behavioral biases induce agents into suboptimal behavior such as reducing earnings from investments (e.g., Duflo et al., 2011; Kremer et al., 2013) or reduce savings (Dupas and Robinson, 2013). A common solution to these biases is to design financial products that –unlike waiting periods– impose restrictions on agents.[2]

The paper proceeds as follows. In Section 2 we describe the setting, sample, and key variables. Section 3 explains the empirical strategy. Section 4 presents the results, and Section 5 concludes.

2 Settings and sample

Our sample consists of borrowers from an online digital lender in Mexico, which provides short term loans through its website. Borrowers interact with the lender using a browser on

[2]Examples include commitment savings accounts, which cannot be drawn down in the face of an unexpected need (Ashraf et al., 2006), and microfinance, which imposes frequent fixed payments on borrowers (Bauer et al., 2012; Field et al., 2013).

a smartphone or a computer. On the digital lender's home page, borrowers are advised that they can get a loan in "minutes." Borrowers need to meet these requirements: 1) proof of citizenship (photo of national identification card), 2) age between 20-65 years, 3) photo taken from the phone or computer camera, 4) regular income (from credit report), 5) cellphone number and e-mail address, and 6) a bank account.

2.1 Application Process

Loan application and pre-approval happen online during a single browsing session. Shortly afterwards, there is a "know your customer" (KYC) verification process for first-time borrowers to convert pre-approvals into approvals. Approved loans are paid via bank transfer to borrowers' bank accounts.

Potential borrowers start their application by selecting the amount and term of the loan they are applying for on the digital lender's home page. Loan amounts range from 1,500 to 3,000 Mexican Pesos (approximately USD 75 to 150),[3] and loan terms vary from seven to 30 days. Like other digital lenders, the APR (called CAT in Mexico) is very high: up to 478.8%. The web page prominently reports the interest rate in addition to other costs–taxes and fees–at the bottom of the window.

For a first-time borrower, after their request is submitted, the digital lender pulls their credit history from a credit bureau, and notifies them whether their application has been pre-approved for credit. Pre-approved applicants undergo the KYC process, which involves receiving a verification call from a customer-service representative. This process is skipped for repeat borrowers. Next, loans are entered into a spreadsheet which serves as a delivery queue. The queue is sent to the lender's bank multiple times per day for batch processing. Loan funds are then immediately transferred to borrowers' bank accounts. Loans can be repaid anytime after they are disbursed, but the repayment amount includes the interest for the full approved duration of the loan.

[3]The exchange rate during the study period is approximately USD 1 for 20 Mexican Pesos.

2.2 Sample

Our estimation sample consists of 11,512 approved loan applications from 7,206 unique borrowers, with loans disbursed between November 2018 and May 2019.[4] 48% of the loans in our sample are from first-time borrowers. For any borrower, we observe up to three loans. Our dataset contains, for each borrower in the sample: the timestamps of all loan application submissions and loan disbursements; demographic characteristics (sex, marital status, number of dependents, and personal income) as reported in their first loan application; loan sequence (whether this is the first, second or third loan); for first-time loans, we also have information on requested and approved loan amount and repayment length.

Appendix Table A5 reports summary statistics. Successful loan applicants are poorer than the average Mexican worker, with self-reported median monthly income of below 1,000 Pesos (52 USD). 45% of applicants are female, and 11% have no prior credit history. On average, first-time applicants receive 1,785 Pesos, which is approximately 25% of average monthly income. Loan processing times, calculated as the difference between loan disbursal and submission times, are longer for first-time loans –26 hours on average from application to delivery of the loan– while they are 9 hours on average for repeat borrowers.

Our outcome variable is the repayment rate. Defaults (non-repayment) on first-time loans are very high: 32%.[5] 11% of first-time loans are repaid after the due date. Default rates are lower (22%) for repeat borrowers, which is expected since repeat loans are given conditional on past repayment. Appendix Table A6 shows the relationship between borrower/loan characteristics and the repayment likelihood. As expected, income and credit score tend to positively correlate with repayment. The term of a loan correlates negatively

[4]The raw data from the lender contain 15,882 loans. Of these loans, 669 had missing submission times, and three were reported disbursed before they were submitted. Sections 3.1 and 3.2 detail the additional steps that take us to the estimation sample.

[5]Unfortunately, it is not possible to tell from the data whether overdue loans have been partially repaid. It is possible that some of the defaulted loans are repaid after we received the data.

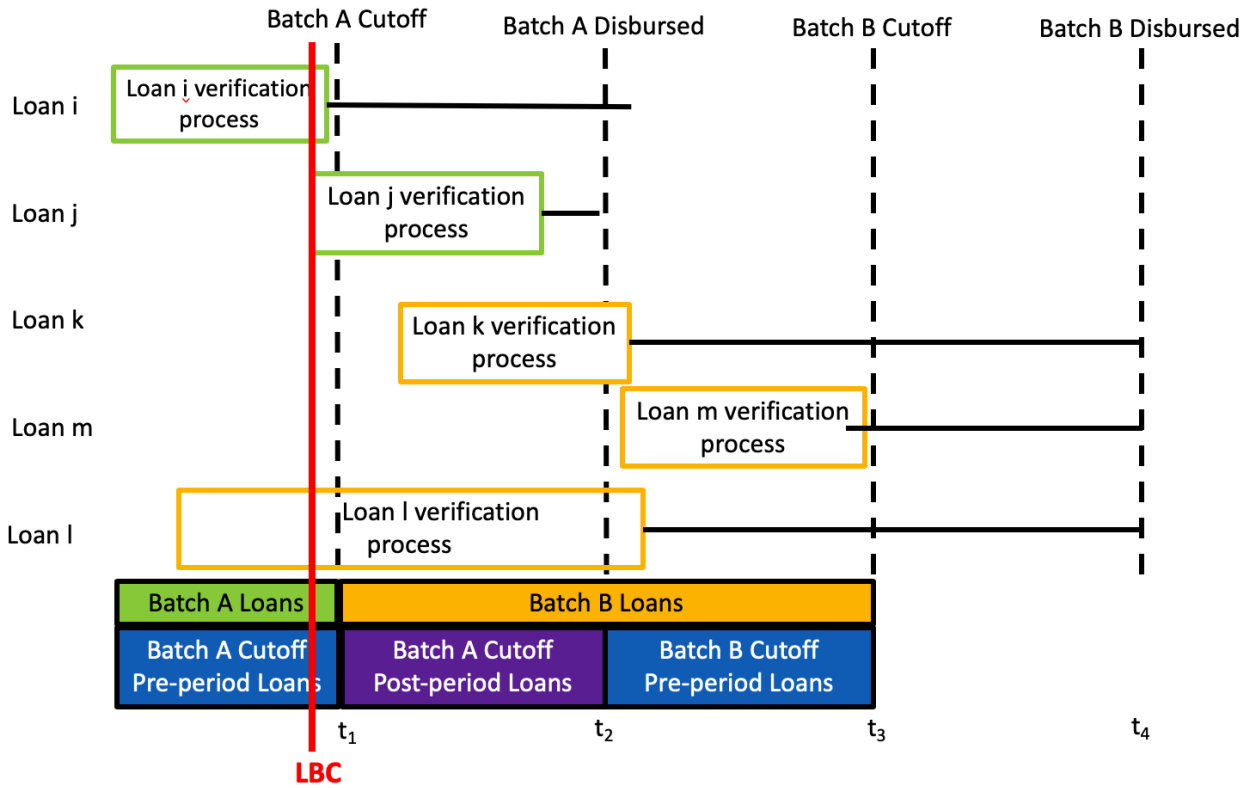
with repayment, but the amount of the loan does not.

3 Empirical strategy

Loan applications happen continuously during the day while loan disbursements are lumpy (Appendix Figure A5). The latter is a consequence of the fact that processed loans are queued and sent to the bank for disbursement in batches. We use the term "batch" to refer to a group of loans that are sent to the bank for disbursement together. By disbursing loans in batches rather than continuously as they are approved, the lender creates random variation in loan processing times. We compare loans that are submitted by the client in time to make it into a particular batch (before the "batch cutoff"), to those that are not (after the "batch cutoff"). Crucially, borrowers (nor the lender, ex-ante) are not aware of these batch cutoff times.

Figure 1 shows a simplified timeline of loan applications and disbursements to illustrate our approach to identification. Applicants apply for loans i, j, k, m, and l at different points in time. All loans go through a verification process, which starts when the loan application is submitted and pre-approved. Verification can take longer for some loans than others (e.g. loan l). Once a loan's verification process is completed –with approval– it is assigned to the current disbursement batch. In the figure, loans i and j are approved prior to the disbursement of Batch A at time t_2. Therefore, they are assigned to Batch A. Loans k, m, and l are approved after the disbursement of Batch A. Therefore, they cannot be assigned to Batch A, and are instead assigned to Batch B, and disbursed at t_4. The batch cutoff is defined as the latest an application could be submitted and make it into that batch. The Batch A cutoff occurs at t_1 shortly after loan j is received. Importantly, not all loans received prior to a batch cutoff are in that batch. For example, loan l is submitted prior to the Batch A cutoff, but because of its extended verification process, it is not approved until after Batch A is disbursed. It ends up in Batch B because it is approved before the Batch B cutoff at

Figure 1: **Hypothetical timeline of loan submission, verification and disbursement**



t_3. On the other hand, no loans received after a batch cutoff can possibly be in that batch. The LBC line stands for "lower bound cutoff," which we discuss in detail in the next section.

Our empirical strategy is best illustrated by the comparison between loans j and k. These loans have been submitted by two separate clients around the same time, but because they fall on different sides of the Batch A cutoff time t_1, loan j is delivered much more quickly even though it takes a similar amount of time to be verified.

In our analysis, we implement this strategy as follows. We start from the list of loan submission times and the list of batch cutoff times. We first assign every loan to the closest batch cutoff (based on its application submission time). Next, we create an indicator called $PostBatch$ that takes the value of one if (similar to loan k) the application was submitted after its assigned cutoff. Then, we compute a continuous variable $DistanceToBatch$ that represents the time of loan application submission minus the assigned batch cutoff time. For

each loan, j, issued to applicant i, we then run the following regression:

$$\begin{aligned} Y_{ij} = \quad & \beta_1 DistanceToBatch_{ij} + \beta_2 PostBatch_{ij} + \\ & \beta_3 DistanceToBatch_{ij} \times PostBatch_{ij} + \delta X_{ij} + \epsilon_{ij} \end{aligned} \tag{1}$$

where Y is the processing time (to demonstrate our first-stage) and loan outcomes (our main results), X controls for individual borrower characteristics, and a variety of application time fixed effects (hour-of-day, day-of-week, and month). The coefficient β_2 identifies the effect of missing a batch cutoff, under the assumption that borrowers who submitted applications just too late to be processed into a particular batch are no different in terms of ex-ante repayment/default likelihood than borrowers who submitted an application in time to be processed into that particular batch.

To estimate Equation (1) and plot results, we use the "rdrobust" suite of commands developed by Calonico et al. (2017). The commands allow for optimal bandwidth selection and automatically provide confidence intervals robust to bias induced by the optimal bandwidth selection.We also report specifications with fixed two-hour bandwidths, that exactly match our discontinuity figures. Because treatment (delay) is assigned at the loan level, we do not cluster standard errors.[6]

3.1 Data construction

The construction of the explanatory variables in equation (1) requires the identification of batches and batch times. Here we outline the procedure for this process, and recommend interested readers to review the Appendix for more specific details.

