DISCUSSION PAPER SERIES

DP13884
(v. 2)

CREDIT BUILDING OR CREDIT CRUMBLING? A CREDIT BUILDER LOAN'S EFFECTS ON CONSUMER BEHAVIOR, CREDIT SCORES AND THEIR PREDICTIVE POWER

Jeremy Burke, Julian C. Jamison, Dean Karlan, Kata Mihaly and Jonathan Zinman

FINANCIAL ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

CREDIT BUILDING OR CREDIT CRUMBLING? A CREDIT BUILDER LOAN'S EFFECTS ON CONSUMER BEHAVIOR, CREDIT SCORES AND THEIR PREDICTIVE POWER

Jeremy Burke, Julian C. Jamison, Dean Karlan, Kata Mihaly and Jonathan Zinman

Discussion Paper DP13884
First Published 22 July 2019
This Revision 15 October 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CREDIT BUILDING OR CREDIT CRUMBLING? A CREDIT BUILDER LOAN'S EFFECTS ON CONSUMER BEHAVIOR, CREDIT SCORES AND THEIR PREDICTIVE POWER

Abstract

How does the large market for credit score improvement products affect consumers and market efficiency? For consumers, we use a randomized encouragement design on a standard credit builder loan (CBL) and find null average effects on scores. But a generalized random forest algorithm finds important heterogeneity, most starkly with respect to baseline installment credit activity. CBLs induce delinquency on pre-existing loan obligations, suggesting that even a seemingly modest additional claim on monthly cash flows is too much for many consumers to manage. For the market, CBL take-up reveals information: takers experience future score improvements relative to non-takers, which, given null average treatment effects, implies positive selection. However, we find suggestive evidence that the CBL weakens the score's power for predicting default in some cases. We propose simple changes, to CBL provider strategy and credit bureau reporting categories, that could produce more uniformly positive effects for both individuals and the market.

JEL Classification: D12, G14, G21

Keywords: subprime, thin file, credit scoring, screening, credit invisibles, household finance, consumer finance

Jeremy Burke - jeremy.burke@usc.edu
University of Southern California

Julian C. Jamison - julison@gmail.com
University of Exeter

Dean Karlan - dean.karlan@gmail.com
Northwestern University and CEPR

Kata Mihaly - kmihaly@rand.org
Rand Corporation

Jonathan Zinman - jzinman@dartmouth.edu
Dartmouth University

Acknowledgements
The credit bureau that provided data to us had the right to: (a) review the paper to ensure that the analysis using credit scores was depersonalized, aggregated, and that the scores received the correct trademark attribution; (b) offer comments about the paper, which the authors agreed ex-ante to consider in good faith. We thank: the St. Louis Community Credit Union, and especially Paul Woodruff, for cooperation; the Consumer Financial Protection Bureau (CFPB), in particular Sarah Bainton Kahn and Daniel Dodd-Ramirez for their assistance; participants at many conferences and seminars for comments; Innovations for Poverty Action, in particular Anna Cash, Lucia Goin, Nora Gregory, and Kayla Wilding, for research support. We gratefully acknowledge research funding provided by the Consumer Financial Protection Bureau (under competitive award CFP-12-Z-00020/0002). The views expressed are those of the authors alone and are not necessarily shared by the CFPB or any other arm of the U.S. government. Institutional Review Board approval for human subjects protocols from Innovations for Poverty Action (no. 14January-001) and RAND Human Subjects Protection Committee (no. 2013-0660). This study was registered with the AEA RCT Registry with the ID number AEARCTR-0000441.

Credit building or credit crumbling?
A credit builder loan's effects on consumer behavior,
credit scores and their predictive power[*]

Jeremy Burke, Julian Jamison, Dean Karlan, Kata Mihaly and Jonathan Zinman

October 2020

ABSTRACT

How does the large market for credit score improvement products affect consumers and market efficiency? For consumers, we use a randomized encouragement design on a standard credit builder loan (CBL) and find null average effects on scores. But a generalized random forest algorithm finds important heterogeneity, most starkly with respect to baseline installment credit activity. CBLs induce delinquency on pre-existing loan obligations, suggesting that even a seemingly modest additional claim on monthly cash flows is too much for many consumers to manage. For the market, CBL take-up reveals information: takers experience future score improvements relative to non-takers, which, given null average treatment effects, implies positive selection. However, we find suggestive evidence that the CBL weakens the score's power for predicting default in some cases. We propose simple changes, to CBL provider strategy and credit bureau reporting categories, that could produce more uniformly positive effects for both individuals and the market.

Keywords: subprime; thin file; credit scoring; screening; credit invisibles; household finance; consumer finance

JEL Codes: D12; G14; G21

[*] Jeremy.burke@usc.edu, julison@gmail.com, karlan@northwestern.edu, kmihaly@rand.org, jzinman@dartmouth.edu. The credit bureau that provided data to us had the right to: (a) review the paper to ensure that the analysis using credit scores was depersonalized, aggregated, and that the scores received the correct trademark attribution; (b) offer comments about the paper, which the authors agreed ex-ante to consider in good faith. The authors retained intellectual freedom to report the results of this study regardless of the outcomes. We thank: the St. Louis Community Credit Union, and especially Paul Woodruff, for cooperation; the Consumer Financial Protection Bureau (CFPB), in particular Sarah Bainton Kahn and Daniel Dodd-Ramirez for their assistance and preparation of a policy brief based this study; participants at many conferences and seminars for comments; Innovations for Poverty Action; and, in particular, Anna Cash, Lucia Goin, Nora Gregory, Peter Lugthart, and Kayla Wilding for research support. We gratefully acknowledge research funding provided by the Consumer Financial Protection Bureau (under competitive award CFP-12-Z-00020/0002). The views expressed are those of the authors alone and are not necessarily shared by the CFPB or any other arm of the U.S. government. Institutional Review Board approval for human subjects protocols from Innovations for Poverty Action (no. 14January-001) and RAND Human Subjects Protection Committee (no. 2013-0660). This study was registered with the AEA RCT Registry with the ID number AEARCTR-0000441.

1. Introduction

Consumer credit histories are important inputs to various markets. Lenders use them in determining willingness to ration or lend, and at what terms.[1] Many landlords, insurers, and employers now use them when evaluating potential customers or employees.[2] Yet about 20% of the U.S. population lacks a traditional credit score due to thin or non-existent credit bureau files. For these "credit invisibles", and many more, there is much information beyond standard credit histories that credit risk modelers could and increasingly do use, including payment behavior on non-traditional products like the credit builder loans (CBLs).[3]

Many programs and products, like CBLs, aim to help consumers signal or improve creditworthiness. These also include financial education, financial coaching and credit counseling programs, and credit repair and credit monitoring services. Whether they truly help consumers is an open question.[4] So too is whether they enhance market efficiency by revealing unobservable information or worsen market efficiency by providing misleading signals. CBLs have become an increasingly common approach. They are widely available, and prominent financial self-help resources like NerdWallet and Credit Karma provide advice on how to access and manage them. The main CBL suppliers are credit unions and community banks, with financial technology companies like Self Inc. entering the market recently.

CBLs are short-term installment contracts on small amounts (typically $500 to $1,000, repaid monthly over 6-24 months) in which the "lender" eliminates its credit risk by inverting the sequence of origination and repayment: "loan" proceeds are held in an escrow account and only released after the contracted payments, which include principal and an administrative fee, are

[1] A majority of credit users in the USA have below-prime credit scores (Brooks et al. 2015), and below-prime credit usage typically increases borrowing costs by several percentage points and hundreds or thousands of dollars per year (Pulliam Weston 2010; Zinman 2015).
[2] See, e.g., Consumer Financial Protection Bureau (2012), Bartik and Nelson (2020), Bos et al. (2018), Clifford and Shoag (2020), Dobbie et al. (2020).
[3] Other examples of relatively novel information sources include alternative credit bureaus focusing on small-dollar products that are largely missing from the "Big Three" traditional bureaus, utility and rent payments, and social media. See, e.g., Brevoort et al. (2015) and Brevoort and Kambara (2017).
[4] We are not aware of any prior randomized evaluations of credit building *products*. Several RCTs have estimated effects of *programmatic* interventions like financial education and counseling on credit behaviors and scores, and Kaiser et al.'s (2020) meta-analysis infers a mean effect size of 0.04 standard deviations. One issue with programmatic interventions for adults is low take-up, as discussed below.

made.[5] The CBL thus operates less like a loan and more like either a costly commitment savings device (if individuals do not withdraw the funds from the restricted account) or a costly sequence of deposits and withdrawals (if individuals choose to withdraw the funds immediately after making each payment). Nevertheless, and critically, credit reporting treats CBLs as standard installment loans, per industry agreements between CBL providers and the three major credit bureaus. And as with standard loans, CBL providers report all CBL payment performance to the bureaus, both positive (<30 days late) and negative (>=30 days late).

Like any credit building intervention, CBLs could have impacts on consumers, lenders, and markets alike. For consumers, CBLs could help them become credit visible, or shift their credit scores up or down. Our descriptive evidence suggests that both shifts likely occur with some frequency; e.g., 40% of CBL users in our sample pay more than 30 days late on their CBL at some point. For lenders, CBLs provide marginal customers a point of entry or re-entry into the mainstream financial system, opening the possibility of cross-sells. For the market, via CBL providers reporting to credit bureaus, CBLs could help or harm market efficiency. If CBL take-up predicts downstream behavior in ways that are not fully captured by other observables, market efficiency could improve. If the post-CBL credit score is a less predictive measure of behavior than the pre-CBL credit score, market efficiency could worsen.

We start by estimating CBL treatment effects on consumers, and on lender cross-sells, using an encouragement design that randomizes take-up requirements. St. Louis Community Credit Union (SLCCU) has offered CBLs since 2009 and worked with the research team from September 2014 through February 2015 to identify a sample of over 1,500 SLCCU members who expressed interest in a CBL. Nearly 20% of our sample lacked a FICO® score at baseline, and scores are low

[5] Operationally, the lender typically first disburses loan proceeds in whole to a locked savings account it controls, and then releases proceeds to the borrower, either in parts after each of the borrower's payments or in whole after the borrower makes all of the payments. In our setting the lender releases proceeds after each payment. This setup imposes modest liquidity demands on the CBL user, who need only come up with $54 on the payment due date and can get $50 of the $54 back within minutes of making the payment. Credit unions tend to calibrate the fee to roughly cover the cost of the staff time required to administer the CBL, with the intent of generating returns downstream through cross-sells and/or helping their membership (credit unions are mutually-owned and often operate like nonprofits).

overall (baseline sample mean FICO® of 560, sd 65 points, compared to a national average about 700), making our sample suitable for studying CBLs.[6]

We then randomly assigned these individuals to one of two arms: a "CBL Arm" that followed SLCCU's standard enrollment process for a CBL, and an "Extra Step Arm" facing an additional requirement to complete five modules of online financial education, taking about 50-60 minutes in total, either onsite or offsite. Only six individuals in the Extra Step Arm even started the online financial education, and thus it should have no treatment effect.[7] The CBL Arm had a take-up rate of 30% within 18 months of entering the study, while the take-up rate in the Extra Step Arm was only 12%. The financial education requirement thus strongly deterred CBL adoption (this has its own policy implications, as we discuss below).

We measure FICO® Scores and credit market behaviors using four data pulls obtained from one of the three major credit bureaus: one at baseline, and three more at endlines of roughly 6 months, 12 months, and 18 months post-random assignment. Our two main outcomes are whether the consumer has a FICO® Score, and their score conditional on having one at baseline. Having a credit score is an important step for consumers in becoming credit-visible and potentially signaling a positive credit history. It is also an important step for lenders and the market in the sense that a scoring company only reports a consumer's score when it has sufficient confidence in its predictive power. The numerical credit score itself is important, as discussed above, because of its widespread use in credit and other markets.

Averaging across the three endlines, we find a null average intent-to-treat effect of the CBL on the likelihood of having a credit score (1.8pp, se 1.5pp, Extra Step Arm baseline mean 0.84). We also find a null average treatment effect on the credit score (-2 points, se 3 points, Extra Step Arm baseline mean 561), among the subsample of individuals with a credit score at baseline.

[6] FICO® is a registered trademark of the Fair Isaac Corporation. For more information on aggregate score trends and distributions see, e.g., https://wallethub.com/edu/cs/average-credit-scores/25578/ , dated May 6 2020 and accessed September 12, 2020.

[7] The financial education content did not include anything specifically about credit builder loans, and participants were not informed about the content of the financial education modules at the time of randomization: they were simply told they needed to "complete five online lessons" that would take about an hour or less.

These null average effects obscure important heterogeneity in treatment effects (HTEs), most starkly by baseline installment credit activity. We are motivated to examine this margin of heterogeneity by theory, practice, and machine learning estimation that is designed to “let the data speak”.

In theory, those with existing loans may benefit less from CBLs since they already have a recent credit history. Those with existing installment loans may struggle to manage their existing loan obligation(s) in tandem with a CBL, e.g., if learning and/or behavioral considerations are important. On the other hand, those with existing loans may have experience and/or better access to liquidity that helps them successfully manage the CBL.

In practice, baseline installment borrowing is prevalent (over 60% at baseline), and readily observable. Should it drive treatment effects, any CBL provider could market and screen on it.

We let the data speak in two steps. First, we use a causal forest aggregate test for overall treatment effect heterogeneity. This test rejects the hypothesis of homogeneous treatment effects on credit scores at the first two endlines (p-values of 0.00 at 6-months and 0.05 at 12-months). Second, we examine readily observable potential correlates of the causal forest’s predicted conditional average treatment effect (CATE) for each consumer and find that the data strongly reject the hypothesis of homogeneity with respect to baseline installment activity. Most strikingly, those in the bottom tercile of the distribution of installment credit activity at baseline have a mean CATE on their 6-month credit score of +23 points (se 7 points), while those in the top tercile have a mean CATE of -13 points (se 6 points). We examine many other potential drivers of HTEs, but none is as robustly significantly correlated with CATEs in statistical or economic terms.

These HTEs imply economically significant magnitudes, whether in intent-to-treat units that identify effects of CBL access or treatment-on-the-treated units that identify effects of CBL usage. We report the former because they are more cleanly interpretable in our setup; a rough estimate of the latter would inflate the ITT coefficients by roughly a factor of five.[8] But even the ITT effects

[8] We get this multiplier by scaling up the ITT estimates by the reciprocal of the differential take-rate between the two experimental arms: 1.00/0.18 = 5.56.

on baseline installment activity sub-groups, or the ITT difference across these sub-groups, are large enough to move someone across credit score bins that affect market access and terms.[9]

This heterogeneity could be due to differential consumer behavior. Or it could be due to differential firm behavior, with FICO scoring the same behavior differently for people with different credit histories. We cannot test the latter hypothesis, as we are not privy to the proprietary model behind the FICO® Score. We can test the former hypothesis, and do find evidence supporting it. Specifically, we find HTEs on two categories of behaviors that factor into credit scoring: credit mix and repayment performance. The repayment performance results are the most striking, with no evidence of TEs on delinquency for those in the lower two terciles of baseline installment activity, but 0.22 sd more delinquency (se 0.08 sd) for those in the CBL arm and the top tercile. The bulk of this effect is likely driven by *non*-CBL delinquency. Thus even though the CBL studied here imposes minimal liquidity constraints in principle, adding a CBL to existing credit obligations seems too much for many borrowers to manage successfully in practice.[10] We lack the ability to identify why, but offer some speculation, and guidance for future work testing models of consumer decision making, in the Conclusion.

