Automated SMS Training and Micro-Entrepreneurship Performance

William Fuchs[*] Ganesh Iyer[†] Przemyslaw Jeziorski[‡]

July 19, 2023

Abstract

We study the effectiveness of a massively scalable automated SMS business training program in improving business outcomes and in promoting financial resilience during a time of economic distress. The technology involves negligible marginal costs and requires only basic cell phone coverage. Our Randomized Control Trial included nearly 13,000 micro-retailers in Kenya and manipulated the timing of the training delivery to establish a lower bound for the intent-to-treat during a holiday period. We find that SMS training significantly increases revenues and profits, promotes formal bookkeeping practices, and improves understanding of financial concepts, and thereby promotes financial resilience. Training can also significantly reduce the need for asset liquidation for the repayment of debts. The effect of training on profits is only significant for retailers with low prior knowledge as proxied by the lack of a college education, which points to knowledge as a mediating mechanism for the impact of training. We also show that the return on investment of the SMS training is an order of magnitude larger than that of conventional in-person training.

[*]william.fuchs@mccombs.utexas.edu, UC3M and UT Austin, McCombs School of Business
[†]giyer@berkeley.edu, UC Berkeley, Haas School of Business
[‡]przemekj@berkeley.edu, UC Berkeley, Haas School of Business

1 Introduction

The literature in economics has long stressed the importance of the adoption of business practices for firm performance.[1] Well-managed firms have greater productivity, profitability, and ultimately higher survival rates (Bloom and Van Reenen, 2007). For example, recent study by Anderson, Iacovone, Kankanhalli, and Narayanan (2022) reports a field experiment amongst retailers in Mexico which shows that adopting modern business practices leads to better sales performance. However, despite the potential benefits, the literature documents a lack of adoption of even basic business practices among millions of small entrepreneurs and firms, who are an essential part of the economy in developing countries.[2] Consequently, significant public and private efforts have been devoted to introducing modern business practices and training small entrepreneurs. An obvious challenge, however, lies in designing cost-effective programs at scale that can reach millions of small and micro-entrepreneurs and improve their business outcomes.

Past training efforts in the literature mostly take the form of traditional in-person business training or intensive business consulting. However, as a recent review by McKenzie (2020) documents, the existing evidence from studies across the world has generated skepticism about the cost-effectiveness of these traditional in-person business training methodologies in improving outcomes for small entrepreneurs. In a similar vein, McKenzie and Woodruff (2014) reviews thirteen of the published studies and report statistically significant effects only in two studies. One possible way to improve cost effectiveness is to deploy scalable and automated online training programs.

This paper provides the first large-scale experimental evidence that an interactive, automated, and scalable form of training based on SMS (short message service or text

[1] see (Abowd, Haltiwanger, Jarmin, Lane, Lengermann, McCue, McKinney, and Sandusky, 2005; Syverson, 2011).

[2] see for example (McKenzie and Woodruff, 2017)

message) technology is effective in inducing the adoption of modern business and financial practices such as book-keeping and lending, and in increasing business profitability. We present the outcome of a country-wide large-scale randomized controlled experiment involving nearly 13,000 small micro-retailers in Kenya. The experiment was conducted in partnership with Mastercard's Center for Inclusive Growth and Arifu which is a Kenya-based company that creates digital content and interactive learning platforms and was responsible for deploying the SMS chatbox technology based training to foster financial resilience. Thus, the intervention in our study is a commercially relevant digital training program deployed by a business partner, rather than a training program developed by the researchers as is typical in many existing studies. The training provided small micro-retailers with the necessary business skills, including financial management, merchandising, safe financial practices, and business funding options. The study involved designing and supervising the randomized control trial to evaluate the effectiveness of the SMS training program.

Our study contributes on a second dimension by providing a unique opportunity to explore the effectiveness of a digital business training program during a period of economy-wide financial distress caused by the COVID-19 pandemic. Little research if any exists that sheds light on whether human capital accumulation through training can help entrepreneurs weather significant negative business shocks. Our objective is to evaluate if this automated and scalable form of training can serve as a viable tool to bolster financial resilience during times of significant economic crisis. This is an especially crucial issue because as McKenzie and Paffhausen (2019) show about half the small firms in developing economies operating at a given point in time die within the next six years (i.e., 8,2% per year). This effect should be even more pronounced in times of wide-spread financial distress.

Performing an impact evaluation during times of financial distress poses unique methodological and ethical challenges. Traditional approaches that withhold support from a control group can raise both ethical concerns (due to full withholding of support when it's needed) and statistical inference issues (especially if the anticipated support affects agent behavior). These concerns should be heightened in periods of financial distress following the pandemic, when the need for support available to small marginal entrepreneurs is likely to be most acute. To deal with these challenges, we designed and deployed an alternative approach that manipulates the timing of the training delivery. Based on the information from our local partners December was established as the ideal time to do the training because the holidays would provide more time for the shopkeepers to engage with the training platform. In our design, a randomly selected subset of retailers received the training in November — on an accelerated schedule compared to the ideal December deployment. Indeed, our data allows us to verify that, as expected, due to the higher opportunity cost of time, there was lower engagement by the group treated in November. The design allows us to use the November group as a quasi-control arm, providing a conservative lower bound estimate of the training's effect in December.