Constructing batches We do not explicitly observe the batch a loan is assigned to, nor we have a timestamp for the submission of the batch to the bank for disbursement. Consid-

[6]See Abadie et al. (2017).

ering Figure 1, this means that we do not observe the batches' disbursement times t_2 and t_4. We construct batches and batch disbursement times from individual loan disbursement timestamps. We note that most loans are disbursed within seconds or milliseconds from one another (that is, in batches), and use a K-means clustering algorithm to reconstruct the batches for each day. See Appendix Section A.1 for more detail. 259 loans are un-batched, and they are dropped from the sample.

Constructing the cutoffs The next challenge we face is determining the batch cutoff times (t_1 and t_3 from Figure 1). These cutoffs times are never observed, even by the lender. They are hypothetical: they are the latest moment a loan could have been received and processed in the currently accumulating batch. For each batch, we proxy for the batch cutoff time with the submission time of the last loan that is included in the batch. In Figure 1, our proxy for t_1 would thus be given by the application submission time for loan j. We refer to these cutoffs as lower-bound cutoffs (LBCs hereafter), as they precede the actual, unobserved cutoff t_1.[7] Since repeat loans are processed more quickly, we calculate separate LBCs for first-time and repeat borrowers. Once we have a LBC for each batch, we assign each loan in the sample to the closest LBC,[8] and code *DistanceToBatch* and *PostBatch* accordingly. See Appendix Section A.2 for more detail. 2,056 loans are more than twelve hours away from a batch cutoff, and they are dropped from the sample.

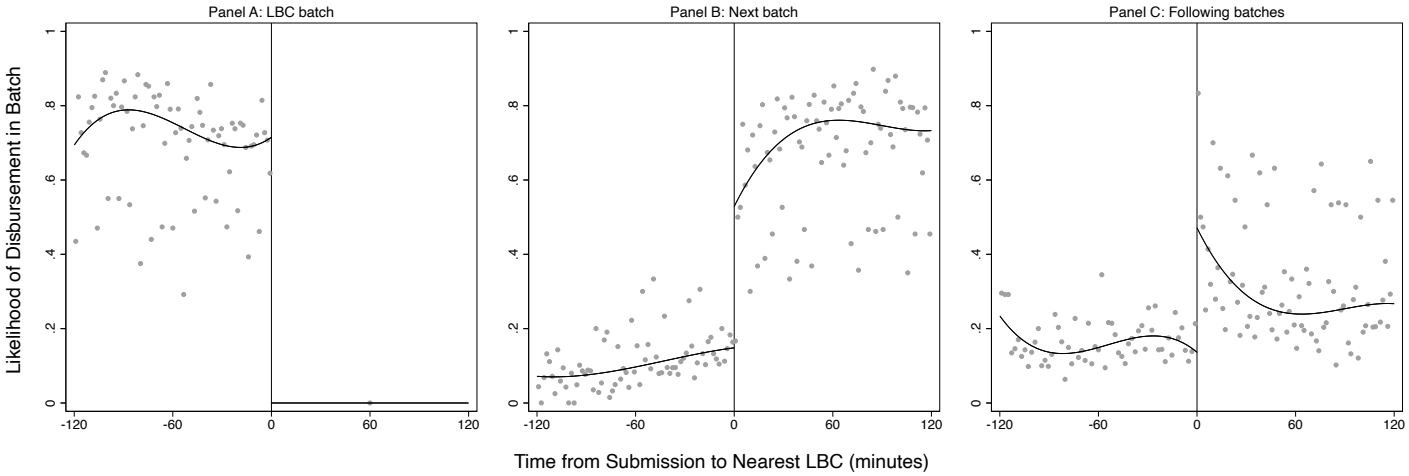
Figure 2 demonstrates the result of this process. It shows the likelihood that a loan is processed in the same batch as the LBC loan, the next batch, or in future batches, as a function of *DistanceToBatch* and the LBC (which is centered at zero).

70% of the loans issued before the cutoff are disbursed within the same batch as the

[7]To be clear, the reason this procedure yields a lower-bound of the batch cutoff is because we cannot know whether any loan received in between the LBC loan and the next observed loan could have been in the same batch as the LBC loan or if it would've been in a subsequent batch.

[8]Recall that there are multiple batches, and therefore multiple batch cutoffs, in a single day. In order to use each loan as a single observation, some assignment rule is necessary.

Figure 2: **Impact of cutoff on likelihood of loan processing in batch**



Notes: Regression discontinuity plots use third degree polynomials, a uniform kernel, and a fixed bandwidth of 120 minutes. We always exclude the LBC loanshist.

LBC loan. Because of the way the LBC is constructed, there are no loans after the LBC time (Panel A) in the LBC batch. Panels B and C show that the likelihood of a loan being processed in subsequent batches jumps immediately after the LBC. Appendix Figures A6 and A7 show these patterns for first-time loans and repeat loans separately. The discontinuity is very sharp for repeat loans, and less clearly defined for first-time loans. This is in line with the expectation that there is more volatility in the length of time it takes to verify a new client.

3.2 Cutoffs and selection

Our approach has three important caveats. First, by defining the LBCs as the set of maximum submission times within batches, the density of submission times after the LBCs will mechanically be lower than the density before.[9] Thus, the density of submissions drops after the LBC, although not due to active manipulation by the applicants or the lender (see Ap-

[9]This is similar to what Miller and Sanjurjo (2018) call "streak selection bias" in the context of collecting data to analyze the hot hand fallacy, and can be shown in a simulation of our data with a uniform density of submission times.

pendix Figure A4). Second, loans that create LBCs are processed quickly. The average delay in disbursing LBC loans is 4.39 hours, against 10.69 hours for loans issued in the five minutes prior to the LBC. Indeed, the speed of processing the LBC loan is precisely the reason why it created an LBC. It is likely that LBC loans are also different along unobservable characteristics. Finally, loans submitted just after the LBC may be more difficult to process than loans submitted just before the LBC. If they were not, they would have been included in the batch with the LBC loan, and become the LBCs themselves. Figure 2 provides a visual confirmation that loans issued immediately after the LBC (within the next 20 minutes or so) are different from subsequent loans: they have a lower likelihood of being processed in the next batch (Panel B), and are more likely to be processed in future batches (Panel C). Additional analysis in Appendix Section A.3 supports the hypothesis that loans immediately after the LBC are negatively selected.

We address all these issues by excluding from the LBC loan (863 loans) and all loans received within 20 minutes after the LBC (512 loans).[10] In other words, we employ a one-sided "half-donut RD," where we drop only the right side of the donut hole.[11] This yields our estimation sample of 11,512 loans. Following these procedures, loan submission density and observable characteristics of the borrower are smooth across the cutoff. See Appendix Section A.3 for details.

[10]This process should have the added benefit of reducing measurement error associated with proxying for the cutoff, so long as we do not overshoot the true cutoff by more than the distance to the LBC.

[11]Note that selection concerns are absent for the loans that were submitted before the LBC, because those loans are processed in either the same batch as the LBC or in subsequent batches, i.e., they are not selected based on their batching. We thus include all loans leading up to the LBC.

4 Results

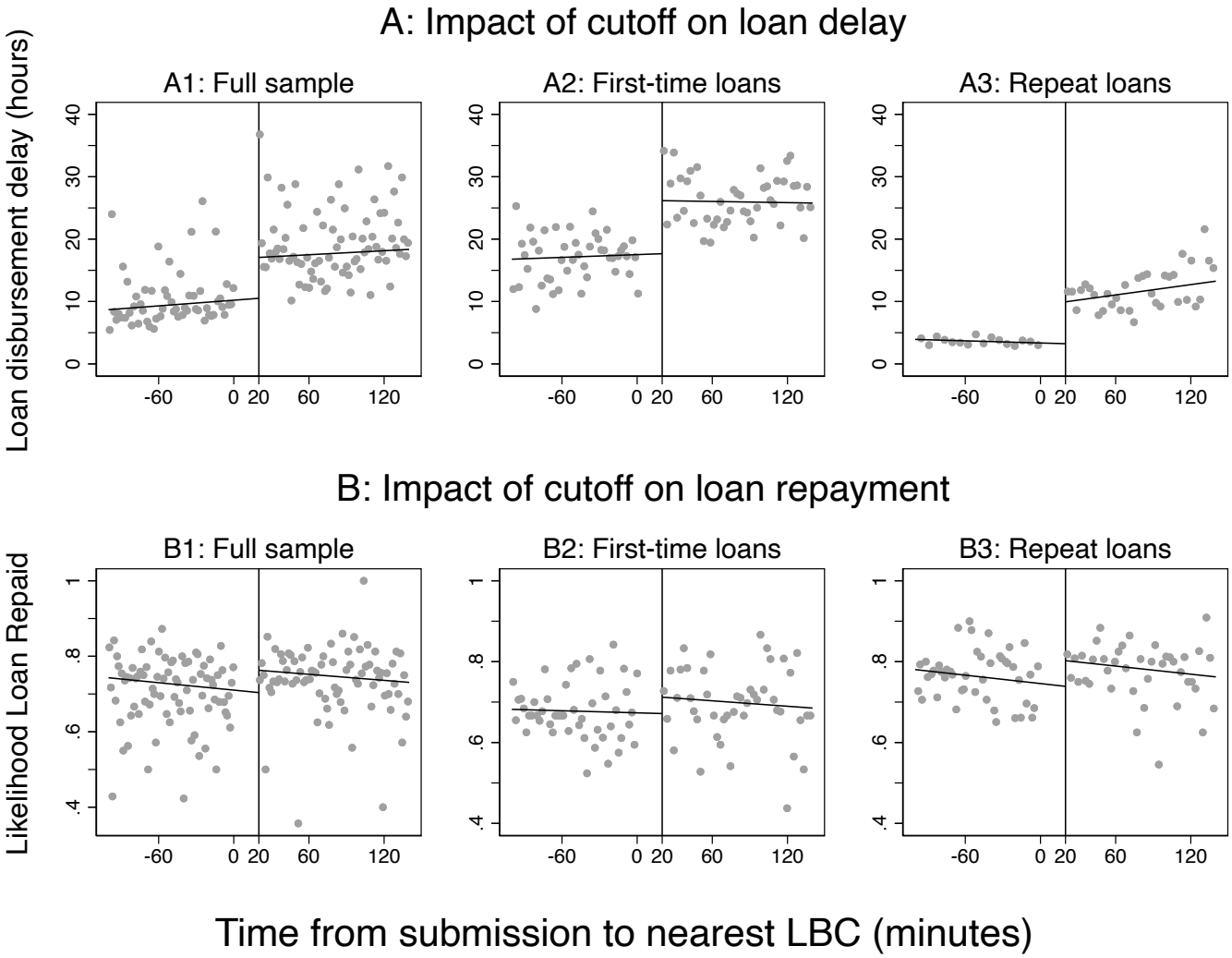
4.1 First stage

We begin by evaluating the degree to which our constructed discontinuity in batch processing captures variation in the overall delay–the elapsed time between application submission and loan disbursement. First, we winsorize the delay distribution at the 90th percentile to account for a large right tail that is not of interest; the longest delay in our sample is just over 27 days, and the 90th percentile is 62.75 hours. Next, we estimate equation 1 with the delay length (in hours) as the dependent variable. We first show these results graphically in Panel A of Figure 3. For the visual regression discontinuities, we use a straightforward approach: we assume a bandwidth of two hours around the LBC, estimate a linear fit, and a use uniform estimation kernel. As noted in the previous section, the LBC loans have been removed from the data, and that we estimate the discontinuity at 20-minutes post-LBC, with the interim data excluded. There is a clear increase in delay at the LBC. The relative size of this effect is more pronounced for repeat loans than first-time loans. This is because the potential speed of delivery is higher for the former.