Turning to treatment effects on other SLCCU products (cross-sells), there is some evidence that the CBL increases savings balances (average treatment effect is $248, se $121). This is consistent with some consumers using the CBL for what it is, functionally, aside from the credit reporting: a costly commitment to save. But this inference is not robust to other functional forms of savings balances, or to counting checking account balances together with savings account balances. For other SLCCU outcomes, we find no evidence of effects on customer retention, and some evidence that non-CBL borrowing from SLCCU increases for those in the bottom tercile of baseline installment activity (4.9 pp, se 2.7).

[9] See, e.g., one of the three main credit bureau's description of credit score bins here, https://www.experian.com/blogs/ask-experian/infographic-what-are-the-different-scoring-ranges/ (June 23, 2020, accessed October 6, 2020). The implied treatment-on-the-treated (ToT) effects would be large enough to move someone across multiple score bins. E.g., consider someone with the baseline mean credit score in our sample, 560. Increasing their score by 23x5 = 115 points—our rough ToT estimate for those with low baseline installment activity at baseline—moves them from being clearly "sub-prime" (a "very poor" score, 579 and below), to "prime" (670 and above).

[10] Footnote 5 and Section 2-A elaborate on liquidity requirements. We attempted to engage participants in qualitative follow-up discussions to better understand participants' experiences with the CBL, particularly regarding cash flow management, but we were stymied by a low response rate.

Last, but not least, we examine impacts of the CBL on market information, using three different predictive tests.

Our first predictive test focuses on self-selection: on whether CBL take-up reveals information about a consumer's future credit score. We find that CBL takers, relative to non-takers in the CBL Arm, show estimated credit score improvements of 13 points (se 4 points) or 17 points (se 5 points) over the post-treatment months, depending on specification. In theory, this upward trend is a combination of selection and the CBL average treatment effect. In practice, since the average ITT effect is a precisely estimated zero, the upward trend reveals strong positive (advantageous) selection: those who choose to open a CBL are improving irrespective of the CBL itself. This suggests that CBL take-up provides a valuable signal to lenders, and that credit bureaus should consider reporting CBLs as a distinct category rather than lumping them together with standard installment loans.

Our second and third predictive tests examine whether CBL-influenced credit scores at 12-months are better or worse predictors of default at 18-months than 12-month scores less-influenced by the CBL. For these tests we focus on the margins where we *do* see treatment effects on credit scores, by comparing gradients and fits within baseline installment credit activity groups across the CBL vs. Extra-Step arms. The results yield some cause for concern that CBL weakens the predictive power of the credit score in cases where it causes the score to decline. How much cause for concern is an open question, as these gradient and fit tests are not as well identified as the rest of our analyses.

All told, we add to extant literatures in several respects. First, we use random variation to help separately identify CBL selection and treatment effects.[11] Our experimental findings deliver some surprising results and implications, principally that CBL providers should consider screening out consumers with existing installment loan obligations.[12] Second, and closely related to the first, our findings that a CBL with modest liquidity requirements causes delinquency on *non*-CBL loans, at least for those with pre-existing installment debt, adds to work on consumer liquidity constraints, cash flow management, and financial distress (e.g., Gelman et al. 2018; Olafsson and Pagel 2018; Dobbie and Song 2020). Third, we replicate and expand on the key finding from prior CBL

[11] See also Liberman et al (2020) on signaling and treatment effects in the U.K. payday loan market.
[12] We discuss other approaches to CBL-seeking consumers with existing loan obligations in the Conclusion.

studies—CBL usage is advantageously selected (Chenven 2014; Wolff 2016)—and infer that credit bureaus could better harness this information revelation by reporting CBLs as a distinct product category. We thereby build bridges to work on credit history as a public good that may lead for-profit firms to under-invest in information acquisition (e.g., Petersen and Rajan 1995), and on whether and how credit bureaus reduce asymmetric information (e.g., de Janvry, McIntosh, and Sadoulet 2010; Hertzberg, Liberti, and Paravisini 2011; Manso 2013; Garmaise and Natividad 2017; Kovbasyuk, Larinsky, and Spagnolo 2019). Fourth, our findings suggest that "product-linked" financial education requirements may be counterproductive, despite strong policy and programmatic interest in that approach (Askari 2009; Sledge, Gordon, and Kinsley 2011; Reyes et al. 2013).

2. Study setting and design

A. Implementing partner and credit building product

We partnered with St. Louis Community Credit Union (SLCCU) to design and implement our study. SLCCU, a certified Community Development Financial Institution (CDFI), serves approximately 51,000 members who live or work in the greater St. Louis area. SLCCU has 11 branches (including three located within social service agencies), provides access to online financial education and phone-based credit counseling and education, and offers numerous financial products designed to improve members' financial stability. SLCCU has offered the "Credit Builder Loan" ("CBL") since 2009 and had originated approximately 4,400 CBLs at the onset of the study.

SLCCU markets and structures the CBL per credit union and CDFI industry standards. It markets the CBL as an opportunity to build credit history and improve credit scores (Figure 1 shows the marketing materials used by SLCCU, both in our study and routinely). The terms are such that no money changes hands at origination. Instead, the credit union places \$600 in a restricted access savings account (an escrow account, basically). Borrowers then make 12 monthly payments of approximately \$54 and the credit union releases \$50 from the restricted savings account back to the consumer's regular savings account immediately upon receipt of payment each month. As such, the payments portion of the CBL functions like a costly commitment savings account, yielding a certain and negative pecuniary return on saving; e.g., if the consumer makes all 12 CBL payments and does not make any withdrawals, they will have invested $648 over the course of the year and yielded $600 at year's end.

CBL payments, both timely and late, are reported to each of the three major credit bureaus as a standard installment loan, using standard definitions of delinquency (e.g., a loan is first reported delinquent if a payment is more than 30 days late). According to SLCCU policy, if a delinquent CBL borrower does not bring her CBL current within 10 days of the delinquency, the credit union closes out the loan by transferring the restricted portion of the loan amount to pay off the remaining principal balance, "successfully" paying off the loan from a credit bureau perspective.

Approximately 40 percent of CBL users in our sample made at least one payment more than 30 days late (Figure 2). This high rate of delinquency indicates that CBLs could backfire, at least for some borrowers.

B. Data

We have three data sources: a baseline survey, SLCCU accounts, and FICO® Scores and credit report attributes from one of the three major credit bureaus. Surveyors administer the baseline survey as part of the CBL marketing process, as described below. The survey captures demographics, some aspects of financial status, and attitudes. SLCCU administrative data is pulled monthly for everyone in our sample. These data capture CBL performance and usage of other loan and deposit products.

The bureau data capture snapshots of borrowing and repayment activity and one widely-used credit score, the FICO® Score.[13] We obtain snapshots at baseline (on a biweekly rolling basis as participants entered the study), at approximately 6 and 12 months post-random assignment, and at >=18 months post-assignment (with a maximum of 24 months, depending on assignment date). The credit bureau did not share loan-level data; e.g., our measure of 30-day delinquency is the number of loans, include any CBL, on which the person is >=30 days late. Some bureau variables are disaggregated to the person*loan type-level—e.g., number and balance of installment or revolving loans—but not delinquency. CBLs are reported as installment loans, both in our data and in the credit reports visible to lenders and other firms.

[13] Per a research agreement with the bureau, we obtain the data through "soft" credit pulls that do not impact credit scores and do not get access to clients' entire reports (hence the lack of "tradeline"-level data).

C. Sampling and experimental design

Figure 2 illustrates our sampling and experimental design. Our goal for *survey sampling* was to create a sample frame of SLCCU members who are generally interested in improving their credit. Between October 2014 and February 2015, research staff ("surveyors") enrolled participants into the study at seven of the SLCCU branches. Surveyors approached individuals in the branch and first asked if they were generally interested in building their credit. Individuals responding affirmatively were escorted to a private office and asked for consent to participate in a "research study focused on credit markets and products".[14] In total, 2,310 individuals consented and started the short baseline survey. Of these 2,310 we infer that 2,269 were SLCCU members at baseline, as evidenced by a match to SLCCU administrative data.

Our goal for *the experiment* was to engineer variation in CBL take-up within a sample of SLCCU members who are interested in a CBL. After the survey, surveyors described the CBL and elicited participant interest in the CBL specifically (as distinct from credit building generally). We remove the 738 "Uninterested" individuals from the experiment sample: we do not randomly assign these individuals to an experimental arm. The remaining 1,531 expressed interest in the CBL and comprise the "experimental sample".[15] Surveyors randomized these 1,531 participants, in real-time and at the individual level,[16] into one of two arms: a "CBL Arm" that is encouraged to open the CBL on the spot, per standard SLCCU procedures;[17] or an "Extra Step Arm" that is

[14] Surveys and treatments were delivered in private spaces within the credit union branches to preserve privacy and minimize the possibility of one applicant hearing about what another applicant receives.

[15] Study participants were compensated for their time (about 15-20 minutes) with a $5 gift card to a local grocery store. SLCCU preferred paper surveys and surveyors overnighted them periodically to research team headquarters; unfortunately, one package containing about 50 surveys was lost (including some who did not receive a random assignment). Thus we have random assignment but no survey data for these 50 individuals. Missingness is balanced across the two experiment arms.

[16] Each surveyor used a random number generator on a computer provided, maintained, and monitored by the research team. We also randomly assigned two other treatments. First, an independent cross-randomization provided half the survey sample (unconditional on CBL interest) with information on phone-based credit counseling and financial education. Second, six months after opening the CBL product, half of CBL takers were invited to set up an automatic transfer from checking to savings that would start six months later, after the last CBL payment. Take up of these two treatments was 2% and 0% and thus we exclude them from the analysis

[17] If a CBL Arm member was ready to open a CBL on the spot, our surveyors would escort them to a credit union representative who would further describe the product, establish payment dates, and originate the CBL. CBL Arm members who were not immediately ready to open a CBL received three forms of follow-

encouraged to open the CBL but told they must first complete approximately 50 minutes of free, online financial education prior to opening.[18] The financial education course is one of SLCCU's standard offerings and clients can complete it from a branch computer or any other web-connected device.

D. Sample characteristics and randomization balance

Table 1 presents baseline summary statistics and randomization balance tests, on our experiment sample, for 17 key outcome variables and sources of potential heterogeneity. Columns 1 and 2 present descriptive statistics, separately for the CBL (N=789) and Extra-Step (N=742) Arms. Column 3 presents an estimate of the difference across the two arms for each variable. The overall pattern is consistent with a valid randomization: only one variable has a difference that is close to statistically significant at conventional cutoffs, and the difference on that variable (age) is economically small. A caveat is that many of the statistically null point estimates here have confidence intervals that include economically meaningful differences.

Demographically, our experiment sample is predominantly female, unmarried, and Black. Only 25% of our sample has a college degree. Mean age is about 43, with an sd of 15, and the support of its distribution spans most working ages.

In terms of credit history, a bit more than 80% of our sample has a FICO® score at baseline. Table 2's transition matrices show that most movement on this variable goes in the direction of obtaining a score: nearly 50% of those unscored at baseline are scored at the 18-month endline, while only about 4% of those scored at baseline lack a score at the 18-month endline. A consumer can have a credit report with information on specific debts, without being scored, if FICO cannot estimate risk with sufficient confidence. Returning to Table 1, scores are low on average among those with scores, albeit with substantial heterogeneity: the mean is about 560 and the sd about 65. FICO® Scores can range from 300 to 850, and most of our sample is well below the cutoffs for a "prime" borrower (usually 640 or 680). Sub-prime consumers typically face high prices and rationing (see e.g., the evidence on utilization in the next paragraph). Many individuals have

up: nudges from a teller any time they transacted in a branch; phone calls attempting to set up an appointment to open a CBL; and two emails.

[18] Participants could satisfy the requirement by completing five (or more) modules out of eight available: Savings and Investments, Mortgages, Overdraft Protection, Payment Types and Credit Cards, Credit Scores and Reports, Identity Protection, Insurance and Taxes, and Financing Higher Education

substantial past borrowing experience, with a mean and sd of lifetime loans of about eight each. And many individuals have outstanding loans at baseline: over 60% have one or installment loans, and over 45% have one or more revolving loans.[19] Nearly 50% of these borrowers have been delinquent during the past 12 months.

Focusing next on liquidity, liquid asset holdings at SLCCU are low for most of the sample: 64% holds less than the required CBL monthly payment amount ($54) in their SLCCU deposit accounts at baseline.[20] And among those with an open credit line at baseline, mean utilization is greater than 100%: the average person with a revolving credit line in our sample has exceeded their credit line(s). Together with prevalent low credit scores and delinquency, these patterns suggest that liquidity constraints bind for most of our sample.

3. Results

A. CBL take-up and average treatment effects

Our randomization induced large differences in CBL take-up, defined as opening a CBL within 18 months of entering the study.[21] Table 3 presents the average treatment effects (ITT) on take-up (column 1), having a credit score (column 2) and credit score (column 3).

Recall that the main proximate goal of a CBL is to help consumers improve their credit scores. We examine whether and how CBLs achieve this goal by using the four credit reports we have per-person, and our *random assignment* to either the CBL or Extra-Step Arm, to estimate intent-to-treat (ITT) effects using OLS equations of the following form:

$$(1) \qquad Y_{it} = \alpha + \beta(CBL\ Arm \times Post_t) + \gamma Post_t + \sum_i \delta_i I_i + \varepsilon$$

[19] The traditional credit bureaus have broad but not entirely comprehensive coverage of borrowing, so some people we classify as non-borrowers may in fact have an outstanding loan.
[20] The 1/0 variables for baseline borrowing activity, delinquency, and liquid assets are not shown in Table 1 because they are each part of broader indices that are shown.
[21] Approximately 50 percent of take-up occurred on the same day as the survey and offer, 71 percent occurred within the first 30 days, and 97 percent occurred within the first year. Appendix Table 1 shows our key baseline characteristics do not have strong univariate correlations with take-up, with one potentially noteworthy exception being that takers in the CBL arm have lower credit scores than non-takers (-14 points, se 6).

Here Y is a credit report variable for person i at time t, where t includes the baseline and the three endlines (pulled roughly 6, 12, and 18 months post-random assignment). *CBL Arm*=1 if i was randomly assigned to that arm; the Extra-Step Arm is the omitted category. The CBL interaction with *Post* identifies the average effect of CBL access across the three endlines. Because we have multiple observations per person we include person fixed effects I_i (thereby absorbing the main effect $CBL\ Arm_i$) and cluster standard errors at the person level (the unit of randomization).