Crucially, we prove that the validity of this approach only rests on a "no-harm" assumption – that the training in November did not have a negative effect on participant outcomes. We provide two justifications for this no-harm assumption. First, we document that the actual empirical incidence of negative effects in the training in the existing literature are rare. Second, and more importantly, we provide an empirical test of the assumption showing that more than 90% of retailers who engaged with the November training found it helpful for mitigating their financial distress. Indeed we believe that using this type of design which leverages the variation in engagement should be valuable in scenarios where establishing a pure control group is infeasible, and where panel data cannot be collected. Moreover, it offers a feasible and theoretically structured method to evaluate the impact of interventions during crises such as the COVID-19 pandemic, when rapid, scalable responses are needed and when fully withholding support for the sake of research is not an option.

The objective of our analysis is to measure the causal impact by evaluating the intent-to-treat (ITT) effect of training on three types of outcomes. First, we solicit self-reported manipulation checks that provide preliminary evidence that the training was more effective in the treatment group. In particular, more subjects in the treatment group reported that they received any business training, and more chose business training as the most helpful type of assistance during the pandemic. In addition, more shop owners in the treatment group reported that SMS training impacted their business practices.

Beyond the manipulation checks, we measured two types of outcomes: intermediate inputs, and final business performance metrics. We find that treated entrepreneurs have higher revenues and are more resilient to the COVID-19 shock. In particular, 8 percentage points (p.p.) fewer entrepreneurs in the treatment group reported that they had liquidated their property to pay off their debts. We also find that significantly more entrepreneurs in the treatment group understand what a financial interest is, engage in some form of book-keeping, and more frequently check their retail credit card balance. These improvements in behavioral measures, business practices and the need to liquidate assets are potential mechanisms to generating higher revenue and greater financial resilience.

Further, we collected longer-run performance measures to demonstrate the effect of training over time and to rule out memory and salience effects. We deployed a shorter version of the survey six months after the training and found a significant impact of the treatment on daily and monthly profits. Finally, we show that the longer-run profit impact of training is only significant for entrepreneurs without extensive formal education, i.e., those without a college education. This heterogeneous effect suggests that SMS training can alleviate potential knowledge gaps caused by limited access to formal education.

1.1 Comparing In-person and Automated SMS Training

The experience of mobile money across various economies in Africa highlights the potential of digital communication technology such as SMS to revolutionize access to training. Like banking, traditional in-person training programs face several inefficiencies: First, in-person training programs have significant difficulty and cost of reaching intended audiences. This friction is substantial when the population is rural and geographically dispersed. Second, in-person training programs involve time investments from small business owners and may require them to take time off during business hours to attend the training. The requirement to take time off may lead to non-trivial opportunity costs of foregone business resulting in low take-up rates.

Digital training through SMS technology has the potential to alleviate these inefficiencies due to its flexibility and its ability to deliver training at negligible marginal costs. Training may take place at the business owner's choice of location and time, and the owner can decide to learn the content in smaller, more manageable packets. In this manner, training through SMS delivery can decentralize the learning to the business owner's choice of what, when, and how much to learn, potentially leading to greater engagement. Although one might be skeptical of the effectiveness of SMS as a medium to teach business knowledge and change behavior, there exists some evidence that over-the-phone solutions may be effective. In particular, Cole and Fernando (2021) find that a mobile phone advice hot-line service staffed by human operators improves intermediate inputs and the final output for farmers in Gujarat, India.[3] The effect is not as significant as when consultants

[3]In addition, Anderson and McKenzie (2022) find that an intermediate solution of mixing online and in-person classes may also be successful. In contrast to both these studies, the delivery of SMS training

are deployed on-site (Bloom, Eifert, Mahajan, McKenzie, and Roberts, 2013) or when firms get intensive in-person training. However, given that the per-unit cost of the decentralized virtual solutions is significantly smaller (our SMS training costs Mastercard approximately $2.30 per user), they may have the potential to deliver higher returns per dollar invested than the more expensive in-person programs.

Compared with previously deployed virtual training, the training technology in this study is automated and does not require human trainers or consultants. Further, access is decentralized because the content is available at any time. Automation also removes the labor costs of trainers, and lowers the marginal cost of the deployment while providing a significantly shorter time-to-market.[4] Lastly, the SMS technology also enables reaching disadvantaged entrepreneurs who are at the fringes of the cellphone towers where robust voice reception is not supported, but where SMS is still operational. Thus, even if the identified effects are modest when compared to human-driven solutions, it is possible that an automated, SMS-based solution may still have a superior return on investment while at the same time improving inclusion. Therefore, an important component of the study was to demonstrate proof of concept for using automated and decentralized digital platforms as an effective, low-cost, and scalable training solution for small merchants. This is facilitated by the fact that the context of our study involves Mastercard offering the Jaza Duka program to more that 40,000 retailers in Kenya and nearly 13,000 of them opened a credit line and were part of the SMS training program. Thus the deployment of the interactive digital training here is at the country wide level and truly large scale.

used in our study can be fully automated.