Appendix Table A4 shows corresponding regression estimates using both a model that exactly matches the specification from Figure 3, as well as optimal-bandwidth models with a variety of control variables: borrower demographics and application submission time fixed effects. There is a large and statistically significant effect of the cutoff on loan delay in every model. We estimate that missing the cutoff essentially doubles the borrower's wait time. Consistent with the regression discontinuity figures, we find that the absolute size of the effect is larger for first-time loans, but estimates are more precise for repeat borrowers. Moreover, repeat loans experience a much larger relative size increase in delay.

One way of thinking about the induced delays is that they force borrowers to "sleep on" the decision of how to use their loan. We can also measure the first stage in a way that captures that intuition. Appendix Table A7 shows identical regression discontinuity specifications

Figure 3: **Regression discontinuity figures**



Notes: Regression discontinuity plots use linear fit with a uniform kernel and a fixed bandwidth of 120 minutes. The LBC loan and loans between the LBC and 20-minutes post-LBC are excluded.

that estimate the impact of missing a batch cutoff on the likelihood a borrower receives her loan on the same day she applied. The results are clear: the induced delays make same-day disbursement about much less likely.

4.2 Effect of delays on repayment

We now estimate the effects of the delay-inducing cutoff on loan repayment rates. Using the same graphical specification for repayment rates as we did in the previous section for loan

delay, we show the main result in Panel B of Figure 3. We observe an increase in repayment likelihood at the 20-minute post-LBC cutoff. Repeat loans show a slightly larger increase in repayment than first-time loans.

Table 1 shows the corresponding regression estimates. The estimates in column (1) are obtained from a specification that matches Figure 3: two-hour bandwidth, uniform kernel, linear estimation. In columns (2)-(4) we use optimal bandwidth selection and a triangular estimation kernel. We allow for an asymmetric optimal bandwidth because the exclusion of loans submitted within 20 minutes following the LBC creates an asymmetry in density around the post-LBC latent cutoff. The models in column (2) feature no control variables, in column (3) we add application submission day-of-week, hour-of-day and month fixed effects, and in column (4) we add borrower and loan control variables.[12] Panel A shows the full sample estimate, and Panels B and C show estimates for first-time loans and repeat loans, respectively. Below each estimate, we include the effect magnitude as a percentage of the pre-cutoff mean delay within two hours of the cutoff for the relevant sample, the optimal bandwidth as determined by the rdrobust command, and the number of observations within that optimal bandwidth. The sample that is fed into the optimal bandwidth algorithm is held fixed across the specifications, but the number of observations within the optimal bandwidth varies slightly across specifications as the optimal bandwidth changes slightly when we add control variables. We include two p-values below each estimate. The first estimate has the advantage of pertaining to the point estimate of interest, but it does not account for potential bias due to bandwidth selection. The second estimate has the advantage of accounting for bias due to bandwidth selection, but it pertains to the quadratic estimate used for bias correction, not the linear estimate of interest. Both are heteroskedasticity-robust.

In the full sample in Panel A, we find that the induced delay (of roughly ten hours on average) increases repayment rates by six percentage points. This is roughly an 8% increase in repayment rates (equivalently, a 20% reduction in the default rate). The effect is very

[12] In Panels A and C, we also add a fixed effect for a borrower's sequential loan number in column (3).

Table 1: **Impact of cutoff on loan repayment**

RD bandwidth:	Two-hour	Optimal		
	(1)	(2)	(3)	(4)
A. Full sample (N = 11,512)				
PostBatch	0.059	0.063	0.060	0.056
	(0.023)	(0.024)	(0.024)	(0.024)
Estimate p-value	0.011	0.008	0.012	0.018
Bias-corrected estimate p-value	0.044	0.017	0.021	0.038
Effect as % of pre-cutoff mean	8%	9%	8%	8%
Optimal bandwidth (mins)		[144,119]	[144,112]	[146,112]
Observations within bandwidth	7,177	7,704	7,602	7,658
B. First-time loans (N = 5,530)				
PostBatch	0.041	0.040	0.041	0.054
	(0.036)	(0.035)	(0.034)	(0.034)
Estimate p-value	0.259	0.251	0.227	0.110
Bias-corrected estimate p-value	0.813	0.326	0.274	0.146
Effect as % of pre-cutoff mean	6%	6%	6%	8%
Optimal bandwidth (mins)		[153,136]	[162,126]	[164,126]
Observations within bandwidth	3,090	3,565	3,554	3,577
C. Repeat loans (N = 5,982)				
PostBatch	0.064	0.083	0.078	0.074
	(0.030)	(0.034)	(0.034)	(0.034)
Estimate p-value	0.037	0.015	0.021	0.029
Bias-corrected estimate p-value	0.015	0.038	0.050	0.067
Effect as % of pre-cutoff mean	8%	11%	10%	10%
Optimal bandwidth (mins)		[123,110]	[127,110]	[123,111]
Observations within bandwidth	4,087	4,036	4,084	4,068
Day-of-week, hour-of-day, month FEs	N	N	Y	Y
Borrower controls	N	N	N	Y

Notes: All estimated discontinuities are from linear models that exclude the LBC loan, and loans received within 20 minutes after the LBC. Column (1) reports a specification with a uniform estimation kernel and a fixed bandwidth of 120 minutes around the 20-minute post-LBC cutoff. Columns (2-4) report specifications with a triangular estimation kernel, and an optimal bandwidth selected from a 12-hour window around the LBC. Heteroskedasticity-robust standard errors of the linear estimates are shown in parentheses below the estimates, calculated using the nearest-neighbor variance estimator with a minimum of three matches. We report both the heteroskedasticity-robust p-values of the linear estimates, and the bias-correction- and heteroskedasticity-robust p-values of the quadratic, bias-corrected estimates. We also report the estimated effect as a percentage of the pre-cutoff mean repayment rate within the two-hour bandwidth. The optimal bandwidths –rounded to the nearest integer– are reported for the specifications in columns (2)-(4), and observations within the used bandwidth are reported below. The overall sample sizes for each panel correspond to all loans within twelve hours of an LBC. The fixed effects added in column (3) include the hour-of-day, day-of-week, and month of application submission. In Panels A and C, a fixed effect for the borrower's sequential loan number is also included. The borrower controls added in column (4) are age, age squared, sex, marital status, number of dependents, log income, and credit score.

similar in magnitude across specifications, and is always statistically significant according to both sets of p-values. Appendix Figure **??** shows that the estimate gets closer to zero when we include data within the 20-minute post-LBC exclusion window, but gets larger if we expand the window beyond 20 minutes. This is consistent with our concerns about selection and measurement error when using the loans in this excluded window.

As expected, the effect is more precise for repeat borrowers, and perhaps slightly larger, although the differences in estimates between Panels B and C is not statistically significant. Column (4) shows a statistically significant 7.4 percentage point (10%) increase in repayment for repeat loans in Panel C, and a not-quite-statistically significant 5.4 percentage point (8%) increase in repayment for first-time loans in Panel B. For first-time loans, the magnitude of the effect grows as our specification becomes more robust, but it never achieves statistical significance.

While these results demonstrate a causal effect of induced delays on repayment, the unit of measurement of the effect is somewhat arbitrary: a 5.6 percentage point increase in repayment in response to an induced additional delay of 9.81 hours (estimates from Panel A, column (4) of Table 1 and Appendix Table A4, respectively). Thus, the back-of-the-envelope calculation implies an increase of 0.6 percentage points per hour of induced additional delay. Alternatively, we can directly estimate the causal effect of loan disbursement delay on repayment rates (albeit at the same margin as the crude calculation) using two-stage least squares. We instrument for loan disbursement delay using our regression discontinuity model from equation 1 using a fixed bandwidth of two hours.[13] This approach yields slightly smaller, but qualitatively similar results; using the most robust specification in the full sample, we estimate that each hour of induced delay increases repayment rates by 0.4 percentage points ($p = 0.016$). Estimates are shown in Appendix Table A8.

[13]The use of least squares implies an uniform estimation kernel.

4.3 Timing of repayment

Without details on how these loans were used, it is difficult to exactly identify the mechanism behind the effect of a disbursement delay on loan repayment. However, by considering the timing of loan repayment, we hope to distinguish between three broad mechanisms. First, it could be that a delay leads individuals to conclude that they did not need the loan in the first place. If so, we would expect to see a higher rate of very early repayments for delayed loans. Second, it could be that the delay leads individuals to divert their loans to relatively higher short-run (prior to due date) return-on-investment opportunities than they would have otherwise.[14] If so, we would expect to see a higher rate of on-time repayments for delayed loans. Third, it could be that even holding short-run return fixed, the delay leads individuals to divert their loans to investments that yield a higher experienced utility, meaning that the costs of failing to pay back the loan –particularly reduced future loan access– loom particularly large. If so, we would expect to see a higher rate of extended repayments –where borrowers are willing to incur fees to avoid default– for delayed loans.[15]

To assess repayment timing, we re-arrange our data as a panel. For each loan in the sample, we define the time dimension as days since loan disbursed, ranging from zero to 356 (the latest repayment we observe). On each day, a loan is classified as repaid or not, and for each day, we estimate the effect of missing a batch cutoff using our regression discontinuity specification. By construction, the estimates will converge to the overall estimate in Table 1, however, by plotting the coefficients (and confidence intervals) over time, we can observe when the difference in the repayment rate emerges by looking for time periods with a positive slope. We use the specification from column (2) of Table 1, which implements the optimal bandwidth procedure, but does not include control variables.[16] Results are shown in Figure