The strong first stage (Column 1, 18pp differential take-up rate with an se of 2pp) serves two purposes. The first is methodological: it enables us to estimate the causal effects of CBL access (in Section 3-B). The second is substantive: it sheds light on the deterrent effect of financial education, even when financial education is offered through a convenient delivery channel and at a seemingly opportune moment. The financial education requirement serves as a deterrent even though it was not enforced: only two of the 86 takers in the Extra Step Arm completed the requirement, because credit union staffers had the discretion to waive it. This is important, as it also removes the possibility that the Extra Arm group benefited from the financial education.

The average treatment effect is null on each of the primary outcome variables. Column 2 shows a 1.8pp point estimate of the CBL ITT effect on the likelihood of having a FICO® score, on a base of a control mean of 87% across the follow-up period. The standard error of 1.5pp implies that the confidence interval includes meaningful but not large effects on the extensive margin of scoring, at least in ITT terms. Column 4 shows a -1.9 point estimate of CBL's effect on the FICO® score, conditional on having a score at baseline, on a base of 567. The standard error of 2.7 points implies a fairly precisely estimated zero in ITT terms. Columns 3 and 5 disaggregate the treatment effect by endline and show no strong evidence of differences or dynamics across endlines.

B. Heterogeneous treatment effects

The null average treatment effects mask important heterogeneity. To examine heterogeneity, we first chose an extensive set of model "inputs"—potential sources of HTEs—for a machine learning model to search across. In doing so we grouped correlated baseline variables into indices, to reduce collinearity and preserve degrees of freedom. The notes to Table 4 detail the inputs.

We then test for overall (sometimes referred to as "aggregate" or "omnibus") heterogeneity with a generalized random forest model (Wager and Athey 2018; Athey and Wager 2019; Athey, Tibshirani, and Wager 2019). Table 4 Panel A reports the coefficient and p-value for each of the model's two key test statistics, separately for each outcome-endline combination. The Mean Forest Prediction tests whether the model predicts the outcome accurately: a substantial deviation from 1 is cause for concern, but we find no such evidence across any of the outcome-endline combinations. The Differential Forest Prediction tests the null of treatment effect heterogeneity, and which we reject for the continuous score outcome at 6 months (p=0.003) and 12 months (p=0.05), but not at 18 months (p=0.62). For the binary outcome of having a credit score, we only find suggestive evidence of heterogeneity at 6 months (p=0.08), which is likely a by-product of there being less variation to predict-- most people already have a credit score at baseline and then keep it over time (Table 2).

Figure 4 plots the generalized random forest's predicted conditional average treatment effect (CATE) for each outcome-endline combination for each consumer. The y-axis shows the estimated treatment effect magnitude, and the x-axis orders observations by that magnitude such that the curve is weakly increasing from left-to-right. Focusing on the continuous score, the range of CATEs illustrates considerable heterogeneity at 6 months and 12 months; e.g., the 27 or so point difference between the lowest and highest TEs is economically large (as discussed above, it is important to keep in mind these that treatment-on-the-treated effects are probably on the order of 5x larger than the ITT ones reported in our exhibits). Another key inference is that these person-specific CATEs fall fairly neatly into three bins: we see about one-third of the sample with a substantial negative TE, about one-third with close to zero, and about one-third with a substantial positive TE. As such we split the sample into CATE terciles in Table 4 Panel B, and find further evidence of economically meaningful heterogeneity: at 6 months, the estimated difference in treatment effects between the top and bottom CATE terciles is 15.06 (+/- 10.98). At 12 months, the difference is 10.87 (+/- 13.32).

Table 5 examines the univariate correlates of the CATEs from the causal forest analysis. This table makes three concessions for the sake of brevity and focus. First, we focus on the continuous credit score outcome instead of the extensive margin outcome, because the omnibus test finds more evidence of HTEs on the former. Second, we focus on the 6- and 12 month-endlines, because the omnibus test does not find evidence of heterogeneity at 18 months. Third, because the analysis is univariate, the selection of correlates is guided by applicability (e.g., demographics and credit history variables that policymakers and lenders could use for targeting) as well theoretical considerations (e.g., including a liquidity index because it may capture the ease with an individual can make each month's payments).

For each of these correlates, Table 5 presents the *correlate's* mean for individuals in the lowest CATE tercile (column 1) and for individuals the highest CATE tercile (column 2), where the CATEs are estimated by the causal forest presented in Table 4 Column 4 for the 6-month endline (Panel A here) and in Table 4 Column 5 for the 12-month endline (Panel B here). Column 3 then reports the p-value of the difference between the correlate's means in Columns 1 and 2. Column 3 permits inference about whether a particular variable correlates with the CATE. But this inference does not reveal whether any correlation is economically important or spans both positive and negative predicted treatment effects. For that, we turn to columns 4-6, where we compare *CATE* means across *correlate* terciles.

A driving correlate—a key source of HTEs—should satisfy four criteria: (1) An economically important difference in the input across bottom and top CATE terciles in Columns 1 vs. 2; (2) That difference being statistically significant (p-value < conventional thresholds in Column 3); (3) An economically important difference in the CATE across the top and bottom input terciles in Columns 4 vs. 5; (4) That difference being statistically significant (p-value < conventional thresholds in Column 6).

The only correlate satisfying each of those criteria at both endlines is the installment activity index calculated from baseline credit reports, and thus we focus on this margin of HTEs in the rest of our analyses.[22] As detailed in the Data Appendix, this index is comprised of three components: number of open installment loans, any open installment loan, and the number of new credit inquiries during the previous 12 months. The latter component covers inquiries for revolving as well as installment loans, but we include it in the installment index because it is strongly correlated with the other installment index components and not with the revolving index components.

Table 5 Columns 1-3 show large differences in the baseline installment activity index across the top and bottom terciles of predicted treatment effects, with 1.32 sd less activity (p-value=0.00)

[22] Because this is our key margin of heterogeneity, Appendix Tables 2a and 2b repeat Table 1's full sample descriptive statistics and balance checks within the top and bottom terciles of baseline installment activity.

for those in the top TE tercile at the 6-month endline, and 0.35 sd less activity at the 12-month endline. Columns 4 and 5 show that those with less baseline installment activity have large positive treatment effects at each endline (23 points and 17 points, with ses of 8 points), while those with more baseline installment activity have negative treatment effects at each endline (-13 points and -10 points, with ses of 6 points). The estimated difference of 35 points at the 6-month endline has a p-value of 0.00, and the estimated difference of 27 points at the 12-month endline has a p-value of 0.01 (Column 6).

Our takeaways from Table 5 are that those less installment activity at baseline fare well with CBLs, and that those with more installment activity fare worse relatively speaking, and poorly absolutely speaking. A practical implication is that CBL providers could secure higher average treatment effects, and more uniformly positive treatment effects, with two simple and complementary strategies. First, target-market to consumers with less installment activity. Second, screen out consumers with more installment activity, or at least discourage them from taking up a CBL.[23]

The results in Table 5 raise the question of whether differences in treatment effects are due to differences in CBL-induced credit behaviors — specifically, in factors used as inputs to the FICO® scoring model. A leading alternative hypothesis is that those with different baseline installment credit activity respond similarly to the CBL, but that their similar behavior is scored differently by the model. This alternative hypothesis is viable given the limited modeling information that Fair-Isaac publicly reveals: "The importance of these categories may vary from one person to another…".[24]

Table 6 uses variants of equation (1) to estimate CBL treatment effects on credit behaviors. Columns 1-5 present estimates for behavior indices measuring four of the five behavior factors FICO states it uses in its scoring model: "New Credit", "Payment History" (delinquency), "Amounts Owed" (which includes both "Balances" and a "Utilization" measure), and "Credit Mix". (We lack a direct measure of the fifth factor behind the FICO® score, "Length of Credit History".) For each measure of each factor we present average treatment effects in Panel A. These

[23] Untabulated results suggest that providers can obtain similar results by simply targeting or screening on one component of our index: whether someone has any outstanding installment loan.
[24] https://www.myfico.com/credit-education/whats-in-your-credit-score , accessed September 23, 2020.

average effects are just for completeness and not the focus of our analysis and discussion, which is on the tests for HTEs by baseline installment activity (Panel B). Columns 6 and 7 present additional results, on CBL delinquency, which is not broken out separately in the bureau (because, as discussed above, the delinquency measure in Column 2 includes CBLs, due to reporting and data limitations) but tracked by our partner credit union.

Table 6 Panel B Column 1 shows little evidence of any HTEs for new credit activity.

But Column 2 shows large (0.20 sd) differences in TEs on delinquency between the top tercile of baseline installment activity and the other terciles. Higher values here indicate more delinquency and default, and so the HTE is driven by a deterioration in performance for the high-installment tercile (0.22 sd increase in poor performance, se 0.08 sd). Columns 6 and 7 suggest the pattern in Column 2 is not driven by the CBL itself, for two reasons. First, we do not see HTEs on CBL delinquency; in particular, there is little evidence that those in the highest tercile of baseline installment activity have higher CBL delinquency. Second, Column 6, which uses three endline snapshots of SLCCU data to measure delinquency and thereby mirrors our credit bureau data structure, shows that the magnitude of any treatment effect on CBL delinquency measured at any single point in time is small. This is because any CBL delinquency only appears "on the books" for less than a month, due to SLCCU's practice of curing any 30-day CBL delinquency with the remaining escrow balance and then immediately closing the CBL account (Section 2-A).[25]

Turning to the "Amounts Owed" factor, Panel B's Columns 3 and 4 show suggestive evidence of larger TEs on the bottom tercile than the others. Even if this pattern were statistically stronger, its implication for scoring would less clear than for the other factors. High utilization is scored negatively but there may be non-monotonicity; e.g., some middle range of utilization may be scored more favorably than none.

Panel B's Column 5 suggests large differences between in TEs on credit mix between the bottom tercile and the others. For those in the bottom tercile, CBL access increases the likelihood of having both an installment and revolving loan open by 0.12 sd (se 0.06 sd). The point estimates

---

[25] Column 7 confirms that measuring CBL delinquency across multiple SLCCU data snapshots produces average TEs on delinquency that are closer to what one would expect given the 18% take-up differential between the CBL and Extra-Step arms.

for the other terciles are negative (-0.16 and -0.09 sd) and substantially different (p-values on the difference from the bottom tercile of 0.01 and 0.04). Since having both loan types open is scored positively, this heterogeneity in credit mix could be driving the installment activity HTEs on credit scores.

Altogether, these results are consistent with the CBL inducing differential responses in credit mix and delinquency that drive the HTEs by baseline installment activity in Table 5.[26]

C. Impacts on usage of other SLCCU products

Table 7 examines CBL treatment effects on the usage of other SLCCU products, using the same specifications we use for Table 6. These results help round out the picture of how consumer financial behavior changes as creditworthiness builds (or deteriorates), on whether the CBL helps individuals build savings (SLCCU does not focus on this extensively in its marketing, but other CBL providers do), and on the bottom-line viability of CBLs from the supply-side perspective. Odd-numbered columns estimate average treatment effects for the full sample across the three endlines, and even-numbered columns estimate treatment effects separately by baseline installment credit activity terciles.

Columns 1 and 2 show no evidence of treatment effects on membership retention (1pp with se 1pp), although the confidence intervals do not rule out economically meaningful effects on attrition given that only 7% of the full sample is no longer an SLCCU member by the 18-month endline. Columns 3 and 4 show no treatment effect of the CBL on non-CBL borrowing from SLCCU on average (1pp, se 2pp, control mean 0.32), but with suggestive evidence of heterogeneity: the TE on those in the bottom tercile of baseline installment credit activity is an estimated 4.9pp (se 2.7pp) increase, while the TEs on those in other terciles are imprecisely estimated nulls (-0.1pp with ses of 3.1 and 4.0pp).

Columns 5-8 examine treatment effects on deposit account balances. These are key outcomes for understanding whether there is a flypaper effect of CBL proceeds. Positive treatment effects on balances would be consistent with members using CBL for what it is, mechanically, aside from its credit reporting feature: a costly commitment savings device. We see some evidence that CBL

[26] Appendix Table 3 shows similar results when we limit the sample to those with a credit score at baseline.

increases the level of savings balances, with the full sample result in Column 5 ($248, se $121) perhaps being driven by those in the upper terciles of installment credit activity at baseline in Column 6. But Appendix Table 4 Columns 1-3 shows this pattern is not entirely robust to alternative functional forms of savings balances. We add checking account balances together with savings in Table 7 Columns 7 and 8 and Appendix Table 4 Columns 4-6, finding imprecisely estimated null TEs on balances in these specifications. Overall, our estimates are too imprecise to yield sharp inferences on CBL effects on deposit account balances.

Summarizing Table 7, we find little evidence that the CBL backfires from the provider's perspective, and some statistically weak hints of benefits.

D. Effects on market information

Next, we investigate how the CBL affects the quality of information available to the market, with three predictive analyses.

The first analysis, in Table 8, tests for self-selection: does CBL take-up help predict someone's future credit score? The idea here is that a consumer's CBL take-up decision may reveal something about their credit risk trajectory that otherwise would be unobserved to lenders. We implement selection tests that predict each of our two main credit score outcomes by: a) restricting the sample frame to the CBL Arm, since the CBL Arm faced the usual take-up process; and b) replacing the random assignment indicator in equation (1) with an indicator for whether someone took-up a CBL. Normally this "naïve" specification would capture an unidentifiable combination of treatment and selection effects, but given a null for average treatment effects (Table 3) the naïve specification identifies selection in the full sample.

Two specifications per outcome take different but complementary approaches to identifying selection. The specification in Table 8's odd-numbered columns assumes that the relevant margin for selection on unobservables is anything not captured by baseline levels (recall that our empirical models include person fixed effects whenever we have multiple observations per person, as we do here). Even-numbered columns assume the relevant margin for selection on unobservables is anything not captured by baseline levels and trends that can vary with the baseline score level (we use Post Double Selection LASSO to select which *Post*Baseline score bin* terms to include).

Table 8 shows strong evidence of positive (advantageous) selection on CBL take-up, for both outcomes and both approaches to controlling for observables. Columns 1 and 2 show that CBL takers are 10pp or 11pp more likely (se 3pp and 2pp) to have a credit score in the endline period than non-takers in the two specifications. Columns 3 and 4 show that CBL takers who enter the sample with a credit score have scores that are 17 points (se 5) or 13 points (se 4) higher during the endline period. Figure 3 suggests that this is due to CBL takers catching up to CBL non-takers: CBL takers increase their scores on average over the first six months and then flatten, while non-taker scores remain roughly constant from baseline through the endlines.

In all, Table 8 implies that CBLs attract consumers who are on an upward trajectory that is not fully captured by baseline observables. This has market implications: lenders can use CBLs to identify consumers whose creditworthiness is about to start improving. We speculate that credit bureaus could facilitate even stronger advantageous self-selection by distinguishing CBLs from standard installment loans in their data.

The second and third predictive tests focus on whether a CBL-influenced credit score is better, or worse, at predicting default, as measured by the score's gradient (second test) and its fit (third test). CBLs might capture valuable information and thereby improve the predictive power of the credit score or distort information and thereby reduce the score's predictive power. As noted at the outset, distortion seems like a real possibility given that the CBL is not a loan in an economic sense—it functions like a commitment contract for saving—yet is reported to credit bureaus as a standard installment loan.