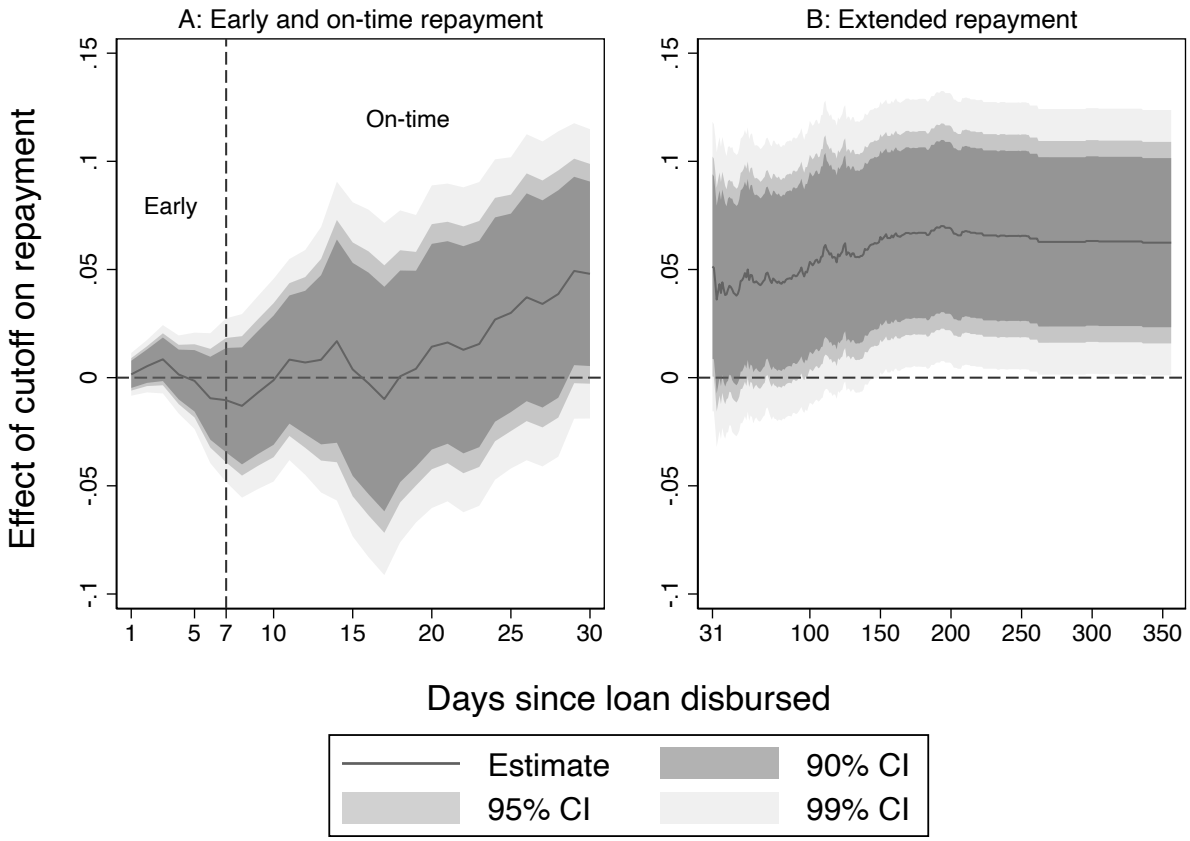
[14]This could include simply holding onto the money for extra temporary liquidity, if some loans are motivated by negative expected return investments.

[15]One source of this utility gain could be long-run (after due date) return on the investment.

[16]Near the beginning of the time dimension the repayment rate is close to zero, so we can reduce co-

4.[17] We split the plot into two parts to account for the right-skewed distribution of repayment timing; Panel A shows one to thirty days after loan disbursement, and Panel B shows 31 to 356 days after loan disbursement. Because the earliest possible due date was seven days after disbursement, we label one to seven days the "early" repayment period. The latest possible due date was thirty days after disbursement, so seven to thirty days is labeled the "on-time" repayment period, and the period after thirty days is referred to as the "extended" repayment period.

Figure 4: **Regression discontinuity estimates over time**



During the early repayment period, we find no effect of the cutoff on repayment, nor is

linearity concerns by omitting the fixed effects and demographic controls.

[17]We use the conventional confidence intervals for the figure because they pertain to the estimated coefficient.

there a positive slope to the estimate plot. Starting about 17 days after disbursement, we begin to observe a positive slope that is then maintained throughout the remainder of the time dimension. By the very end of the on-time repayment period, we can begin to detect a significant effect of the cutoff. On day 30, we estimate an effect of the cutoff of 4.8 percentage points ($p = 0.064$), which represents roughly three-quarters of the overall effect.[18] During the extended repayment period, the slope remains positive, explaining the remaining effect.

While we cannot offer direct evidence that the delay changed how loans were used, the timing data is consistent with the idea that delayed loans may have provided higher returns than immediate loans, and that borrowers with delayed loans ended up valuing the product more.[19] One clear pathway through which a disbursement delay may lead to a change in loan use is through household bargaining; without a delay, an individual may be able to apply for, obtain, and use the loan without confronting their partner. This may be more difficult if disbursement is delayed overnight. Indeed, for a married borrower, the effect of missing the batch cutoff is to increase repayment by 10.8 percentage points ($p = 0.003$), and for an unmarried borrower, the effect is 1.2 percentage points ($p = 0.708$). Heterogeneity analysis also reveals evidence that borrower willingness to repay may be important. We find larger effects for borrowers with *higher* income and credit scores. Splitting the sample, we find an effect of 7.9 percentage points ($p = 0.011$) for above-median income, and an effect of 2.8 percentage points ($p = 0.424$) for below-median income. For borrowers assessed by the lender to have a "better" or "best" credit score, we find an effect of 14.2 percentage points ($p = 0.001$), and for borrowers assessed as "average," "marginal," or "none," we find an effect of 3.3 percentage points ($p = 0.251$.)[20]

[18] The bias-correction robust p-value is 0.082.

[19] If many defaults are strategic, it could be that even the on-time repayment period effect operates more through borrower value than financial return, but we cannot separately identify this.

[20] All heterogeneity estimates are in Appendix Table A9, and are obtained using the specification from column (4) of Table 1. It is difficult to estimate these heterogeneous effect as interaction terms using our optimal bandwidth approach, and so we present them as separate models.

5 Discussion

The fintech industry provides digital loans through mobile apps and text messages to millions in developing countries. The speed and ease of access to credit makes digital loans very appealing. Nevertheless, despite their small size, these loans often become expensive and many borrowers struggle to repay. In addition, defaulting on a digital loan, however small, comes at a cost: beyond losing the opportunity for future credit from the same provider, defaulters could be reported to a credit bureau (as in our context), or suffer disruptions in the mobile services they rely on (as in cases where the lender happens to operate through a mobile service provider).

We study whether one of the primary features of digital credit–the speed of delivery of funds–affects the likelihood that a loan is repaid. We leverage the fact that our lender processes payments in discrete batches to construct a regression-discontinuity estimate of the effect of loan delivery speed on loan outcomes. The first stage of the design is very strong: applicants who submitted their application just too late to be verified in time to be added to the current disbursement queue receive their loan nearly ten hours later than those applicants who submitted just in time (from a baseline of ten hours). We find that this induced delay increases repayment rates by 6 percentage points. This is a large increase; the default rate in our sample is 20%.

Governments, central bank governors, and policy makers have recently started to advocate for supervision and consumer-protection measures in the digital credit market (Donovan and Park, 2019). While the nature of our data prevent us from making firm statements about consumer welfare, we think our study is an important first step in exploring a hallmark feature of new digital credit markets. While the immediacy offered by digital credit has the advantage of providing access to those needing liquidity in case of emergency or a limited-time investment opportunity, it might negatively affect borrowers' repayment rates, investment decisions, their overall amount of debt, and their ability to get future loans through the formal financial system. Understanding these trade-offs is of importance to policymakers

as well as to the private sector. Further research on delivery speed, especially using data on how digital credit is used, and how delays affect the demand for credit, could speak to whether mandatory waiting periods are a sensible consumer protection in this market.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Ashraf, N., D. Karlan, and W. Yin (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics 121*(2), 635–672.

Bauer, M., J. Chytilová, and J. Morduch (2012). Behavioral foundations of microcredit: Experimental and survey evidence from rural india. *American Economic Review 102*(2), 1118–39.

Bharadwaj, P., W. Jack, and T. Suri (2019). Fintech and household resilience to shocks: Evidence from digital loans in kenya. Technical report, National Bureau of Economic Research.

Björkegren, D. and D. Grissen (2018). The potential of digital credit to bank the poor. In *AEA Papers and Proceedings*, Volume 108, pp. 68–71.

Brooks, A. W. (2015). Emotion and the art of negotiation. *Harvard Business Review*.

Brownback, A., A. Imas, and M. A. Kuhn (2019). Behavioral interventions increase the effectiveness of healthy food subsidies.

Burgess, H. (2004). *Beyond Intractability*, Chapter Cooling-Off Periods. Conflict Information Consortium, University of Colorado, Boulder.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal 17*(2), 372–404.

Cattaneo, M. D., M. Jansson, and X. Ma (2018). Manipulation testing based on density discontinuity. *The Stata Journal 18*(1), 234–261.

DeJarnette, P. (2018). Temptation over time: Delays help. Technical report, Working paper.

Donovan, K. P. and E. Park (2019). Perpetual debt in the silicon savannah.

Duflo, E., M. Kremer, and J. Robinson (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American economic review 101*(6), 2350–90.

Dupas, P. and J. Robinson (2013). Why don't the poor save more? evidence from health savings experiments. *American Economic Review 103*(4), 1138–71.

Edwards, G., E. Nesson, J. J. Robinson, and F. Vars (2018). Looking down the barrel of a loaded gun: The effect of mandatory handgun purchase delays on homicide and suicide. *The Economic Journal 128*(616), 3117–3140.

Field, E., R. Pande, J. Papp, and N. Rigol (2013). Does the classic microfinance model discourage entrepreneurship among the poor? experimental evidence from india. *American Economic Review 103*(6), 2196–2226.

Francis, E., J. Blumenstock, and J. Robinson (2017). Digital credit: A snapshot of the current landscape and open research questions. *CEGA White Paper*.

Heidhues, P. and B. Kőszegi (2010). Exploiting naivete about self-control in the credit market. *American Economic Review 100*(5), 2279–2303.

Imas, A., M. Kuhn, and V. Mironova (2016). Waiting to choose.

Izaguirre, J. C., M. Kaffenberger, and R. Mazer (2018, September). It's time to slow digital credit's growth in east africa. Technical report, CGAP Blog.

Karlan, D. and J. Zinman (2010). Expanding credit access: Using randomized supply decisions to estimate the impacts. *The Review of Financial Studies 23*(1), 433–464.

Koenig, C., D. Schindler, et al. (2016). Dynamics in gun ownership and crime-evidence from the aftermath of sandy hook. Technical report, Working paper.

Kremer, M., J. Lee, J. Robinson, and O. Rostapshova (2013). Behavioral biases and firm behavior: Evidence from kenyan retail shops. *American Economic Review 103*(3), 362–68.

Lang, K., K. Leong, H. Li, and H. Xu (2019). Lending to the unbanked: Relational contracting with loan sharks. Technical report, National Bureau of Economic Research.

Makles, A. (2012). Stata tip 110: How to get the optimal k-means cluster solution. *The Stata Journal 12*(2), 347–351.

Malingha, D. (2019, August). This nobel prize-winning idea is instead piling debt on millions. Bloomberg Future Finance.

Mazer, R. and A. Fiorillo (2015). Digital credit: Consumer protection for m-shwari and m-pawa users. *CGAP. April 21.*

McKee, K., M. Kaffenberger, and J. M. Zimmerman (2015). Doing digital finance right: The case for stronger mitigation of customer risks. *CGAP Focus Note 103.*

Meier, S. and C. Sprenger (2010). Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics 2*(1), 193–210.

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

Miller, J. B. and A. Sanjurjo (2018). Surprised by the hot hand fallacy? a truth in the law of small numbers. *Econometrica 86*(6), 2019–2047.

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

Skiba, P. M. and J. Tobacman (2019). Do payday loans cause bankruptcy? *The Journal of Law and Economics 62*(3), 485–519.

A Appendix for Online Publication

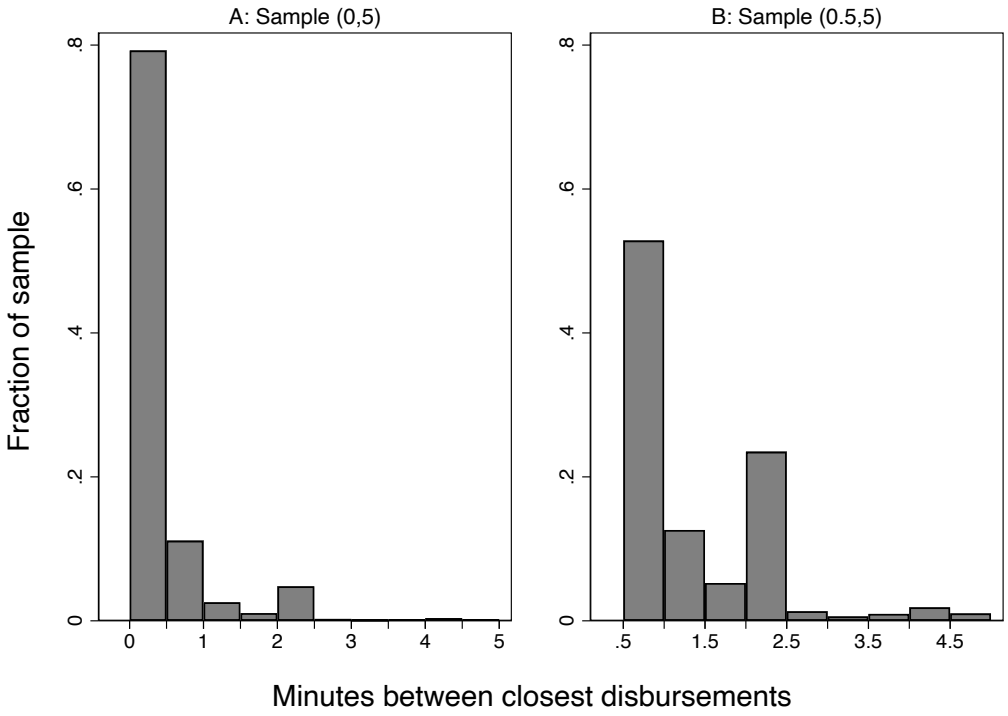
A.1 Data construction: Batching identification

We do not observe batches' disbursement times (t_2 and t_4 in Figure 1). Instead, we construct the batches and batch cutoffs from our data on disbursement and submission times. Because loans are disbursed from a batch, we observe a series of loan disbursements in quick succession to one another. For example, the median gap between any loan disbursement and the nearest other disbursement in our sample is six seconds, and 94% of loans are disbursed within a minute of another loan. Some loans are processed in isolation of others and appear "unbatched" (not belonging to any particular batch), we exclude these from the data. In particular, our exclusion criteria is to drop all loans that are not disbursed within 2.5 minutes of other loans, as the detectable density of loans falls sharply at that cutoff (see Appendix Figure A1). This drops 259 loans. Alternative approaches, including using k-means clustering to identify batched versus unbatched loans using the minimum distance to another loan produces similar results.

Among the "batched" loans, we use the k-means clustering algorithm to assign each loan to a specific batch within a given day. There are two parts of this process. First, we assume that each day consists of six batches, and let the algorithm assign each loan to one of the six batches under that assumption. Then, we repeat that process assuming five, four, three, two, and one batches per day. Next, we use the maximum proportional reduction of error (PRE) statistic to select the optimal number of batches for each day (Makles, 2012). Of the 142 days in our sample, one is a 1-batch day, 37 are 2-batch days, 47 are 3-batch days, 38 are 4-batch days, eleven are 5-batch days, and eight are 6-batch days.

On some days with smaller samples, random initial batch assignments lead to final batch assignments that overlap, likely representing a locally –rather than globally– optimal assignment. As such, we initialize the algorithm on each day by assigning every k^{th} observation to one of k batches in sequential fashion. This is essentially a stratified randomization proce-

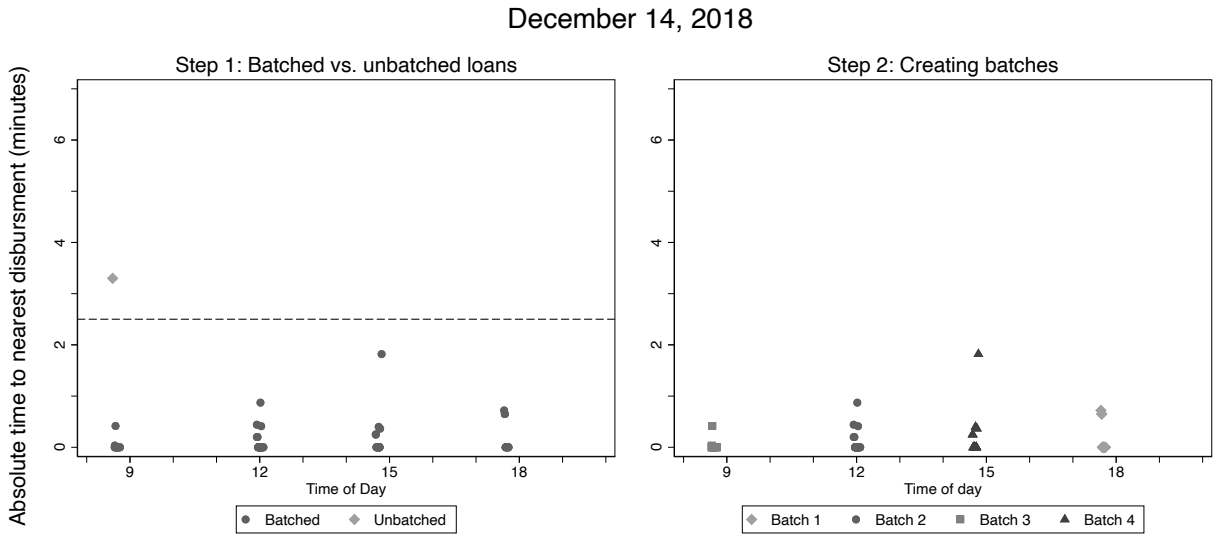
Figure A1: **Time between loan deliveries in a batch**



dure to ensure a neutral starting point even in a small sample. In four of 142 days, we still end up with overlapping batches with this approach. By switching to a segmented initial batch assignment, whereby the first N/k observations are assigned to batch one, and the second N/k observations are assigned to batch two, etc. (when assigning N observations to k clusters), we extract non-overlapping batches for these four days (although this initialization does much worse on the overall sample). On one of 142 days, the data clearly suggest one batch is appropriate, but the cluster optimization procedure cannot return an answer of one, so we manually assign all observations on this day to a single batch.

Appendix Figure A2 shows an example of the batching process applied to December 14, 2018. One loan on this day was processed more than 2.5 minutes from the other loans, and is thus removed from the data (left panel). The clustering, applied to the remaining loans, produces four distinct batches around 9am, 12pm, 3pm, and 6pm (right panel).

Figure A2: **Example of batching process**



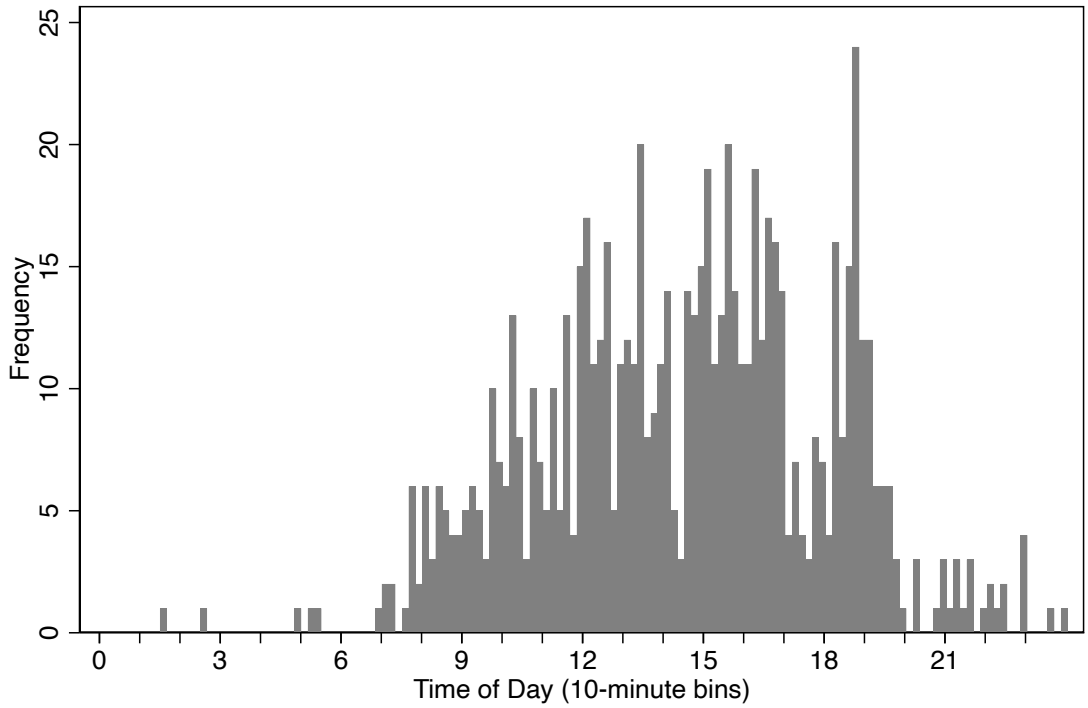
Notes: The left panel shows the first step of our batching process, where we drop one loan that was disbursed 2.5 or more minutes apart from any other loan. Remaining "batched" loans are fed to the k-means clustering algorithm. The right panel shows the batching results from the procedure.

A.2 Constructing the cutoffs

The lower-bound cutoff (LBC) is defined as the latest application submission time within a batch. Appendix Figure A3 shows the distribution of the LBCs in our sample.

We code $DistanceToBatch$ as the difference (in minutes) between loan application submission time and the relevant cutoff, and $PostBatch$ as an indicator for whether $DistanceToBatch$ is positive. Accordingly, in Figure 1, the application submission (start of the verification process) for loan k is closer to the Batch A cutoff than to the Batch B cutoff. Therefore, loan k is assigned to the Batch A cutoff, with a positive value of $DistanceToBatch$ ($PostBatch = 1$). The start of the verification for loan m is closer to the Batch B cutoff. Therefore, loan m is assigned to the Batch B cutoff, with a negative value of $DistanceToBatch$ ($PostBatch = 0$).