Our tests compare the 12-month endline credit score's default gradient or fit for the delinquency index from the 18 month-endline, across the CBL versus Extra Step arms. (The predicted outcome here is the same delinquency index we use in prior tables.) Focusing on the predictive power of 12-month endline scores allows time for the CBL to exert any salutary or distortionary influence on the predictive content of the credit score. Since the CBL is more likely to exert influence if it changes scores, we focus here on our key margin of HTEs, although we also present results on the full sample for completeness. If the CBL changes the scores' predictive power, then e.g., the *12-month score*CBL Arm*Bottom tercile baseline installment activity* coefficient or fit will differ from the *12-month score*Extra-Step Arm*Bottom tercile baseline installment activity* coefficient or fit. A caveat is that these tests may be underpowered and/or biased-against finding distortion due to two limitations with our setup. First, the six months between 12 and 18 months may not be enough time for marginal delinquencies to emerge. Second, our measures of delinquency and default include CBL delinquency, due to data limitations described in Section 2-B, and because there is mechanically more CBL delinquency in the CBL Arm by virtue of the strong first-stage, our predictive tests are biased towards finding an improvement in predictive accuracy of the credit scores (although, as discussed vis a vis Table 6, that bias might be small due to reporting and measurement nuances).

Table 9 presents results from the gradient tests. As expected, given the null average TE on credit scores, Column 1 shows no statistically significant difference in the default-score gradient across the CBL and Extra-Step arms (p-value 0.26), and the point estimate on the difference is small in economic terms: a 0.01 sd difference per 100-point change in credit score. Column 2 decomposes this average gradient for our key margin of heterogeneity and here finds economically small differences in predictive power within each of the baseline installment loan activity terciles. However, the hint that CBL weakens predictive power for those in the top tercile (a 0.02 sd flatter gradient, with a 0.08 p-value on the difference) generates some cause for concern given the aforementioned caveat that these estimates are biased in the opposite direction, i.e., against finding distortion.

Figure 5 presents the results from the fit tests. Specifically, we test whether the CBL changes the 12-month endline credit score's ability to explain the variance of our delinquency index, using receiver operator characteristic (ROC) curves. A greater area under the curve (AUC) indicates a better fit. The 45-degree line shows what the ROC curve would be if the 12-month endline credit score had no power to predict delinquency 6 months later. Because a ROC curve requires a discrete predicted (outcome) variable, we cut the delinquency index at its median, with those above the median defaulting more. We then compare the AUCs for the CBL vs. Extra Step arms, calculating standard errors and p-values using the DeLong et al (1988) method. As expected in the full sample, there is little difference in the AUCs across the CBL vs. Extra-Step arms (p-value on the difference of 0.64). We also find no evidence for distortion in the lowest-tercile installment activity group, where the fit for those in the CBL arm suggests weakly greater predictive power: 0.76 vs. 0.73 for those in the Extra-Step arm (p-value 0.65). But as with the gradient test, the fit results for the top tercile of baseline installment activity generate some cause for concern, as here the 12-month credit score explain less variance in delinquency in the CBL arm (0.84 vs. 0.89). This difference has a p-value of 0.20, but its true magnitude is likely somewhat greater and economically meaningful.

All told, our predictive tests suggest that CBLs induce self-selection on upward credit score trajectory but may weaken the predictive power of those scores. Together with the weakly positive average treatment effect on the extensive margin of scoring (Table 3 Column 2)-- on moving consumers from “credit invisible” to visible – the full picture suggests that CBLs provide valuable information to the market but that changes to CBL marketing, screening, and/or reporting may be needed to minimize distortion. We discuss such changes in the next section.

4. Conclusion

We use a randomized encouragement design and predictive modeling to examine impacts of a credit-builder loan (CBL) on borrowers, providers, and credit market information. The results are mixed, but promising.

The CBL studied here has null average treatment effects on consumer credit scores, but these average effects obscure important heterogeneity on a readily observable margin: baseline installment borrowing. Those with more activity at baseline experience large credit score drops from the CBL, while those with less obtain the intended large credit score increase. Perhaps most strikingly, our results suggest that the CBL increases overall *non*-CBL delinquency among borrowers with higher levels of baseline installment activity. Together with high delinquency rates on the CBL itself (approximately 40%), this suggests that adding CBL’s seemingly modest liquidity requirement is too much for many consumers to manage.

CBL effects at the market level also show some signs of being mixed. On the positive side, we find that CBL takers are substantially more likely to obtain or improve their credit scores over the next 6-18 months on average, conditional on their baseline score, implying that lenders can use CBLs to advantageously select borrowers who are on an upward trajectory. As such our results also illustrate how merely comparing outcomes before versus after product take-up, a common advertising strategy of CBL providers, is misleading. On the potentially negative side, we find some suggestive evidence that CBLs distort market information weakens the predictive power of the credit score in cases where it causes the score to decline.

With respect to overall efficiency, our estimates of the CBL's effects on consumers, providers, and the market suggest that CBLs could be efficient, and perhaps Pareto-improving, with some modest design changes. Providers should consider remediating or screening out those with pre-existing installment debt (both because they have negative average treatment effects, and because market efficiency may worsen for this sub-group as per our predictive tests). Credit bureaus should consider reporting CBLs as a distinct category rather than as a traditional installment loan (as they do with distinct categories for unsecured vs. secured credit cards). In short, our results suggest a path to CBL designs that make nearly everyone better off while doing little harm to anyone else.

Expanding a bit on implications for providers, we see three potential product/program design implications to explore going forward. First, it may be counterproductive to try building consumers' financial knowledge with "product-linked" financial education. We find that a modest financial education requirement decreases product (CBL) take-up by nearly 20 percentage points, even among our sample of consumers that had expressed interest in credit building generally and the CBL specifically. Second, providers should test various approaches to dealing with the possibility that CBLs backfire for those with pre-existing installment debt. Possibilities include: screening out existing borrowers; offering or requiring a scaffolded approach that focuses first on timely repayment of existing obligations and then segues into another traditional loan or CBL; offering or requiring help with cash flow management; informing and/or reminding users that they need only part with $54 for a few minutes on the payment due date, as $50 of each payment is available to be returned to the customer upon demand. Third, automation of marketing, screening, and payment functions is likely essential for CBL providers to operate at scale, as the small deal sizes required to meet consumer needs and constraints imply a high ratio of fixed costs to potential revenues. The recent emergence of fintech providers is encouraging in this regard, and it will be interesting to see whether credit unions and other providers with strong digital operations follow suit.

Testing CBL design changes, together with testing whether our results replicate, offers exciting possibilities for revealing insights into fundamental aspects of consumer decision making. The differential effects we find on baseline installment debt activity beg for particular scrutiny. Is coming up with a very short-term outlay of $54 really so disruptive to customers with pre-existing installment loans, and if so… why? And why don't consumers with a pre-existing installment loans anticipate this disruption and simply decline the CBL?

We suspect our results are best explained by a behavioral model with limited attention to future liquidity constraints and/or over-confidence about making future payments, or limited capacity to manage multiple tasks due to scarcity in time, effort, and/or attention. Perhaps such biases or limitations are more prevalent among those with more pre-existing installment activity, or more binding for those with more claims on future cash flows or more logistical claims on their time and effort. Concepts of scarcity as put forward by Mullainathan and Shafir (2013) can lead to predictions of both negative treatment effects (e.g., lack of capacity to manage one more obligation) and positive treatment effects (e.g., via tunneling and thus hyper-attention to particularly salient tasks, see Kaur et al. (2020); Lichand and Mani (2020); Ong et al. (2019)), and so a key challenge going forward is developing testable models that sharpen understanding of whether and how scarcity leads to better or worse decision making.

Altogether our results highlight some key questions for future research and policy/product development. For research, we need to better understand how to model the decision making of very resource-constrained consumers. For policy and product development, efforts to help households build stronger credit records need to consider how to target more effectively and how such efforts affect market efficiency as well as consumers.

REFERENCES

Askari, A. 2009. “Banks and Financial Education: Integrating Practice, Products, and Partnerships.” *Community Investments* 21 (2): 23–26.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. “Generalized Random Forests.” *The Annals of Statistics* 47 (2): 1148–78.

Athey, Susan, and Stefan Wager. 2019. “Estimating Treatment Effects with Causal Forests: An Application.” *Observational Studies* 5: 36–51.

Ballance, Joshua, Robert Clifford, and Daniel Shoag. 2020. “‘No More Credit Score’: Employer Credit Check Bans and Signal Substitution.” *Labour Economics* 63 (April): 101769.

Bartik, Alex, and Scott Nelson. 2020. “Credit Reports as Resumes: The Incidence of Pre-Employment Credit Screening.”

Bos, Marieke, Emily Breza, and Andres Liberman. 2018. “The Labor Market Effects of Credit Market Information.” *Review of Financial Studies* 31 (6): 2005–37.

Brevoort, Kenneth P, Philipp Grimm, and Michelle Kambara. 2015. “Data Point: Credit Invisibles.” Consumer Financial Protection Bureau.

Brevoort, Kenneth P, and Michelle Kambara. 2017. “Data Point: Becoming Credit Visibles.” Consumer Financial Protection Bureau.

Brooks, Jennifer, Kasey Wiedrich, Lebaron Sims, and Solana Rice. 2015. “Excluded from the Financial Mainstream: How the Economic Recovery Is Bypassing Millions of Americans.” Corporation for Enterprise Development.

Chenven, Sarah. 2014. “The Power of Credit Building: Credit Building Strategies for Funders.” Asset Funders Network.

Consumer Financial Protection Bureau. 2012. “Key Dimensions and Processes in the U.S. Credit Reporting System: A Review of How the Nation’s Largest Credit Bureaus Manage Consumer Data.” CFPB.

DeLong, Elizabeth R., David M. DeLong, and Daniel L. Clarke-Pearson. 1988. “Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach.” *Biometrics* 44 (3): 837–45.

Dobbie, Will, Paul Goldsmith-Pinkham, Neale Mahoney, and Jae Song. 2020. “Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports.” *The Journal of Finance* 75 (5): 2377–2419.

Dobbie, Will, and Jae Song. 2020. “Targeted Debt Relief and the Origins of Financial Distress: Experimental Evidence from Distressed Credit Card Borrowers.” *American Economic Review* 110 (4): 984–1018.

Garmaise, Mark J., and Gabriel Natividad. 2017. “Consumer Default, Credit Reporting, and Borrowing Constraints.” *The Journal of Finance* 72 (5): 2331–68.

Gelman, Michael, Shachar Kariv, Matthew D. Shapiro, Dan Silverman, and Steven Tadelis. 2018. “How Individuals Respond to a Liquidity Shock: Evidence from the 2013 Government Shutdown.” *Journal of Public Economics*, July.

Hertzberg, Andrew, Jose Liberti, and Daniel Paravisini. 2011. “Public Information and Coordination: Evidence from a Credit Registry Expansion.” *The Journal of Finance* 66 (2): 379–412.

Janvry, Alain de, Craig McIntosh, and Elisabeth Sadoulet. 2010. “The Supply and Demand Side Impacts of Credit Market Information.” *Journal of Development Economics* 93 (2): 173–88.

Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban. 2020. “Financial Education Affects Financial Knowledge and Downstream Behaviors.”

Kaur, Supreet, Sendhil Mullainathan, Frank Schilbach, and Suanna Oh. 2020. “Does Financial Strain Lower Worker Productivity?”

Kovbasyuk, Sergey, Artem Larinsky, and Giancarlo Spagnolo. 2019. “Limited Records and Credit Cycles.”

Liberman, Andres, Daniel Paravisini, and Vikram Pathania. 2020. “High-Cost Debt and Perceived Creditworthiness: Evidence from the U.K.” *Review of Economic Studies* forthcoming.

Lichand, Guilherme, and Anandi Mani. 2020. “Cognitive Droughts.”

Manso, Gustavo. 2013. “Feedback Effects of Credit Ratings.” *Journal of Financial Economics* 109 (2): 535–48.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means so Much*. New York: Times Books, Henry Holt and Company.

Olafsson, Arna, and Michaela Pagel. 2018. “The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software.” *The Review of Financial Studies* 31 (11): 4398–4446.

Ong, Qiyan, Walter Theseira, and Irene Y. H. Ng. 2019. “Reducing Debt Improves Psychological Functioning and Changes Decision-Making in the Poor.” *Proceedings of the National Academy of Sciences* 116 (15): 7244–49.

Petersen, Mitchell A., and Raghuram G. Rajan. 1995. “The Effect of Credit Market Competition on Lending Relationships.” *The Quarterly Journal of Economics* 110 (2): 407–43.

Pulliam Weston, Liz. 2010. “Lifetime Cost of Bad Credit: $201,712.” *MSN Money* (blog). February 19, 2010.

Reyes, B., E. Lopez, S. Phillips, and K. Schroeder. 2013. “Building Credit for the Underbanked: Social Lending as a Tool for Credit Improvement.”

Sledge, Joshua, Sarah Gordon, and Matthew Kinsley. 2011. “Making the Shift from Financial Education to Financial Capability: Evidence from the Financial Capability Innovation Fund.” Center for Financial Services Innovation.

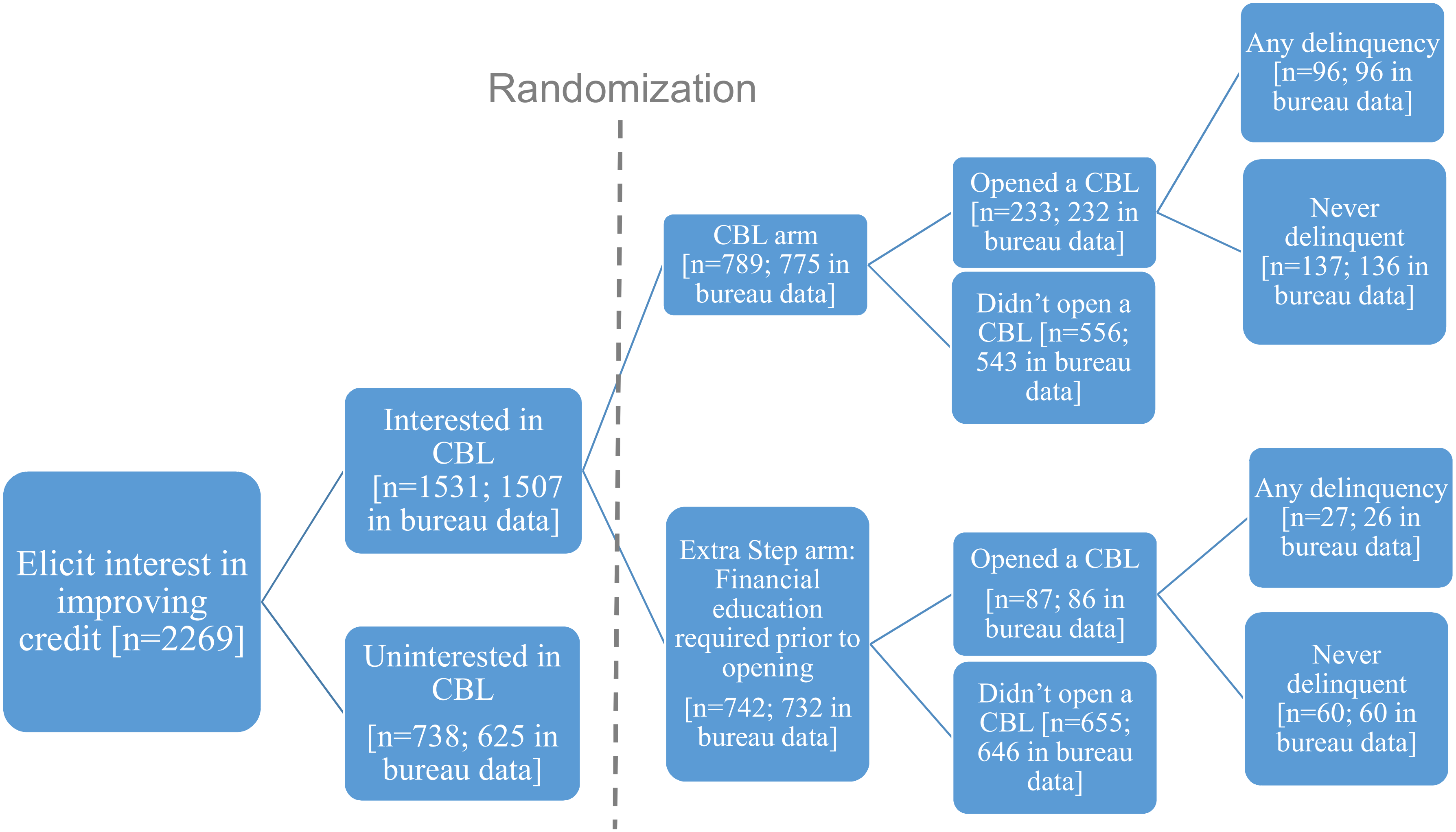
Wager, Stefan, and Susan Athey. 2018. “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests.” *Journal of the American Statistical Association* 113 (523): 1228–42.

Wolff, Sarah. 2016. “Providing a Fresh Start: An Analysis of Self-Help Federal Credit Union’s Fresh Start Product.” Self-Help Credit Union.

Zinman, Jonathan. 2015. “Household Debt: Facts, Puzzles, Theories, and Policies.” *Annual Review of Economics* 7 (1): 251–76.

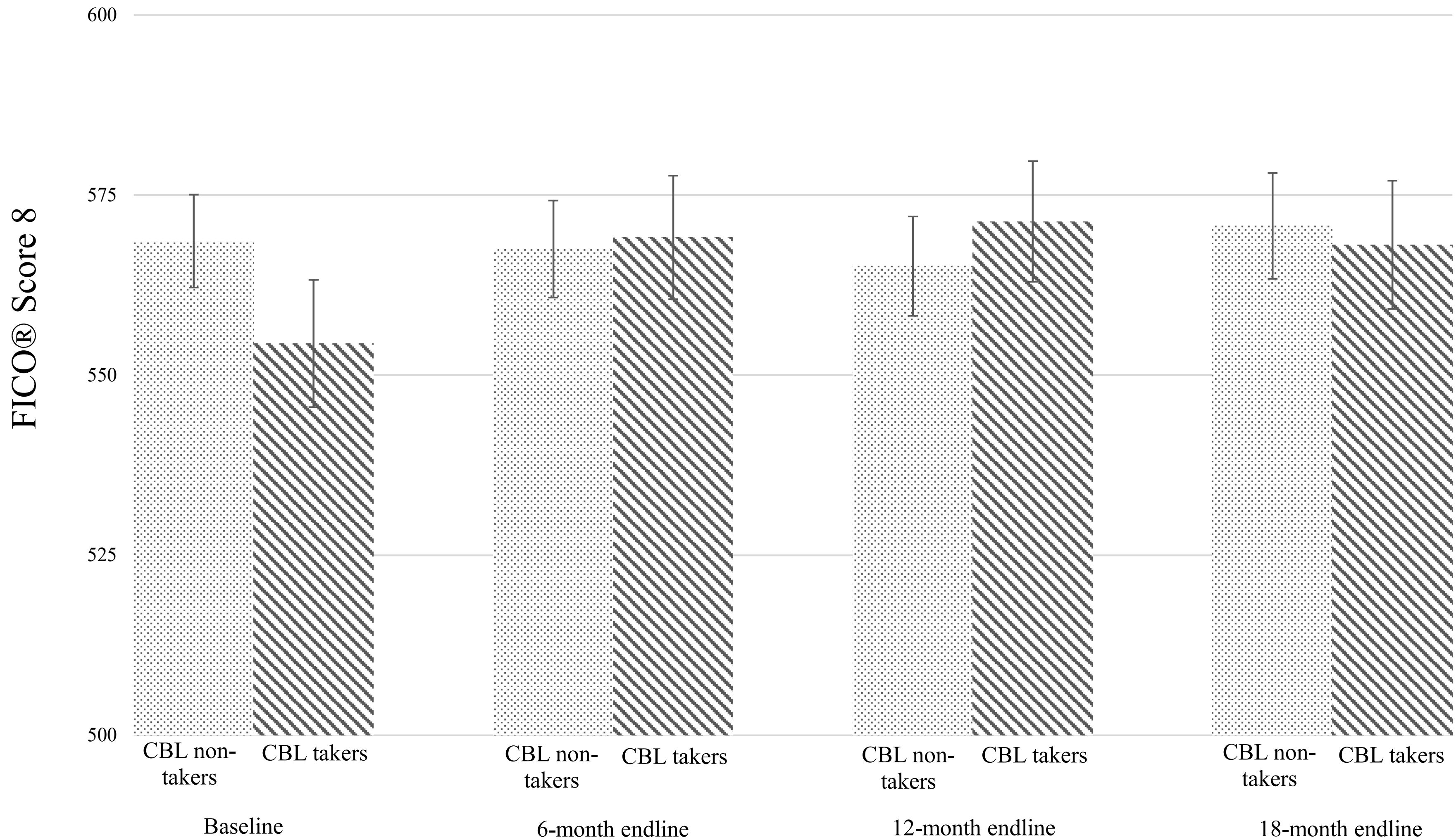
Figure 1. CBL Marketing Materials

Figure 2. Sample construction, experimental design, and CBL usage



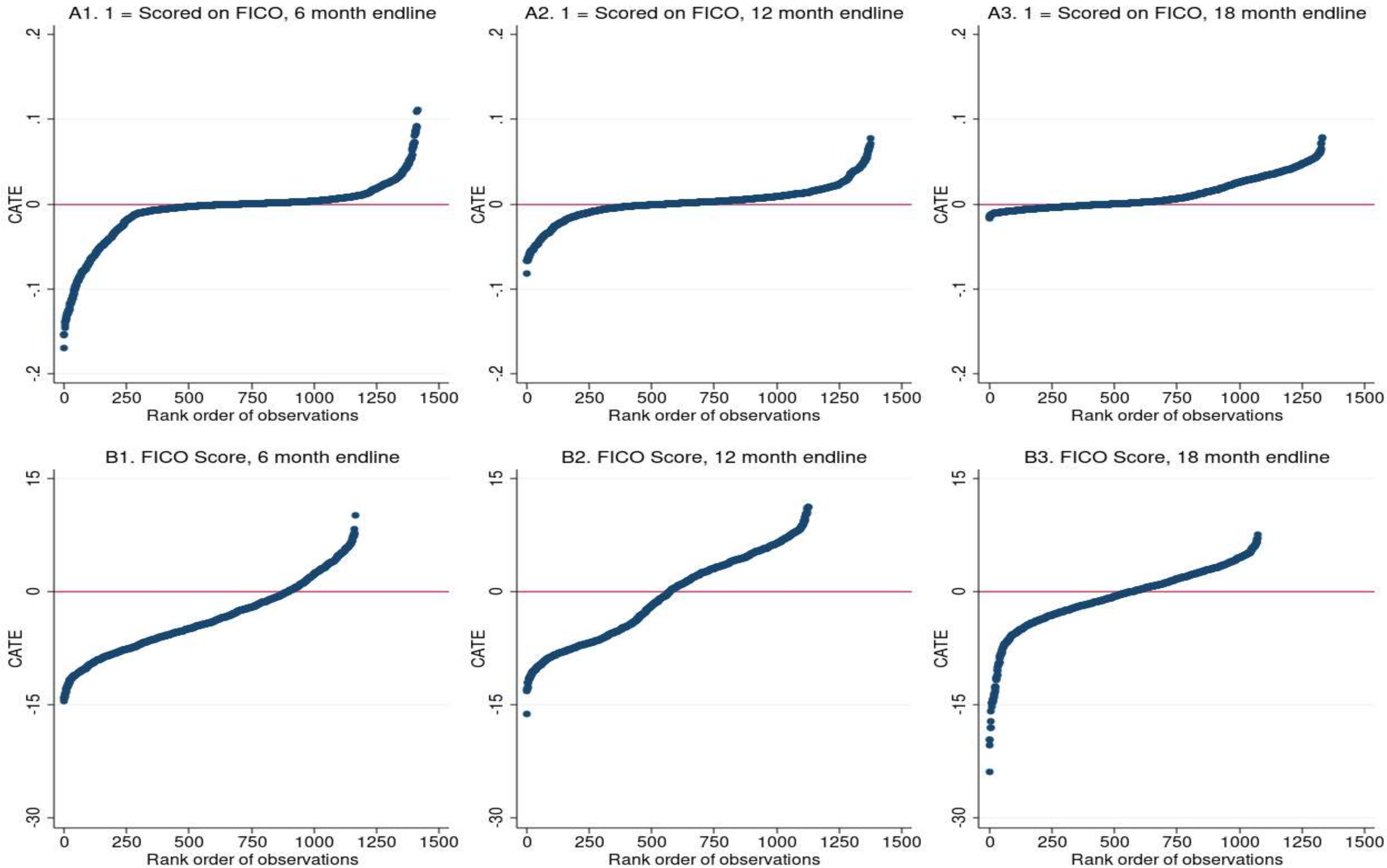
Note: "CBL"= Credit Builder Loan. Sample sizes include only those matched to the credit union's administrative data and hence inferred to be a credit union member at baseline. The sample sizes shown to be "in bureau data" are those in the study sample whom we were able to match to a credit report at baseline.

Figure 3. Selection into CBL:
Mean FICO® Score 8 over time (CBL Arm Only)



Note: Error bars show the 95% confidence intervals.

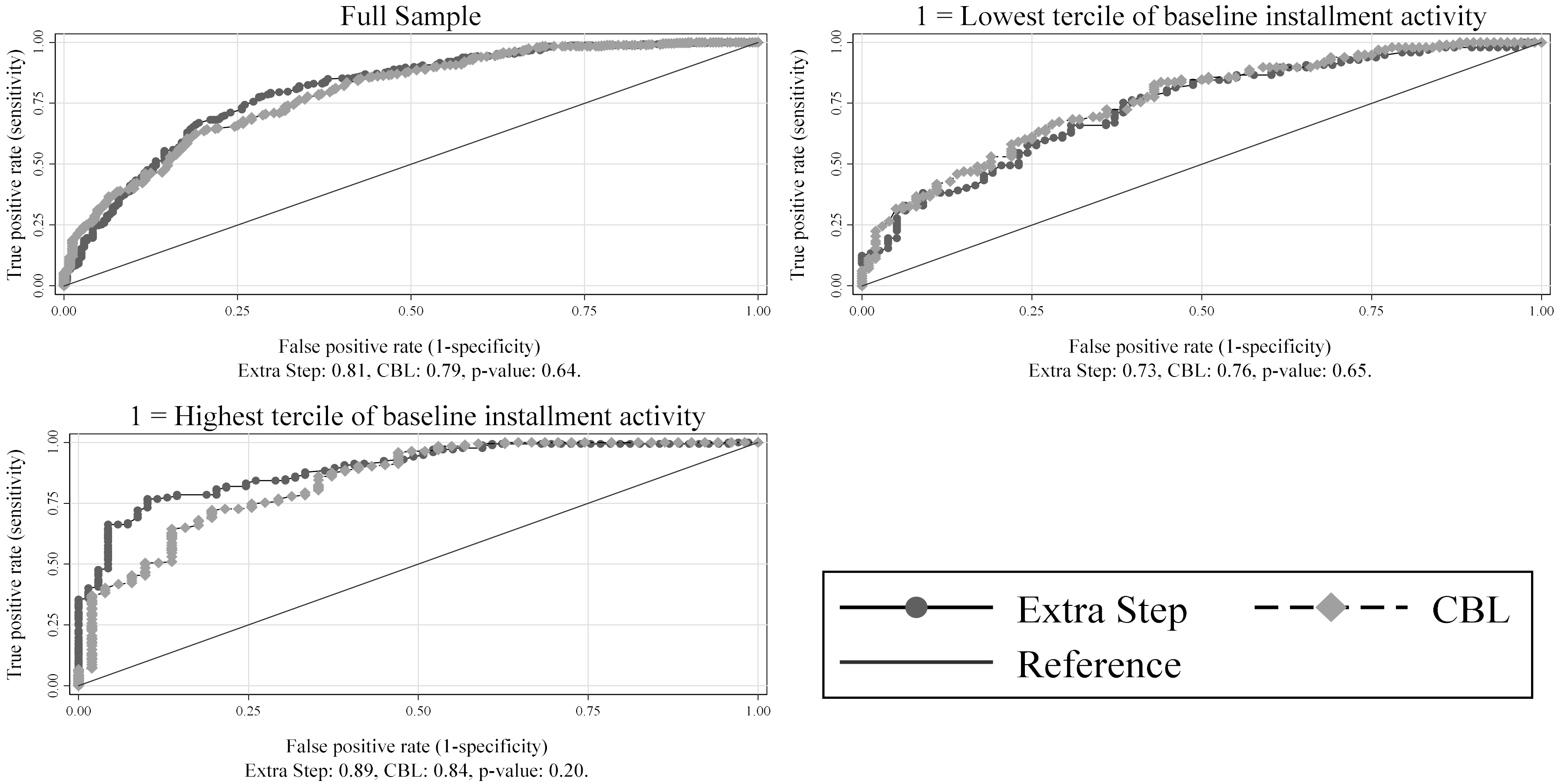
Figure 4. CATE plots for each outcome at each endline



Note: Conditional Average Treatment Effects (CATE), estimated using Generalized Random Forests package in R. Sample sizes are lower here (than e.g., the number of individuals with data for each outcome in Table 3) because we are doing each outcome-endline combination separately, and because of missing values on input variables.

Figure 5. Do CBLs change credit scores' predictive power?