Figure A3: **Distribution of lower-bound cutoffs**

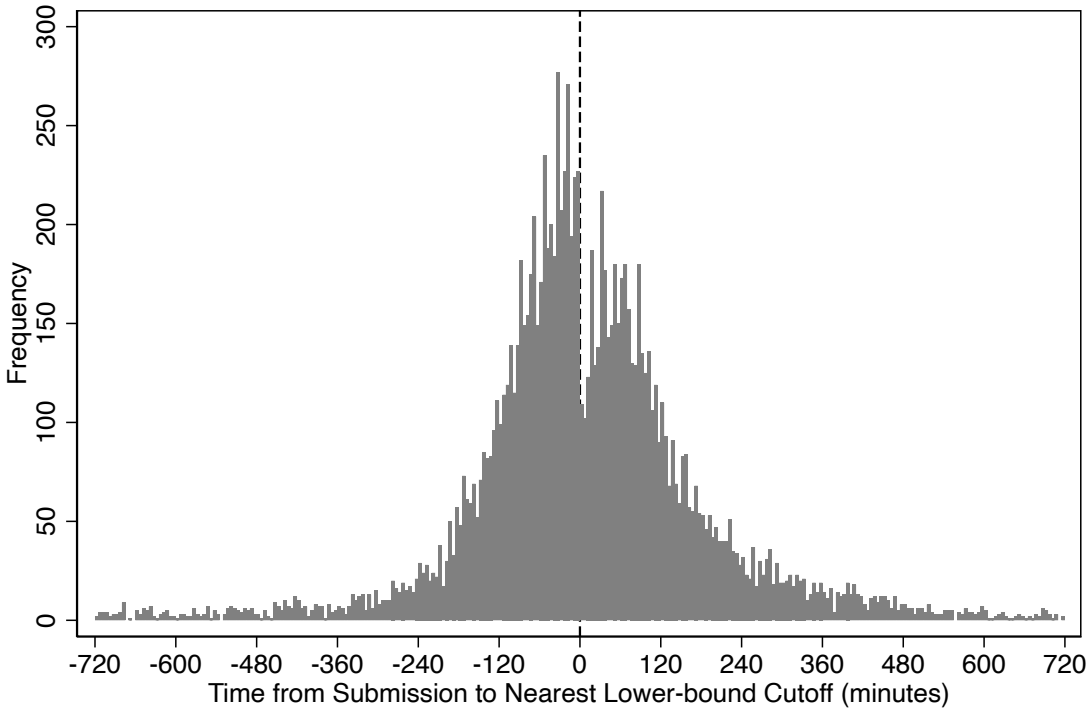


A.3 Selection around the cutoffs

Because the density of submissions will drop at the LBC,this means that the typical regression discontinuity validity test–smoothness of the density of the running variable–is not informative for out context. Appendix Figure A4 shows the density of loan submissions within a 4-hour window of an LBC, excluding the LBC loan. Note that because every window around an LBC must, by definition, contain a submission for which $DistanceToBatch = 0$, zero will be overrepresented in the distribution. We exclude those observations from the figure. It shows a very clear violation of smoothness at the LBC.

Additionally, there is a selection issue for both LBC loans and submitted just after the LBC. Perhaps loans after the LBC were processed in subsequent batches because they were more difficult to process; if they had been easy to process, they would have been included in the batch with the LBC loan, and become the LBCs themselves. The LBC loan is likely to be selected in the opposite manner. If processing difficulty is negatively correlated with

Figure A4: **Density of** *DistanceToBatch*, **12-hour window**



Notes: Five-minute bins.

borrower quality, a failure to fix these issues could lead to biased estimates of the β_2 coefficient in equation (1) towards indicating harmful effects of induced delays. For example, a failure of the borrower to pick up the phone the first time they are called for identity verification could be correlated with borrower quality. The average time from submission to disbursement is 19.6 hours for loans submitted within 20 minutes after the LBC, and 17.7 hours for loans submitted 20-60 minutes after the LBC. This supports the idea that loans right after the LBC take longer to process, and that these loans could be negatively selected.

To determine how to exclude these loans, we first consider the smoothness of the density of the running variable above the LBC. We use the "rddensity" suite of commands developed by Cattaneo et al. (2018) to determine where the right side of the density shown in Appendix Figure A4 achieves smoothness, starting from the LBC; where does it shift from outlier loans that couldn't be processed quickly enough to be the LBC to typical loans that simply missed the previous batch? Starting at five minutes post-LBC, we test for smoothness through

each five-minute increment above the LBC, up to one hour. We use the optimal bandwidth approach, with bias-correction robust standard errors. Appendix Table A1 shows the p-value associated with each test, along with the optimal bandwidth and effective observation count.

Table A1: **Density-smoothness tests of post-LBC application submissions**

Minutes post-LBC	p-value	Optimal bandwidth	Obs. in bandwidth
5	0.003	[3,41]	1,276
10	0.050	[6,44]	1,508
15	0.228	[25*,49]	1,971
20	0.380	[11,60]	2,281
25	0.805	[9,78]	2,723
30	0.922	[18,132]	3,892
35	0.257	[12,103]	3,213
40	0.144	[13,77]	2,706
45	0.447	[23,73]	2,911
50	0.279	[20,70]	2,673
55	0.745	[19,67]	2,468
60	0.577	[26,64]	2,595

Notes: *: this bandwidth is outside the range of the data. Optimal bandwidths are rounded to the nearest integer. Discontinuities are estimated with a quadratic fit of the density and a triangular kernel. We use distinct optimal bandwidths left and right of the cutoffs to allow for the larger amount of data to the right of these cutoffs to improve precision. p-values are from the heteroskedasticity and bias-correction robust standard errors, calculated using the nearest-neighbor variance estimator with a minimum of three matches.

The first failure to reject is at 15-minutes post LBC, although the estimated optimal bandwidth exceeds the given range (a test with a symmetric bandwidth of just under 15 minutes yields a p-value of 0.175). Beginning with 20-minutes post-LBC, we always reject the null with a well-defined bandwidth.

Does a 20-minute exclusion window make sense using other approaches? We now test directly for smoothness in a key observable –creditworthiness– through post-LBC cutoffs. Using the rdrobust optimal bandwidth approach, in Table A2, we show how the assessed credit score category of borrowers changes when a batch cutoff is missed. Of the 65 p-values in the table, only two are less than 0.05, and both of these are associated with fewer "best" score borrowers being in the sample after the LBC –consistent with our concern

regarding negative selection in this period right after the LBC. Focusing on that credit score category, the discontinuities at zero, five, ten, and 15 minutes post-LBC are at least marginally statistically significant, and we fail to reject smoothness at twenty minutes.

Table A2: **Borrower credit score smoothness through post-LBC cutoffs**

Credit score:	None		Marginal		Average		Better		Best	
Minutes post-LBC	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
0	-0.004	0.834	-0.033	0.187	0.038	0.141	0.011	0.559	-0.028	0.015
5	-0.009	0.627	-0.033	0.229	0.041	0.133	0.008	0.708	-0.024	0.061
10	-0.005	0.809	-0.025	0.364	0.025	0.366	0.013	0.543	-0.027	0.035
15	0.006	0.713	-0.024	0.401	0.025	0.393	0.005	0.843	-0.023	0.090
20	-0.006	0.841	-0.003	0.935	0.011	0.807	0.009	0.728	-0.017	0.256
25	-0.006	0.837	-0.003	0.995	0.002	0.898	0.014	0.552	-0.015	0.310
30	-0.003	0.913	-0.005	0.976	0.004	0.941	0.010	0.717	-0.012	0.441
35	-0.011	0.501	0.010	0.490	0.005	0.920	0.013	0.661	-0.022	0.091
40	-0.018	0.219	0.019	0.288	0.011	0.705	0.005	0.894	-0.021	0.080
45	-0.019	0.153	0.005	0.713	0.012	0.603	0.016	0.456	-0.020	0.105
50	-0.022	0.071	-0.001	0.995	0.011	0.567	0.024	0.215	-0.020	0.143
55	-0.020	0.078	0.003	0.998	0.009	0.445	0.015	0.480	-0.019	0.255
60	-0.012	0.119	0.001	0.631	0.006	0.428	0.018	0.350	-0.019	0.424

Notes: All estimated discontinuities are from linear models that exclude loans received between the LBC loan and the minutes post-LBC. Estimates are from specifications with a triangular estimation kernel, and an optimal bandwidth selected from a 12-hour window around the LBC. p-values are from the heteroskedasticity-= and bias-correction robust standard errors, calculated using the nearest-neighbor variance estimator with a minimum of three matches.

Finally, in Table A3, we test for whether observable borrower characteristics is smooth through the 20-minute post-LBC cutoff. We again use the rdrobust optimal bandwidth approach. Note that while these borrower characteristics are all fixed at the individual level (we only observe loan amount and length for the first loan), the unit of observation is loan: a particular borrower can experience both sides of the cutoff. Therefore, we use the full sample of loans here. We fail to measure any significant or large jump at the cutoff for any variables. We thus use the 20-minute post-LBC latent cutoff as our preferred specification.

Table A3: **Borrower variable smoothness at 20-minute post-LBC cutoff**

N = 11,512	Coef.	S.E.	p-value	Effect size	Optimal BW	Obs. in BW
Age	0.525	0.534	0.341	1%	[148,98]	7,436
Female	-0.008	0.030	0.989	-2%	[119,102]	6,828
Married	-0.002	0.035	0.787	-0%	[92,144]	6,641
Dependents	-0.053	0.081	0.416	-4%	[90,109]	6,077
Log income	0.003	0.040	0.952	0%	[142,108]	7,485
Credit score	0.010	0.065	0.966	1%	[115,129]	7,182
Loan amount	-7.799	20.847	0.589	-0%	[124,122]	7,275
Loan length	0.217	0.415	0.556	1%	[123,133]	7,447

Notes: All estimated discontinuities are from linear models that exclude the LBC loan, and loans received within 20 minutes after the LBC. All models use a triangular estimation kernel, and an optimal bandwidth selected from a 12-hour window around the LBC. Heteroskedasticity-robust standard errors of the linear estimates are shown, calculated using the nearest-neighbor variance estimator with a minimum of three matches. We also report the bias-correction- and heteroskedasticity-robust p-values of the quadratic, bias-corrected estimates. The reports effect size is as a percentage of the pre-cutoff mean value of the borrower characteristic within the two-hour bandwidth. The optimal bandwidths –rounded to the nearest integer– are reported along with observations within the used bandwidth. The overall sample size for all models in the table corresponds to all loans within twelve hours of an LBC.

One test that we do not report in full is the test for density smoothness of the running variable –$DistanceToBatch$– with loans right after the LBC excluded. This is because a failure to reject here can come either from densities that match up nicely at the post-LBC cutoff, or from an increase in the standard error in the exclusion window to the left of the cutoff. Without data in the region around the cutoff, the uncertainty about the density is large. In the main analysis, with a null hypothesis of no effect of missing a batch cutoff, this simply reduces our power. However, in this analysis, where we are seeking a region where

we cannot reject smoothness, it may lead us to be too confident in the selection of a smaller exclusion window. That said, the 15-minute post-LBC cutoff is where we first fail to reject a discontinuity in density.

A.4 First stage Results

Table A4 presents estimates of the effect of missing a batch cutoff on the loan delay experience by a borrower. The estimates in column (1) are obtained from a specification that matches Figure 3: two-hour bandwidth, uniform kernel, linear estimation. In columns (2)-(4) we use optimal bandwidth selection and a triangular estimation kernel. We allow for an asymmetric optimal bandwidth because the exclusion of loans submitted within 20 minutes following the LBC creates an asymmetry in density around the post-LBC latent cutoff. The models in column (2) feature no control variables, in column (3) we add application submission day-of-week, hour-of-day and month fixed effects, and in column (4) we add borrower and loan control variables. In Panels A and C, we also add a fixed effect for a borrower's sequential loan number in column (3). Panel A shows the full sample estimate, and Panels B and C show estimates for first-time loans and repeat loans, respectively. Below each estimate, we include the effect magnitude as a percentage of the pre-cutoff mean delay within two hours of the cutoff for the relevant sample, the optimal bandwidth as determined by the rdrobust command, and the number of observations within that optimal bandwidth. The sample that is fed into the optimal bandwidth algorithm is held fixed across the specifications, but the number of observations within the optimal bandwidth varies slightly across specifications as the optimal bandwidth changes slightly when we add control variables. Because all p-values – both conventional and bias-correction robust– are less than 0.001 in all specifications, we omit them from the table. Heteroskedasticity-robust standard errors are shown in parentheses.