Testing for differences in fit using Area Under the Curve (AUC) comparisons



Note: Each graph shows receiver operating characteristic (ROC) curves. We discretize at each 12-month endline credit score as a cutoff and predict more-risky behavior (i.e. defaulting) for those with scores below the cutoff and less-risky behavior (i.e. not defaulting) for those with scores above the cutoff. This aligns with what credit scores are constructed to do, which is to predict default ordinally. We then compare each person's prediction based on their 12-month score with their true value of the 18-month endline discretized delinquency index to calculate the true and false positive rates. ROCs require a discrete classification of the outcome to be predicted, so we discretize our 18-month endline delinquency index (see Data Appendix for details) at the median index value. As before, a higher value--the above median indicator--indicates more default. The true positive rate, on the y-axis, is (number of people correctly classified as more-risky at 12 months)/(number of observed more-risky people at 18 months). The false positive rate, on the x-axis, is (number of people incorrectly classified as more-risky at 12 months)/(number of observed less-risky people at 18 months). The areas under the curve (AUCs) for the Extra Step and CBL arms are shown below each graph along with the p-value of a chi-squared test of their equality (Delong, Delong, Clarke-Pearson 1988). The Reference (45-degree) line shows a ROC with no predictive power.

Table 1. Baseline characteristics and randomization balance for experiment sample

	(1)	(2)	(3)
	Mean (sd)		Univariate diff:
Sample:	CBL Arm N=789	Extra Step Arm N=742	(2) - (1) (se)
Age	43.823	42.475	-1.348
	(15.056)	(15.328)	(0.777)
Female	0.642	0.655	0.014
	(0.480)	(0.476)	(0.024)
Married	0.241	0.229	-0.012
	(0.428)	(0.421)	(0.022)
Number of adults in household	1.611	1.629	0.019
	(0.788)	(0.791)	(0.041)
Number of children in household	0.845	0.807	-0.038
	(1.237)	(1.229)	(0.064)
Race - Black	0.875	0.883	0.008
	(0.331)	(0.322)	(0.017)
College or more	0.264	0.253	-0.011
	(0.441)	(0.435)	(0.023)
Financial risk-taking scale (standardized)	0.000	0.039	0.039
	(1.000)	(1.008)	(0.052)
Self-control and credit knowledge index (standardized)	0.000	0.051	0.051
	(1.000)	(0.947)	(0.050)
Liquidity index (standardized)	0.000	-0.005	-0.005
	(1.000)	(0.928)	(0.049)
Delinquency index (standardized)	0.000	-0.074	-0.074
	(1.000)	(0.925)	(0.050)
1 = Higher than median of index of default outcomes	0.595	0.598	0.004
	(0.491)	(0.491)	(0.025)
1 = Scored on FICO	0.840	0.809	-0.031
	(0.367)	(0.394)	(0.020)
Baseline FICO Score	561.489	564.256	2.767
	(64.317)	(66.749)	(3.727)
Installment credit activity at baseline index (standardized)	0.000	-0.047	-0.047
	(1.000)	(1.000)	(0.052)
Revolving credit activity at baseline index (standardized)	0.000	0.006	0.006
	(1.000)	(1.026)	(0.052)
Number of prior loans, lifetime	7.773	7.220	-0.553
	(9.131)	(7.725)	(0.445)

Unit of observation is an individual. Index variables are standardized to the Extra Step Arm; see Data Appendix for details on index components and construction. Sample size varies across rows due to missing observations.

Table 2. Transition matrix for having a credit score

	(1)	(2)	(3)	(4)	(5)	(6)
	CBL Arm			Extra Step Arm		
	N=	Have score at 18-month endline	Do not have score at 18-month endline	N=	Have score at 18-month endline	Do not have score at 18-month endline
N=		668	91		632	85
Have score at baseline	622	96.95%	3.05%	609	95.07%	4.93%
Do not have score at baseline	137	47.45%	52.55%	108	49.07%	50.93%

Unit of observation is an individual. Sample size is slightly reduced from baseline because here it is limited to persons with a credit report at our 18-month endline.

Table 3. CBL average treatment effects on take-up and main outcomes

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Take-up: 1 = opened CBL	1 = has FICO® Score 8		FICO® Score 8	
Sample:	Full	Full		Have score at baseline	
CBL Arm	0.184				
	(0.020)				
CBL Arm * Post		0.018		-1.888	
		(0.015)		(2.730)	
CBL Arm * 6 month endline			0.008		-2.428
			(0.014)		(2.615)
CBL Arm * 12 month endline			0.020		-1.267
			(0.017)		(3.262)
CBL Arm * 18 month endline			0.028		-1.981
			(0.020)		(3.745)
Observations	1531	5978	5978	4865	4865
Individuals		1507	1507	1238	1238
Mean dependent variable in Extra Step Arm	0.119	0.873	0.873	567	567
SD dependent variable in Extra Step Arm	0.324	0.333	0.333	67	67
Mean dependent variable in Extra Step Arm at baseline	0.119	0.840	0.840	561	561
SD dependent variable in Extra Step Arm at baseline	0.324	0.367	0.367	64	64

We define CBL take-up as opening a CBL within 18 months of random assignment. Column (1) presents results from a single OLS regression of the dependent variable described in the column heading on the variable(s) shown in the applicable rows. Unit of observation for column (1) is a person at baseline. Columns (2) and (3) present results from a single OLS regression of the dependent variables described in the column heading on the variables shown in the applicable rows, Post, and person fixed effects. Unit of observation for columns (2) and (3) is a person-credit report, with four observations for most persons: baseline, and three endlines at 6, 12, and 18 months post-treatment assignment, all three of which are included in the Post indicator for the experiment period. Number of observations is lower than the number of individuals x 4 credit reports, because a small number of credit reports lack information on one or more dependent variables, including whether the person is scored. Standard errors, in parentheses, are clustered at the person-level.

Table 4. Causal forest aggregate test for CBL treatment effect heterogeneity

		(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable:	1 = has FICO® Score 8			FICO® Score 8		
	Endline:	6 mo	12 mo	18 mo	6 mo	12 mo	18 mo
Panel A: Aggregate test for treatment effect heterogeneity							
Differential forest prediction	Coefficient:	0.946	-0.974	-2.398	1.366	0.789	-0.274
	p-value:	0.077	0.831	0.993	0.003	0.053	0.623
Mean forest prediction	Coefficient:	0.994	0.955	1.011	1.004	1.023	1.087
	p-value:	0.293	0.451	0.205	0.057	0.382	0.405
Panel B: Average treatment effect by terciles of conditional average treatment effect							
Bottom tercile of CATE		-0.03	0.01	0.01	-8.30	-5.56	4.46
		(0.03)	(0.03)	(0.00)	(3.99)	(5.22)	(6.13)
Top tercile of CATE		0.00	0.01	0.01	6.76	5.31	-1.15
		(0.02)	(0.03)	(0.04)	(3.93)	(4.34)	(4.98)
Difference of top tercile - bottom tercile		0.03	0.00	0.00	15.06	10.87	-5.61
95% confidence interval range (+/-)		0.07	0.08	0.08	10.98	13.32	15.48
Number of observations		1413	1374	1330	1164	1126	1073

Unit of observation is a person-endline. For each column in this table-- each outcome-endline combination-- we ran a casual forest using the generalized random forest (GRF) package in R (Athey et al. 2019; R version 1.0.1, grf version 0.10.4) to predict the outcome listed in the column heading and obtain the CBL's conditional average treatment effects (CATE) on it. Panel A shows the test calibration of each forest: the differential forest coefficient and its p-value, and the mean forest coefficient and its p-value. Using the forest prediction on held-out data, these tests compute the best linear fit with two regressors, the target estimand and the mean forest prediction. The p-value of the "differential forest prediction" coefficient is the key omnibus test for the presence of heterogeneity, with rejection of the null of 0 implying rejection of the null of homogeneous treatment effects. A coefficient of 1 for "mean forest prediction" suggests that the mean forest prediction is accurate. Panel B uses the predicted CATE for each observation to divide observations into CATE terciles (see Figure 4 and its discussion in the text for why terciles are warranted) and reports statistics for the top and bottom terciles. The right hand side variables included in the causal forest for the binary outcome "1 = Has FICO® Score 8" are: age; number of adults in the household; number of children in the household; standardized risk taking score; number of open trade lines; savings balance and combined savings and checking balance (both in hundreds of dollars, winsorized at 95th percentile); dummies equal to one if baseline survey is missing, credit report is missing, the participant is female, the participant's race is Black, the participant is married, the participant has attended college, the participant's household income is less than 30k, the participant is still an SLCCU member, and the participant has a non-CBL loan; and standardized indices of insecurity, self-control, attention to credit status, credit process knowledge, delinquency, new credit, and lack of liquidity. The right hand side variables included in the causal forest for the continuous outcome of FICO® Score 8 are those listed above, with the addition of baseline FICO® score and a standardized index of the amount that the respondent owes based on account balances. Sample sizes are lower here (than e.g., the number of individuals with data for each outcome in Table 3) because we are doing each outcome-endline combination separately, and because of missing values on input variables.

Table 5. Potential sources of CBL treatment effect heterogeneity on FICO® Score

	(1)	(2)	(3)	(4)	(5)	(6)
	Mean (se) by treatment effect terciles			Average treatment effect on credit score by baseline variable terciles		
	Observations in lowest tercile (=0 for binary variables)	Observations in highest tercile (=1 for binary variables)	p-value (1) = (2)	Observations in lowest tercile (=0 for binary variables)	Observations in highest tercile (=1 for binary variables)	p-value (1) = (2)
Panel A. 6 -month endline						
Age	36.75	50.02	0.00	-12.13	12.68	0.01
	(0.65)	(0.65)		(6.29)	(7.55)	
Female	0.67	0.64	0.41	-6.45	-0.29	0.46
	(0.02)	(0.02)		(7.03)	(4.65)	
Married	0.26	0.22	0.28	-2.16	-0.65	0.86
	(0.02)	(0.02)		(4.51)	(7.59)	
Number of adults in household	1.57	1.55	0.80	-0.51	-17.97	0.22
	(0.04)	(0.04)		(5.15)	(13.14)	
Number of children in household	1.02	0.67	0.00	1.45	-10.48	0.21
	(0.06)	(0.06)		(5.38)	(7.74)	
Race - Black	0.88	0.88	0.91	-10.41	-1.04	0.52
	(0.02)	(0.02)		(13.92)	(3.95)	
College or more	0.36	0.19	0.00	-0.54	-5.85	0.55
	(0.02)	(0.02)		(4.38)	(7.71)	
Financial risk-taking scale (standardized)	0.12	-0.04	0.02	6.65	5.23	0.91
	(0.05)	(0.05)		(6.29)	(11.06)	
Self-control and credit knowledge index (standardized)	0.44	-0.14	0.00	5.63	-2.72	0.37
	(0.05)	(0.05)		(6.63)	(6.63)	
Liquidity index (standardized)	0.05	-0.08	0.07	-4.00	7.58	0.23
	(0.05)	(0.05)		(5.43)	(8.03)	
Baseline FICO® Score	577.65	543.80	0.00	2.98	-1.58	0.55
	(3.10)	(3.10)		(4.02)	(6.43)	
Installment credit activity at baseline index (standardized)	0.75	-0.52	0.00	22.79	-12.70	0.00
	(0.04)	(0.04)		(7.74)	(5.61)	
Revolving credit activity at baseline index (standardized)	0.33	-0.09	0.00	-4.12	2.90	0.45
	(0.05)	(0.05)		(4.72)	(7.92)	
Number of prior loans, lifetime	10.37	5.13	0.00	5.94	-3.47	0.35
	(0.40)	(0.40)		(7.57)	(6.66)	
Panel B. 12- month endline						
Age	47.66	43.46	0.00	-2.38	9.97	0.22
	(0.79)	(0.79)		(6.75)	(7.62)	
Female	0.62	0.69	0.03	-9.40	5.10	0.11
	(0.02)	(0.02)		(7.57)	(4.78)	
Married	0.27	0.24	0.41	3.29	-6.51	0.30
	(0.02)	(0.02)		(4.69)	(8.14)	
Number of adults in household	1.62	1.61	0.87	5.04	-24.26	0.04
	(0.04)	(0.04)		(5.51)	(13.31)	
Number of children in household	0.58	0.90	0.00	3.41	-14.02	0.07
	(0.06)	(0.06)		(5.67)	(7.88)	
Race - Black	0.83	0.88	0.07	-5.72	1.47	0.63
	(0.02)	(0.02)		(14.18)	(4.15)	
College or more	0.35	0.23	0.00	0.71	-0.10	0.93
	(0.02)	(0.02)		(4.65)	(7.88)	
Financial risk-taking scale (standardized)	-0.03	0.08	0.12	6.48	2.63	0.77
	(0.05)	(0.05)		(6.43)	(11.81)	
Self-control and credit knowledge index (standardized)	0.11	0.09	0.70	1.76	2.64	0.93
	(0.05)	(0.05)		(7.11)	(6.92)	
Liquidity index (standardized)	0.38	-0.14	0.00	0.56	5.16	0.65
	(0.05)	(0.05)		(5.58)	(8.47)	
Baseline FICO® Score	632.23	511.69	0.00	8.40	-7.89	0.05
	(2.08)	(2.09)		(4.29)	(7.07)	
Installment credit activity at baseline index (standardized)	0.29	-0.06	0.00	17.42	-10.04	0.01
	(0.05)	(0.05)		(8.05)	(6.01)	
Revolving credit activity at baseline index (standardized)	0.57	-0.13	0.00	-1.76	5.29	0.47
	(0.05)	(0.05)		(5.19)	(8.16)	
Number of prior loans, lifetime	8.71	7.14	0.01	0.14	3.05	0.79
	(0.40)	(0.40)		(8.10)	(7.12)	

Notes: For columns (1) and (2), each row with a continuous (binary) variable shows the results of an OLS regression (t-test) of the variable listed in the row on a dummy for those observations classified in the highest treatment effect tercile (equal to 1) and a dummy for those observations classified in the lowest treatment effect tercile (equal to 0). Column (3) shows the results of the t-test that the two coefficients (means) are equal. For each of columns (4) and (5), each row with a continuous (binary) variable shows the results of an OLS regression of FICO Score on treatment for those observations classified in the tercile (binary value) listed in the column header of the variable listed in the row. Column (6) shows the results of the t-test that the two coefficients are equal. Terciles of treatment effect are determined using the conditional average treatment effects (CATEs) from the causal forest estimated in Table 4 Column 4 (Panel A here) or Table 4 Column 5 (Panel B here). See Data Appendix for details on index components and construction.