Table A4: **Impact of cutoff on loan delay (in hours)**

RD bandwidth:	Two-hour	Optimal		
	(1)	(2)	(3)	(4)
A. Full sample (N = 11,512)				
PostBatch	6.56	10.76	9.85	9.81
	(0.87)	(1.50)	(1.25)	(1.08)
Effect as % of pre-cutoff mean	69%	113%	103%	103%
Optimal bandwidth (mins)		[81,49]	[95,53]	[132,55]
Observations within bandwidth	7,177	4,180	4,858	5,974
B. First-time loans (N = 5,530)				
PostBatch	8.50	12.25	11.18	10.91
	(1.57)	(2.00)	(1.77)	(1.76)
Effect as % of pre-cutoff mean	49%	71%	65%	63%
Optimal bandwidth (mins)		[129,62]	[122,72]	[123,72]
Observations within bandwidth	3,090	2,626	2,683	2,695
C. Repeat loans (N = 5,982)				
PostBatch	6.71	8.60	8.34	8.25
	(0.89)	(1.27)	(1.09)	(1.09)
Effect as % of pre-cutoff mean	185%	237%	230%	228%
Optimal bandwidth (mins)		[107,66]	[118,71]	[117,71]
Observations within bandwidth	4,087	3,189	3,426	3,426
Day-of-week, hour-of-day, month FEs	N	N	Y	Y
Borrower controls	N	N	N	Y

Notes: All estimated discontinuities are from linear models that exclude the LBC loan, and loans received within 20 minutes after the LBC. Column (1) reports a specification with a uniform estimation kernel and a fixed bandwidth of 120 minutes around the 20-minute post-LBC cutoff. Columns (2-4) report specifications with a triangular estimation kernel, and an optimal bandwidth selected from a 12-hour window around the LBC. Heteroskedasticity-robust standard errors of the linear estimates are shown in parentheses below the estimates, calculated using the nearest-neighbor variance estimator with a minimum of three matches. All estimates are statistically significant with $p < 0.001$ according to both the heteroskedasticity-robust p-values of the linear estimates, and the bias-correction- and heteroskedasticity-robust p-values of the quadratic, bias-corrected estimates. We also report the estimated effect as a percentage of the pre-cutoff mean delay within the two-hour bandwidth. The optimal bandwidths –rounded to the nearest integer– are reported for the specifications in columns (2)-(4), and observations within the used bandwidth are reported below. The overall sample sizes for each panel correspond to all loans within twelve hours of an LBC. The fixed effects added in column (3) include the hour-of-day, day-of-week, and month of application submission. In Panels A and C, a fixed effect for the borrower's sequential loan number is also included. The borrower controls added in column (4) are age, age squared, sex, marital status, number of dependents, log income, and credit score.

A.5 Additional Figures and Tables

Figure A5: **Distributions of loan application submissions and loan disbursements**

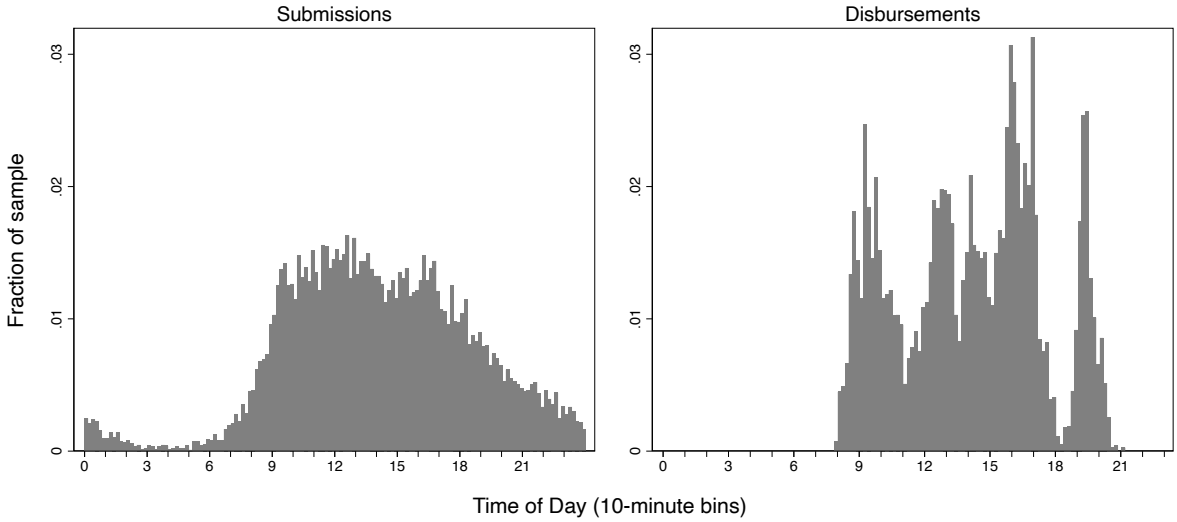
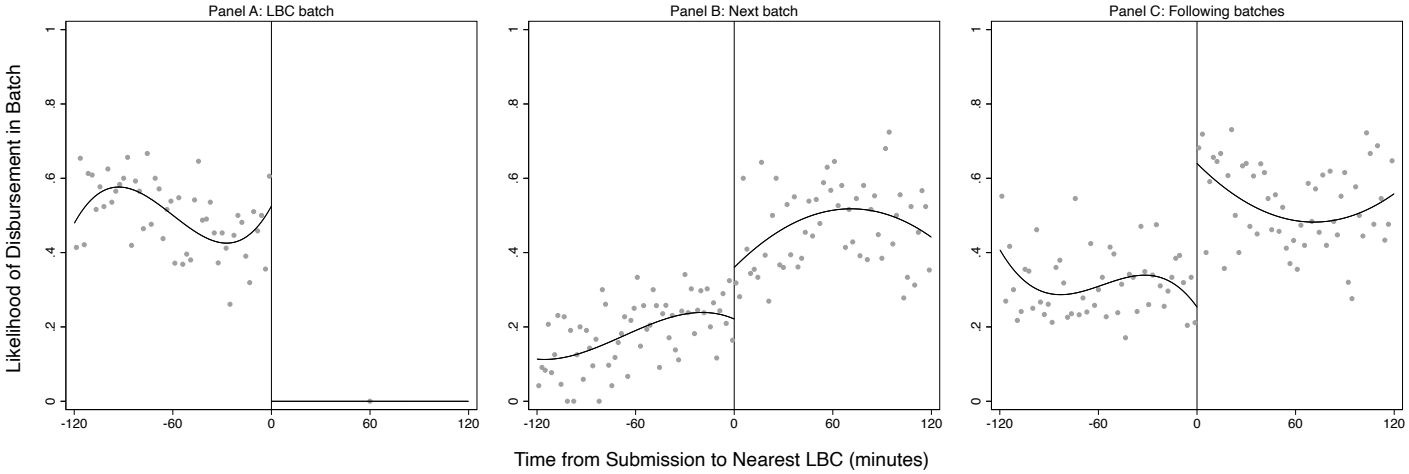
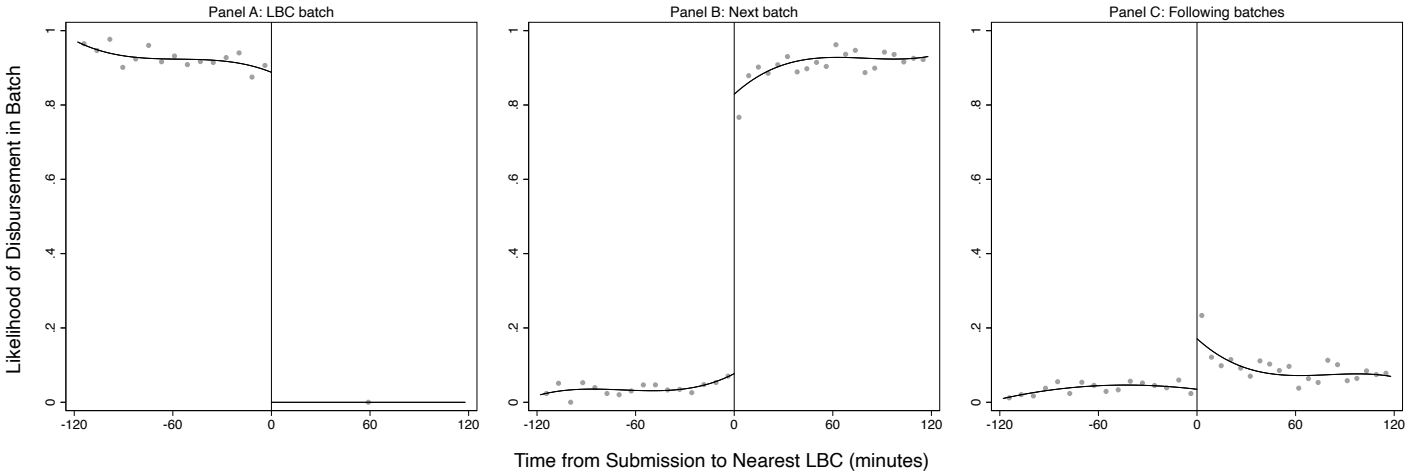


Figure A6: **Impact of cutoff on likelihood of loan processing in batch (first-time loans)**



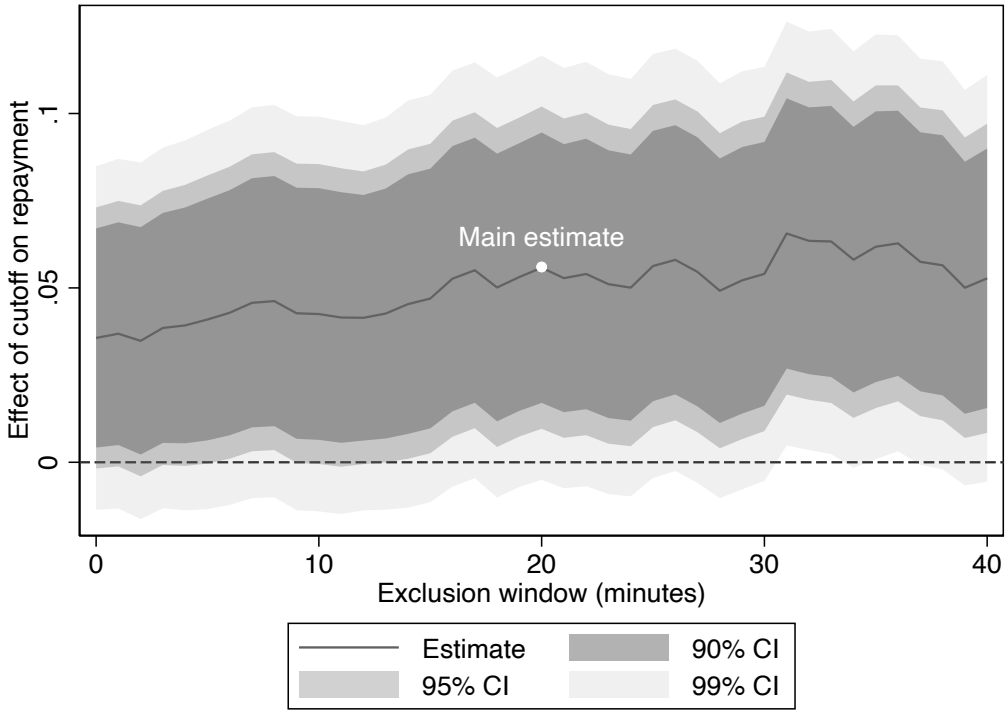
Notes: Regression discontinuity plots use third degree polynomials, a uniform kernel, and a fixed bandwidth of 120 minutes. We always exclude the LBC loans.