Table 6. CBL average and heterogeneous treatment effects on credit behaviors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FICO® Score 8 Factor:	New credit	Delinquency	Amounts owed		Credit Mix		
Dependent variable (index) includes:	Inquiries, number of accounts	10 measures of delinquency, collections, & derogatories (higher values = less timely repmt). Includes CBL delinquency.	Balances: Revolving, auto loans, other Installment	Utilization: 4 discrete measures of credit limit usage and outstanding balances; # open installment loans	1=(open installment and open revolving loan)	Currently delinquent on CBL	Ever delinquent on CBL
Sample:	Full	Full	Full	Full	Full	Full	Full
Data source:	Credit Bureau	Credit Bureau	Credit Bureau	Credit Bureau	Credit Bureau	SLCCU Admin	SLCCU Admin
Panel A. Average effects							
CBL Arm						0.014	0.085
						(0.004)	(0.014)
CBL Arm * Post	0.004	0.081	-0.058	-0.002	-0.032		
	(0.036)	(0.039)	(0.041)	(0.041)	(0.043)		
Observations	5981	5981	5488	5981	5981	4577	1531
Individuals	1507	1507	1425	1507	1507	1531	1531
Panel B. Heterogeneity by baseline credit access							
CBL Arm * Bottom tercile of installment credit activity at baseline index (i)						0.017	0.099
						(0.006)	(0.021)
CBL Arm * Middle tercile of installment credit activity at baseline index (ii)						0.006	0.075
						(0.005)	(0.021)
CBL Arm * Top tercile of installment credit activity at baseline index (iii)						0.019	0.090
						(0.008)	(0.022)
CBL Arm * Post * Bottom tercile of installment credit activity at baseline index (iv)	0.013	0.026	0.021	0.096	0.117		
	(0.040)	(0.060)	(0.065)	(0.069)	(0.056)		
CBL Arm * Post * Middle tercile of installment credit activity at baseline index (v)	-0.046	0.001	-0.109	-0.041	-0.159		
	(0.047)	(0.062)	(0.079)	(0.070)	(0.081)		
CBL Arm * Post * Top tercile of installment credit activity at baseline index (vi)	0.023	0.224	-0.068	-0.082	-0.085		
	(0.084)	(0.081)	(0.063)	(0.074)	(0.082)		
p-value of (i) = (ii) or (iv) = (v)	0.335	0.775	0.205	0.163	0.005	0.123	0.419
p-value of (ii) = (iii) or (v) = (vi)	0.474	0.028	0.689	0.687	0.522	0.119	0.611
p-value of (i) = (iii) or (iv) = (vi)	0.919	0.049	0.321	0.078	0.042	0.816	0.778
Observations	5970	5970	5482	5970	5970	4490	1502
Individuals	1502	1502	1423	1502	1502	1502	1502
Mean dependent variable in Extra Step Arm, baseline	0.000	0.000	0.000	0.000	0.000	NA	NA
Mean dependent variable in Extra Step Arm, Post	-0.095	-0.036	0.099	0.100	0.152	0.005	0.036

Standard errors, in parentheses, are clustered at the person-level. In Columns 1-5 unit of observation is a person-credit report, with four observations for most persons: baseline, and three endlines at 6, 12, and 18 months post-treatment assignment, all three of which are included in the Post indicator for the experiment period. Number of observations is lower than the number of individuals x 4 credit reports because some credit reports lack information on one or more dependent variables. Unit of observation in Column 6 is a person-SLCCU data snapshot, with those snapshots timed to coincide roughly with the credit report endlines. Columns 6 and 7 use endline data only, because no one in our sample had a CBL at baseline. Those who did not open a CBL are a coded as zero in columns 6 and 7. Each panel-column presents results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the applicable rows, with the regressions in Panel A columns 1 - 5 also including person fixed effects and Post, and the regressions in Panel B columns 1 - 5 also including person fixed effects and Post interacted with each of the baseline installment credit activity terciles. Index variables are standardized to be mean zero and standard deviation one in the Extra Step Arm at baseline; see Data Appendix for details on index components and construction.

Table 7. CBL treatment effects on usage of other SLCCU products

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	1 = Remain an SLCCU member		1 = Any non-CBL loan with SLCCU outstanding		Balances of all savings accounts ($ hundreds)		Balances of all savings + checking accounts ($ hundreds)	
Sample:	Full							
Panel A: Main effects								
CBL Arm * Post	-0.008		0.009		2.476		1.297	
	(0.011)		(0.019)		(1.214)		(1.669)	
Panel B: Heterogeneity by baseline credit access								
CBL Arm * Post * Bottom tercile of installment credit activity at baseline index (i)		-0.010		0.049		0.149		0.077
		(0.021)		(0.027)		(1.355)		(1.590)
CBL Arm * Post * Middle tercile of installment credit activity at baseline index (ii)		-0.012		-0.011		5.519		5.456
		(0.021)		(0.040)		(3.064)		(3.164)
CBL Arm * Post * Top tercile of installment credit activity at baseline index (iii)		-0.001		-0.011		2.122		-1.479
		(0.015)		(0.031)		(1.727)		(3.890)
p-value of (i) = (ii)		0.730		0.148		0.369		0.711
p-value of (ii) = (iii)		0.685		0.997		0.334		0.167
p-value of (i) = (iii)		0.960		0.217		0.109		0.129
Observations	6124	6008	6124	6008	6124	6008	6124	6008
Individuals	1531	1502	1531	1502	1531	1502	1531	1502
Mean dependent variable in Extra Step Arm at baseline	1.000	1.000	0.322	0.327	4.987	5.040	7.435	7.536

Unit of observation is a person-SLCCU data snapshot, with four observations for most persons at roughly the same timing as our credit report pulls: baseline, and three endlines at 6, 12, and 18 months post-treatment assignment, all three of which are included in the Post indicator for the experiment period. Standard errors, in parentheses, are clustered at the person-level. Each column presents results from a single OLS regression of the dependent variable described in the column heading on the variable(s) shown in the rows, Post and person fixed effects (odd columns), with even columns including Post * Bottom tercile of installment credit activity at baseline index, Post * Middle tercile of installment credit activity at baseline index, and Post * Top tercile of installment credit activity at baseline index instead of the Post indicator. All outcome variables here are calculated from SLCCU administrative data. Balances are recorded as zero for those who leave the credit union.

Table 8. Selection into CBL

	(1)	(2)	(3)	(4)
Dependent variable:	1 = Has FICO® Score 8		FICO® Score 8	
Sample:	Full CBL Arm		CBL Arm participants who have score at baseline	
Took up CBL * Post	0.095	0.105	16.617	12.949
	(0.025)	(0.016)	(4.639)	(4.140)
Controls for baseline variables * Post	No	Yes	No	Yes
Number of people in sample that took up a CBL	231	231	191	191
Observations	3065	3065	2466	2466
Individuals	772	772	625	625
Mean dependent variable in CBL Arm at baseline	0.810	0.810	564	564

Unit of observation is a person-credit report, with four observations for most persons: baseline, and three endlines at 6, 12, and 18 months post-treatment assignment, all three of which are included in Post indicator for the experiment period. Standard errors, in parentheses, are clustered at the person-level. Each column presents results from a single OLS regression of the dependent variable described in the column heading on the row variables described in the table, Post, and person fixed effects. Even-numbered columns include interactions with the Post indicator selected by Post Double Selection LASSO: baseline FICO® Score 8, 1 = baseline FICO® Score 8 in the 400s,1 = baseline FICO® Score 8 in the 500s, and 1 = baseline FICO® Score 8 in the 600s. Heterogeneous treatment effects by baseline installment activity (Table 5) imply that we cannot identify a pure selection effect separately for those sub-groups, and so we only estimate average selection effects here.

Table 9. Do CBLs change the predictive power of credit scores? Testing for differences in the default-score gradient

	(1)	(2)
Dependent variable:	Delinquency index: includes 10 measures of delinquency, collections, & derogatories (higher values = less timely repmt). Includes CBL delinquency.	
	(18 month endline)	
FICO® Score 8 (hundreds) at baseline	0.043	-0.047
	(0.053)	(0.051)
FICO® Score 8 (hundreds) at 12 month endline * CBL Arm (i)	-0.822	
	(0.051)	
FICO® Score 8 (hundreds) at 12 month endline * Extra Step Arm (ii)	-0.831	
	(0.051)	
FICO® Score 8 (hundreds) at 12 month endline * CBL Arm * Bottom tercile of installment credit activity at baseline index (iii)		-0.828
		(0.050)
FICO® Score 8 (hundreds) at 12 month endline * Extra Step Arm * Bottom tercile of installment credit activity at baseline index (iv)		-0.834
		(0.051)
FICO® Score 8 (hundreds) at 12 month endline * CBL Arm * Middle tercile of installment credit activity at baseline index (v)		-0.770
		(0.050)
FICO® Score 8 (hundreds) at 12 month endline * Extra Step Arm * Middle tercile of installment credit activity at baseline index (vi)		-0.772
		(0.050)
FICO® Score 8 (hundreds) at 12 month endline * CBL Arm * Top tercile of installment credit activity at baseline index (vii)		-0.710
		(0.052)
FICO® Score 8 (hundreds) at 12 month endline * Extra Step Arm * Top tercile of installment credit activity at baseline index (viii)		-0.733
		(0.050)
p-value of (i) = (ii)	0.255	
p-value of (iii) = (iv)		0.687
p-value of (v) = (vi)		0.839
p-value of (vii) = (viii)		0.080
Observations	1217	1217
Mean dependent variable in Extra Step Arm	0.066	0.066

Unit of observation is a person. Standard errors, in parentheses, are Huber-White. Each column presents results from a single OLS regression of the dependent variable described in the column heading on the variables shown in the rows and FICO® Score 8 at baseline. Index variables are standardized to be mean zero and standard deviation one in the Extra Step Arm at baseline; see Data Appendix for details on index components and construction. Sample here is limited to persons for whom we could obtain a credit report at our 18-month endline and who have a credit score at baseline and the 12-month endline.

Data Appendix

Index construction rules

1. Standardize each component with respect to the Extra Step Arm.
2. Calculate the person-level mean across non-missing components (if someone is missing all components their index value is missing).
3. Standardize each index with respect to the Extra Step Arm.

Variable definition details not fully specified in the tables or main text

Baseline financial risk-taking scale (Measured from baseline survey, higher values indicate greater risk tolerance)

In Tables 1 & 5; Appendix Tables 1, 2a, & 2b.

1. Q: "I am willing to take a risk financially if there is a chance of substantial gain."
   A: 1 = strongly disagree, 2= disagree, 3 = feel neutrally, 4= agree, 5 = strongly agree

Baseline self-control and credit knowledge index (12 components, each measured from baseline survey, higher values indicate more self-control)

In Tables 1 & 5; Appendix Tables 1, 2a, & 2b.

1. Q: "Before I buy something I carefully consider whether I can afford it."
   A: 1 = strongly disagree, 2= disagree, 3 = feel neutrally, 4= agree, 5 = strongly agree
2. Q: "I tend to live for today and let tomorrow take care of itself."
   A: 1 = strongly agree, 2= agree, 3 = feel neutrally, 4= disagree, 5 = strongly disagree
3. Q: "I set long term financial goals of five years or more and strive to achieve them."
   A: 1 = strongly disagree, 2= disagree, 3 = feel neutrally, 4= agree, 5 = strongly agree
4. Q: "I often find that I regret spending money. I wish that when I had cash, I was better disciplined and saved my money rather than spent it."
   A: 1 = strongly agree, 2= agree, 3 = feel neutrally, 4= disagree, 5 = strongly disagree
5. Q: "I have trouble finishing or completing my tasks."
   A: 1 = strongly agree, 2= agree, 3 = feel neutrally, 4= disagree, 5 = strongly disagree
6. Q: "In the past 12 months, have you checked your credit score?"
   A: 0 = No, 1 = Yes
7. Q: "In the past 12 months, have you obtained a copy of your credit report?"
   A: 0 = No, 1 = Yes

8. Correctly answered "Could your credit rating affect the amount of interest you would pay on a bank loan?" **(Yes)**
   A: 0 = No, 1 = Yes
9. Correctly answered: "Could your health affect the amount of interest you would pay on a bank loan?" **(No)**
   A: 0 = No, 1 = Yes
10. Correctly answered: "Could your age affect the amount of interest you would pay on a bank loan?" **(No)**
    A: 0 = No, 1 = Yes
11. Correctly answered: "Could how much you borrow overall affect the amount of interest you would pay on a bank loan?" **(Yes)**
    A: 0 = No, 1 = Yes
12. Correctly answered: "Could how long you take to repay the loan affect the amount of interest you would pay on a bank loan?" **(Yes)**
    A: 0 = No, 1 = Yes

Baseline liquidity index (7 components, measured from baseline survey and baseline SLCCU data, higher values indicate more liquidity)
In Tables 1 & 5; Appendix Tables 1, 2a, & 2b.

1. Q: "My financial situation is a source of stress in my life."
   A: 1 = strongly agree, 2= agree, 3 = feel neutrally, 4= disagree, 5 = strongly disagree
2. Q: "In a typical month, it is difficult for me to cover my expenses and pay all my bills."
   A: 1 = strongly agree, 2= agree, 3 = feel neutrally, 4= disagree, 5 = strongly disagree
3. Q: "I am confident that I could come up with $2000 if an unexpected need arose within the next month"
   A: 1 = strongly disagree, 2= disagree, 3 = feel neutrally, 4= agree, 5 = strongly agree
4. Q: "How would you describe your overall financial situation? Would you say…"
   A: 1 = bad, 2 = not very good, 3 = okay, 4 = very good, 5 = excellent
5. HH income is greater than $30k (0 = income less than or equal to $30K, 1 = greater than $30K)
6. Savings Balance ($ hundreds, top-coded at 95%)
7. More than $60 in savings (0 = less than or equal to $60 in savings, 1 = more than $60 in savings)

Delinquency index (10 components, each measured from credit bureau data; higher values indicate more default, delinquency, collection activity on accounts)

In Figure 5 as outcome; Tables 1 (baseline), 6 (outcome), & 9 (outcome); Appendix Tables 1 (baseline), 2a and 2b (baseline), & 3 (outcome).

1. Account 30 days past due in the last 12 months (0 = does not have account past due, 1 = has account past due)
2. Account 90 days past due in the last 12 months (0 = does not have account past due, 1 = has account past due)
3. Account in collection (0 = does not have account in collection, 1 = has account in collection)
4. Has amount past due (0 = does not have amount past due, 1 = has amount past due)
5. Account with a major derogatory event (0 = does not have major derogatory event, 1 = has major derogatory event)
6. Number of accounts 30 days past due in the last 12 months
7. Number of accounts 90 days past due in the last 12 months
8. Number of accounts in collection
9. Amount past due ($)
10. Number of accounts with a major derogatory event

Baseline installment credit activity index (3 components, each measured from credit bureau, higher values indicate more installment credit)

In Figure 5; Tables 1, 5, 6, 7 & 9; Appendix Tables 1, 2a, 2b, 3 & 4.

1. Number of open installment loans (transformed by taking inverse hyperbolic sine)
2. Any open installment loan (0 = no open installment loan, 1 = any open installment loan)
3. Number of inquiries made within last 12 months (transformed by taking inverse hyperbolic sine

Baseline revolving credit activity index (3 components, each measured from credit bureau, higher values indicate more revolving credit access)

In Tables 1 & 5; Appendix Tables 1, 2a, & 2b.

1. Number of open revolving loans (transformed by taking inverse hyperbolic sine)
2. Any open revolving loan (0 = no open revolving loan, 1 = any open revolving loan)
3. Utilization of revolving loans (transformed by taking inverse hyperbolic sine)

Baseline number of prior loans, lifetime (Measured from credit bureau, higher values indicate more loans)

In Tables 1 & 5; Appendix Tables 1, 2a, & 2b.

1. Total number of open and closed loans.

New credit (2 components; each measured from credit bureau; higher values indicate more new credit)

In Table 6; Appendix Table 3.