Figure A7: **Impact of cutoff on likelihood of loan processing in batch (repeat loans)**



Notes: Regression discontinuity plots use third degree polynomials, a uniform kernel, and a fixed bandwidth of 120 minutes. We always exclude the LBC loans.

Figure A8: **Impact of cutoff on loan repayments by post-LBC cutoff**



Notes: Estimates are from the same model as Table 1, column (4), estimated for each post-LBC exclusion window from zero to forty (in one-minute increments).

Table A5: **Summary statistics**

Variables	**Mean**	**SD**	**Min**	**Median**	**Max**
A. Borrower characteristics (N = 7,206)					
Age	37.45	9.55	20	36	65
Female	0.4	0.50	0	0	1
Married	0.49	0.50	0	0	1
Dependents	1.24	1.14	0	1	5
Monthly income (pesos)	1,718.66	8,279.59	291.67	916.67	125,000.00
Credit score - none	0.13	0.33	0	0	1
Credit score - marginal	0.30	0.46	0	0	1
Credit score - average	0.31	0.46	0	0	1
Credit score - better	0.22	0.41	0	0	1
Credit score - best	0.04	0.21	0	0	1
Credit score - linear (0-4)	1.76	1.07	0	2	4
B: First-time loans (N = 5,530)					
Amount received (pesos)	1,759.29	348.53	1,000	1,500	3,000
Loan term (days)	21.36	7.13	7	21	30
Processing time (hours)	23.63	21.32	0.60	18.07	63.10
Loan repaid	0.68	0.46	0	1	1
C: Repeat loans (N = 5,982)					
Processing time (hours)	8.95	14.84	0.15	2.99	63.10
Loan repaid	0.78	0.42	0	1	1

Notes: Borrower characteristics are collected at the time of the first loan application. Income is winsorized at the top 0.5% due to a couple extreme outliers. Loan amounts and lengths are only available for first loans. Processing times measure the time between loan application and loan disbursement. Processing times are winsorized at the top 10% due to a large right tail.

Table A6: **Borrower/loan characteristics and loan repayment**

Sample:	Full sample		First-time loans	
	(1)	(2)	(3)	(4)
Age	-0.004	-0.004	-0.008	-0.009
	(0.003)	(0.003)	(0.005)	(0.005)
	$p = 0.274$	$p = 0.218$	p =0.103	p =0.081
Age^2	0.000052	0.000055	0.000096	0.000103
	(0.000041)	(0.000040)	(0.000062)	(0.000062)
	$p = 0.211$	p =0.172	p =0.123	p =0.099
Female	0.013	0.011	0.018	0.018
	(0.009)	(0.008)	(0.013)	(0.013)
	$p = 0.151$	p =0.177	p =0.154	p =0.154
Married	-0.010	-0.011	-0.014	-0.015
	(0.010)	(0.010)	(0.015)	(0.015)
	$p = 0.292$	p =0.251	p =0.339	$p = 0.318$
Dependents	-0.007	-0.006	-0.002	-0.001
	(0.004)	(0.004)	(0.007)	(0.007)
	$p = 0.116$	p =0.142	p =0.812	p =0.937
Log monthly income (pesos)	0.023	0.020	0.013	0.011
	(0.005)	(0.005)	(0.008)	(0.008)
	$p < 0.001$	$p < 0.001$	p =0.117	p =0.172
Credit score (0-4)	0.026	0.034	0.065	0.073
	(0.004)	(0.004)	(0.009)	(0.009)
	$p < 0.001$	$p < 0.001$	p <0.001	p <0.001
Log amount received (pesos)			0.029	0.003
			(0.043)	(0.045)
			p =0.502	p =0.955
Loan term (days)			-0.003	-0.003
			(0.001)	(0.001)
			p =0.001	p =0.002
Day-of-week, hour-of-day, month FEs	N	Y	N	Y
Observations	11,512	11,512	5,530	5,530
Clusters	7,206	7,206		
Sample mean [SD]	0.733 [0.442]		0.685 [0.465]	

Notes: All estimates are from linear probability models of repayment. Columns (1) and (2) use the entire estimation sample of loans, with standard errors clustered at the borrower level. Columns (3) and (4) use only first-time loans, with heteroskedasticity-robust standard errors. In columns (2) and (4), we include fixed effects for the hour-of-day, day-of-week, and month of application submission. In column (2) the set of fixed effects also includes a borrower's sequential loan number.

Table A7: **Impact of cutoff on likelihood of same-day loan**

RD bandwidth:	Two-hour	Optimal		
	(1)	(2)	(3)	(4)
A. Full sample (N = 11,512)				
PostBatch	-0.148	-0.212	-0.231	-0.237
	(0.022)	(0.043)	(0.034)	(0.028)
Effect as % of pre-cutoff mean	-19%	-27%	-30%	-31%
Optimal bandwidth (mins)		[66,48]	[73,53]	[94,55]
Observations within bandwidth	7,177	3,582	4,097	4,873
B. First-time loans (N = 5,530)				
PostBatch	-0.191	-0.284	-0.264	-0.261
	(0.037)	(0.049)	(0.040)	(0.040)
Effect as % of pre-cutoff mean	-34%	-50%	-47%	-46%
Optimal bandwidth (mins)		[112,55]	[125,61]	[127,61]
Observations within bandwidth	3,090	2,371	2,573	2,592
C. Repeat loans (N = 5,982)				
PostBatch	-0.161	-0.179	-0.201	-0.198
	(0.025)	(0.035)	(0.025)	(0.025)
Effect as % of pre-cutoff mean	-17%	-19%	-21%	-21%
Optimal bandwidth (mins)		[95,74]	[108,89]	[117,71]
Observations within bandwidth	4,087	3,085	3,548	3,509
Day-of-week, hour-of-day, month FEs	N	N	Y	Y
Borrower controls	N	N	N	Y

Notes: All estimated discontinuities are from linear models that exclude the LBC loan, and loans received within 20 minutes after the LBC. Column (1) reports a specification with a uniform estimation kernel and a fixed bandwidth of 120 minutes around the 20-minute post-LBC cutoff. Columns (2-4) report specifications with a triangular estimation kernel, and an optimal bandwidth selected from a 12-hour window around the LBC. Heteroskedasticity-robust standard errors of the linear estimates are shown in parentheses below the estimates, calculated using the nearest-neighbor variance estimator with a minimum of three matches. All estimates are statistically significant with $p \leq 0.001$ according to both the heteroskedasticity-robust p-values of the linear estimates, and the bias-correction- and heteroskedasticity-robust p-values of the quadratic, bias-corrected estimates. We also report the estimated effect as a percentage of the pre-cutoff mean likelihood of same-day disbursement within the two-hour bandwidth. The optimal bandwidths –rounded to the nearest integer– are reported for the specifications in columns (2)-(4), and observations within the used bandwidth are reported below. The overall sample sizes for each panel correspond to all loans within twelve hours of an LBC. The fixed effects added in column (3) include the hour-of-day, day-of-week, and month of application submission In Panels A and C, a fixed effect for the borrower's sequential loan number is also included. The borrower controls added in column (4) are age, age squared, sex, marital status, number of dependents, log income, and credit score.

Table A8: **IV estimates of impact of loan delay on loan repayment**

	(1)	(2)	(3)
A. Full sample (N = 7,177)			
Loan Delay (hours)	0.0026	0.0043	0.0042
	(0.0013)	(0.0018)	(0.0017)
Estimate p-value	0.041	0.014	0.016
B. First-time loans (N = 3,090)			
Loan Delay (hours)	0.0026	0.0035	0.0042
	(0.0019)	(0.0027)	(0.0027)
Estimate p-value	0.172	0.193	0.123
C. Repeat loans (N = 4,087)			
Loan Delay (hours)	0.0025	0.0048	0.0046
	(0.0017)	(0.0022)	(0.0022)
Estimate p-value	0.127	0.031	0.038
Day-of-week, hour-of-day, month FEs	N	Y	Y
Borrower controls	N	N	Y

Notes: All estimates are from two-stage-least-squares models where the regression-discontinuity specification from equation 1 instruments for the experienced delay in receiving a loan (jn hours). The sample limited to a two-hour window around the 20-minute post-LBC cutoff. Heteroskedasticity-robust standard errors are shown in parentheses below the estimates. All models feature first stages with joint F-statistics that are statistically different from zero with $p < 0.001$. The fixed effects added in column (2) include the hour-of-day, day-of-week, and month of application submission. In Panels A and C, a fixed effect for the borrower's sequential loan number is also included. The borrower controls added in column (3) are age, age squared, sex, marital status, number of dependents, log income, and credit score.

Table A9: **Heterogeneity in repayment effects**

	(1)	(2)
A: Marital Status	**Single/Divorced/Widowed**	**Married**
PostBatch	0.012	0.108
	(0.032)	(0.037)
Estimate p-value	0.708	0.003
Bias-corrected estimate p-value	0.833	0.007
Effect as % of pre-cutoff mean	2%	15%
Optimal bandwidth (mins)	[142,133]	[132,101]
Observations within bandwidth	4,054	3,447
Total Observations	5,903	5,609
B: Income	**Below median**	**Above median**
PostBatch	0.028	0.079
	(0.035)	(0.031)
Estimate p-value	0.424	0.011
Bias-corrected estimate p-value	0.492	0.023
Effect as % of pre-cutoff mean	4%	11%
Optimal bandwidth (mins)	[142,121]	[148,121]
Observations within bandwidth	4,017	4,017
Total Observations	5,876	5,876
C: Credit Score	**None/Marginal/Average**	**Better/Best**
PostBatch	0.033	0.142
	(0.028)	(0.044)
Estimate p-value	0.251	0.001
Bias-corrected estimate p-value	0.333	0.004
Effect as % of pre-cutoff mean	5%	18%
Optimal bandwidth (mins)	[148,125]	[117,81]
Observations within bandwidth	5,599	1,834
Total Observations	8,140	3,372

Notes: All estimated discontinuities are from linear models that exclude the LBC loan, and loans received within 20 minutes after the LBC. All estimates are from specifications with a triangular estimation kernel, and an optimal bandwidth selected from a 12-hour window around the LBC. Heteroskedasticity-robust standard errors of the linear estimates are shown in parentheses below the estimates, calculated using the nearest-neighbor variance estimator with a minimum of three matches. We report both the heteroskedasticity-robust p-values of the linear estimates, and the bias-correction- and heteroskedasticity-robust p-values of the quadratic, bias-corrected estimates. We also report the estimated effect as a percentage of the pre-cutoff mean repayment rate within the two-hour bandwidth. The optimal bandwidths –rounded to the nearest integer– are reported, observations within the used bandwidth are reported below, and all observations within twelve hours of an LBC below that. All estimates feature fixed effects for the hour-of-day, day-of-week, month of application submission, and the borrower's sequential loan number. All estimates feature controls for age, age squared, sex, marital status, number of dependents, log income, and credit score. These controls drop out when they are the heterogeneous variable of interest.