1. Number of inquiries made in the last 12 months (bureau data)
2. The number of accounts (bureau data)

Amounts owed: Balances (3 components, each measured from credit bureau; higher values indicate larger amounts owed)

In Table 6; Appendix Table 3.

1. Outstanding revolving loan balance
2. Outstanding installment loan balance
3. Outstanding auto loan balance

Amounts owed: Utilization (5 components, each measured from credit bureau data; higher values indicate more utilization)

In Table 6; Appendix Table 3.

1. Revolving utilization is over 30% (0 = below 30%, 1 = above 30%; missing if no credit line)
2. Number of open installment loans
3. Outstanding revolving loan balance (0 = no outstanding balance, 1 = outstanding balance)
4. Outstanding auto loan balance (0 = no outstanding balance, 1 = outstanding balance)
5. Outstanding installment loan balance (0 = no outstanding balance, 1 = outstanding balance)

Credit mix (Measured from credit bureau data; higher value indicates more credit types open)

In Table 6; Appendix Table 3.

1. Has an open installment loan and an open revolving loan (0 = no loan, 1 = has loan)

Appendix Table 1. Do baseline observable characteristics help predict takeup of the CBL?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	CBL Arm (N=789)			Extra-Step Arm (N=742)			
	Takers Mean (se)	Takers - Non-takers Diff (se)	p-value diff = 0	Takers Mean (se)	Takers - Non-takers Diff (se)	p-value diff = 0	p-value (1) = (4)
Age	43.00	0.74	0.53	42.39	-1.62	0.35	0.74
	(1.00)	(1.19)		(1.63)	(1.73)		
Female	0.63	-0.04	0.27	0.60	-0.05	0.37	0.64
	(0.03)	(0.04)		(0.05)	(0.05)		
Married	0.23	0.01	0.81	0.28	0.05	0.34	0.39
	(0.03)	(0.03)		(0.05)	(0.05)		
Number of adults in household	1.60	-0.04	0.57	1.71	0.11	0.22	0.30
	(0.05)	(0.06)		(0.09)	(0.09)		
Number of children in household	0.92	0.16	0.11	0.88	0.04	0.75	0.85
	(0.08)	(0.10)		(0.13)	(0.14)		
Race - Black	0.90	0.03	0.23	0.86	-0.02	0.67	0.26
	(0.02)	(0.03)		(0.04)	(0.04)		
College or more	0.27	0.02	0.49	0.38	0.14	0.01	0.05
	(0.03)	(0.03)		(0.05)	(0.05)		
Financial risk-taking scale (standardized)	0.14	0.14	0.09	0.02	0.02	0.84	0.38
	(0.07)	(0.08)		(0.11)	(0.12)		
Self-control and credit knowledge index (standardized)	0.01	-0.06	0.42	0.07	0.08	0.49	0.63
	(0.06)	(0.08)		(0.10)	(0.11)		
Liquidity index (standardized)	0.00	0.01	0.86	-0.14	-0.16	0.16	0.22
	(0.06)	(0.08)		(0.10)	(0.11)		
Baseline FICO Score	554.38	-14.23	0.01	561.02	-0.53	0.95	0.43
	(4.74)	(5.68)		(8.06)	(8.53)		
Installment credit activity at baseline index (standardized)	-0.03	0.02	0.83	0.14	0.16	0.16	0.17
	(0.07)	(0.08)		(0.11)	(0.11)		
Revolving credit activity at baseline index (standardized)	-0.06	-0.09	0.25	-0.05	-0.06	0.62	0.95
	(0.07)	(0.08)		(0.11)	(0.12)		
Number of prior loans, lifetime	7.22	0.00	1.00	8.79	1.15	0.28	0.17
	(0.57)	(0.68)		(0.93)	(0.99)		

Unit of observation is an individual. As in Table 3, we define CBL take-up as opening a CBL within 18 months of random assignment. All row variables measured at baseline, with most having sample sizes slightly lower than the full-sample N reported in the column headings, due to survey non-response or credit report missing information. Please see Data Appendix for details on index components and construction.

Appendix Table 2a. Baseline characteristics
(Same as Table 1 but sample here is restricted to those in the top tercile of installment credit activity index at baseline)

	(1)	(2)	(3)
	Mean (sd)		Univariate diff:
Sample:	CBL Arm N= 246	Extra Step Arm N= 244	(2) - (1) (se)
Age	41.881	41.516	-0.365
	(14.670)	(14.578)	(1.321)
Female	0.664	0.736	0.072
	(0.473)	(0.442)	(0.041)
Married	0.331	0.260	-0.070
	(0.471)	(0.440)	(0.042)
Number of adults in household	1.616	1.612	-0.004
	(0.719)	(0.793)	(0.069)
Number of children in household	1.004	0.851	-0.153
	(1.230)	(1.199)	(0.111)
Race - Black	0.895	0.942	0.048
	(0.308)	(0.234)	(0.025)
College or more	0.388	0.365	-0.023
	(0.488)	(0.482)	(0.044)
Financial risk-taking scale (standardized)	0.017	0.104	0.087
	(0.991)	(1.043)	(0.093)
Self-control and credit knowledge index (standardized)	0.239	0.271	0.032
	(0.999)	(0.959)	(0.089)
Liquidity index (standardized)	0.079	0.045	-0.033
	(1.010)	(0.918)	(0.087)
Delinquency index (standardized)	0.331	0.283	-0.048
	(1.050)	(0.954)	(0.091)
1 = Higher than median of index of default outcomes	0.734	0.760	0.027
	(0.443)	(0.428)	(0.039)
1 = Scored on FICO	0.988	0.996	0.008
	(0.110)	(0.064)	(0.008)
Baseline FICO Score	565.714	561.910	-3.803
	(56.460)	(53.620)	(4.994)
Installment credit activity at baseline index (standardized)	1.085	1.067	-0.017
	(0.353)	(0.378)	(0.033)
Revolving credit activity at baseline index (standardized)	0.354	0.301	-0.053
	(1.031)	(1.064)	(0.095)
Number of prior loans, lifetime	10.639	9.756	-0.883
	(10.608)	(8.443)	(0.866)

Unit of observation is an individual. Index variables are standardized to be mean zero and standard deviation one in the Extra Step Arm; see Data Appendix for details on index components and construction. Sample size varies across rows due to missing observations.

Appendix Table 2b. Baseline characteristics
(Same as Table 1 but sample here is restricted to those in the bottom tercile of installment credit activity index at baseline)

	(1)	(2)	(3)
	Mean (sd)		Univariate diff: (2) - (1) (se)
Sample:	CBL Arm N= 283	Extra Step Arm N= 243	
Age	42.745	41.943	-0.801
	(14.301)	(15.842)	(1.325)
Female	0.568	0.594	0.026
	(0.496)	(0.492)	(0.043)
Married	0.148	0.184	0.036
	(0.356)	(0.388)	(0.033)
Number of adults in household	1.662	1.631	-0.031
	(0.900)	(0.843)	(0.077)
Number of children in household	0.692	0.811	0.119
	(1.090)	(1.229)	(0.103)
Race - Black	0.865	0.855	-0.010
	(0.342)	(0.353)	(0.031)
College or more	0.152	0.135	-0.017
	(0.360)	(0.343)	(0.031)
Financial risk-taking scale (standardized)	0.025	-0.060	-0.084
	(1.023)	(0.998)	(0.090)
Self-control and credit knowledge index (standardized)	-0.101	-0.134	-0.033
	(0.978)	(0.943)	(0.085)
Liquidity index (standardized)	-0.099	-0.148	-0.049
	(0.909)	(0.873)	(0.078)
Delinquency index (standardized)	-0.282	-0.407	-0.125
	(0.852)	(0.778)	(0.071)
1 = Higher than median of index of default outcomes	0.465	0.424	-0.041
	(0.500)	(0.495)	(0.043)
1 = Scored on FICO	0.605	0.534	-0.071
	(0.490)	(0.500)	(0.043)
Baseline FICO Score	529.027	542.854	13.827
	(61.548)	(73.668)	(7.874)
Installment credit activity at baseline Index (standardized)	-1.205	-1.178	0.027
	(0.343)	(0.333)	(0.029)
Revolving credit activity at baseline index (standardized)	-0.429	-0.409	0.021
	(0.799)	(0.826)	(0.071)
Number of prior loans, lifetime	4.077	3.980	-0.098
	(5.589)	(4.969)	(0.490)

Unit of observation is an individual. Index variables are standardized to be mean zero and standard deviation one in the Extra Step Arm; see Data Appendix for details on index components and construction. Sample size varies across columns due to missing observations.

Appendix Table 3. CBL average and heterogeneous treatment effects on credit behaviors
(Same as Table 6 but here sample is restricted to those who have a score at baseline)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FICO® Score 8 Factor:	New Credit	Delinquency	Amounts Owed		Credit Mix		
Dependent variable (index) includes:	Inquiries, Number of Accounts	10 measures of delinquency, collections, & derogatories (higher values = less timely repayment)	Balances: Revolving, auto loans, other installment	Utilization: 4 discrete measures of credit limit usage and outstanding balances; # open installment loans	1 = (open installment and open revolving loan)	Currently delinquent on CBL	Ever delinquent on CBL
Sample:				Have score at baseline			
Panel A. Main effects							
CBL Arm						0.011	0.091
						(0.004)	(0.015)
CBL Arm * Post	0.000	0.096	-0.064	-0.021	-0.042		
	(0.041)	(0.043)	(0.043)	(0.047)	(0.050)		
Observations	4945	4945	4929	4945	4945	3701	1238
Individuals	1238	1238	1238	1238	1238	1238	1238
Panel B. Heterogeneity by baseline credit access							
CBL Arm * Bottom tercile of installment credit activity at baseline index (i)						0.013	0.120
						(0.007)	(0.030)
CBL Arm * Middle tercile of installment credit activity at baseline index (ii)						0.005	0.072
						(0.005)	(0.022)
CBL Arm * Top tercile of installment credit activity at baseline index (iii)						0.015	0.090
						(0.007)	(0.022)
CBL Arm * Post * Bottom tercile of installment credit activity at baseline index (iv)	-0.026	0.009	0.006	0.115	0.185		
	(0.048)	(0.070)	(0.073)	(0.104)	(0.086)		
CBL Arm * Post * Middle tercile of installment credit activity at baseline index (v)	-0.016	0.023	-0.106	-0.041	-0.156		
	(0.043)	(0.063)	(0.082)	(0.074)	(0.084)		
CBL Arm * Post * Top tercile of installment credit activity at baseline index (vi)	0.030	0.218	-0.069	-0.085	-0.078		
	(0.085)	(0.081)	(0.063)	(0.074)	(0.082)		
p-value of (i) = (ii) or (iv) = (v)	0.875	0.879	0.308	0.221	0.005	0.333	0.182
p-value of (ii) = (iii) or (v) = (vi)	0.628	0.058	0.717	0.671	0.508	0.223	0.545
p-value of (i) = (iii) or (iv) = (vi)	0.564	0.051	0.440	0.117	0.027	0.834	0.407
Observations	4945	4945	4929	4945	4945	3701	1238
Individuals	1238	1238	1238	1238	1238	1238	1238
Mean dependent variable in Extra Step Arm at baseline	0.139	0.156	0.057	0.151	0.139	NA	NA

Standard errors, in parentheses, are clustered at the person-level. In Columns 1-5 unit of observation is a person-credit report, with four observations for most persons: baseline, and three endlines at 6, 12, and 18 months post-treatment assignment, all three of which are included in the Post indicator for the experiment period. Number of observations is lower than the number of individuals x 4 credit reports because some credit reports lack information on one or more dependent variables. Unit of observation in Column 6 is a person-SLCCU data snapshot, with those snapshots timed to coincide roughly with the credit report endlines. Columns 6 and 7 use endline data only, because no one in our sample had a CBL at baseline. Those who did not open a CBL are a coded as zero in columns 6 and 7. Each panel-column presents results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the applicable rows, with the regressions in Panel A columns 1 - 5 also including person fixed effects and Post, and the regressions in Panel B columns 1 - 5 also including person fixed effects and Post interacted with each of the baseline installment credit activity terciles. Index variables are standardized to be mean zero and standard deviation one in the Extra Step Arm at baseline; see Data Appendix for details on index components and construction.

Appendix Table 4. CBL treatment effects on SLCCU account balances
(Same as Table 7 Columns 5-8 but here with different functional forms of outcome variables)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Account balances ($ hundreds)					
	Savings			Savings + checking		
	Winsorized (95%)	Winsorized (99%)	Inverse Hyperbolic Sine	Winsorized (95%)	Winsorized (99%)	Inverse Hyperbolic Sine
Sample:	Full					
Panel A. Main effects						
CBL Arm * Post	0.329	0.970	0.069	0.128	1.034	0.036
	(0.279)	(0.615)	(0.058)	(0.495)	(0.909)	(0.084)
Observations	6124	6124	6124	6124	6124	6124
Individuals	1531	1531	1531	1531	1531	1531
Panel B. Heterogeneity by installment credit activity at baseline						
CBL Arm * Post * Bottom tercile of installment credit activity at baseline index (i)	0.112	-0.143	0.013	0.185	0.152	-0.023
	(0.405)	(0.991)	(0.090)	(0.713)	(1.508)	(0.134)
CBL Arm * Post * Middle tercile of installment credit activity at baseline index (ii)	0.082	1.376	0.102	-0.509	0.721	0.033
	(0.552)	(1.116)	(0.106)	(0.983)	(1.683)	(0.156)
CBL Arm * Post * Top tercile of installment credit activity at baseline index (ii)	0.801	1.847	0.095	0.764	2.416	0.133
	(0.509)	(1.153)	(0.106)	(0.903)	(1.643)	(0.150)
p-value of (i) = (ii)	0.965	0.309	0.522	0.568	0.801	0.786
p-value of (ii) = (iii)	0.339	0.769	0.964	0.340	0.471	0.645
p-value of (i) = (iii)	0.290	0.191	0.555	0.615	0.310	0.439
Observations	6008	6008	6008	6008	6008	6008
Individuals	1502	1502	1502	1502	1502	1502
Mean dependent variable in Extra Step Arm at baseline	2.160	3.724	0.739	4.053	6.088	1.016

Unit of observation is a person-SLCCU data snapshot, with four observations for most persons at roughly the same timing as our credit report pulls: baseline, and three endlines at 6, 12, and 18 months post-treatment assignment, all three of which are included in the Post indicator for the experiment period. Standard errors, in parentheses, are clustered at the person-level. Each column-panel presents results from a single OLS regression of the dependent variable described in the column heading on the variable(s) shown in the rows, Post and person fixed effects (Panel A), with Panel B regressions including Post * Bottom tercile of installment credit activity at baseline index, Post * Middle tercile of installment credit activity at baseline index, and Post * Top tercile of installment credit activity at baseline index instead of the Post indicator. All outcome variables here are calculated from SLCCU administrative data. Balances are recorded as zero for those who leave the credit union. Outcomes in columns (1) and (4) replace observations greater than the 95th percentile with the value of the observation at the 95th percentile. Outcomes in columns (2) and (5) replace observations greater than the 99th percentile with the value of the observation at the 99th percentile.