[4]Indeed, our partners were able to deploy the SMS training to thousands of entrepreneurs within a matter of weeks.

2 Experimental Design

The randomized field experiment manipulates the timing of access to the Arifu SMS training deployed across Kenya. The experiment was executed as part of 'Jaza Duka,' an innovative credit initiative for small micro-retail entrepreneurs launched at the beginning of 2019 by Mastercard in partnership with Unilever and KCB Bank Kenya Limited.[5] Jaza Duka offers a credit line to small retailers to alleviate financial constraints by providing them with working capital. Notably, the credit qualification of the retailers relied only on the history of purchases from Unilever. It did not require the retailer to have a prior credit history, bring collateral, or be part of a lending group. For these reasons, Jaza Duka became the first genuinely accessible loan for many small retail entrepreneurs. MasterCard offered the program to more than 40,000 retailers across Kenya, of which nearly 13,000 retailers opened the credit line.

Jaza Duka is a relatively complex credit product, resembling a credit card with a 17-day interest-free grace period and minimum payment requirements. Therefore, the roll-out was accompanied by several training programs executed by MasterCard's Center for Inclusive Growth. The training programs varied from intensive, high-cost, traditional in-person business training, to a lighter-touch low-cost, scalable, interactive Short Message Service (SMS) training, which is the subject of this study.

The SMS training was designed and executed by Arifu, an experienced Kenyan interactive digital learning platform provider. The SMS training was executed as part of an extensive COVID credit relief initiative for Jaza Duka merchants. As of September 2020, due to the adverse impact on sales following COVID restrictions, nearly 50% of the Jaza Duka credit lines were restricted for non-payment or else were closed entirely. In collaboration with Unilever, MasterCard designed a credit relief package that included

[5] 'Jaza Duka' means "fill up your store" in Swahili.

credit relief, accounts top-ups, and an SMS based financial resiliency training program.

The primary objective of this study is to quantify *intent-to-treat* (ITT) of the SMS-based training program that was implemented during the holiday season of December 2020. The design was driven by the documented difficulty of providing statistically precise estimates of the causal effect of business training in situations of urgent need. Given this difficulty, our study strives to find ideal conditions for training delivery which would amplify the program's efficacy and also result in necessary statistical power. Based on detailed consultations with the training provider and based on their experience, it was agreed that the holiday periods would yield optimal results for training deployment and engagement. The underlying logic stems from the historical data which show elevated participation in SMS training during these periods. Greater engagement can be due to lower opportunity costs associated with interacting with the chat-bot during holidays, a time when individuals typically have more leisure time at their disposal. However, it is important to acknowledge that our choice also involves possible caveats. We elaborate on these caveats and discuss the robustness of the results to despite their presence.

Under the standard classical framework of a Randomized Controlled Trial (RCT), assessing the effectiveness of the training program in December 2020 would rely on the creation of a control group. This group would either be exempted from the training altogether or would be allowed to participate in the program only at a significantly later point in time. However, applying such design to any highly anticipated training program in the time of need following the pandemic poses significant statistical as well as from an ethical considerations, making this conventional design less than ideal. The crux of the issue lies in the potential ramifications of depriving a substantial number of stores that are in default on the Jaza Duka credit of timely training.

Specifically, if we were to withhold training from a large cohort of retailers, it's likely that the Stable Unit Treatment Assumption (SUTVA) - an essential tenet of a well-designed RCT - could be violated. This is due to the ubiquity of awareness across Kenyan retailers regarding the impending training program, regardless of their group assignment in the trial. If retailers in the control group became aware of their exclusion from training, or the intentional delay of the same, it could provoke a negative reactions towards the program. Retailers in the control group could also learn about whether other proximate retailers in their local market received training and could react strategically, further jeopardizing SUTVA. This limitation presents a challenge to any such large-scale intervention, especially when it is deployed in a time of need.

An obvious solution is to study programs which are not during times of need or to restrict deployment to smaller test markets. But in this case Mastercard deployed the training program as part of a relief package precisely because of the need and the financial distress created by the pandemic. Further, we believe that providing evidence for the impact of a highly scalable program deployed during a time of financial distress could be a useful contribution in and of itself, and this cannot be achieved through a restricted deployment. Regarding the ethical concerns about the withholding of treatment in a control group, a typical solution would be a staggered randomized design. In such a design, subjects are randomly assigned to receive an intervention at different times; but one obtains a baseline measurement to serve as a control before receiving the intervention. Such a design was however infeasible to implement given the urgency of the situation, since a robust baseline measurement would have taken at least a month to set up and implement. Another possibility would be a clustered randomization approach, which does not alleviate the ethical issues but helps in establishing SUTVA. Unfortunately, in our case, the signal to noise ratio when evaluating business training is known to be unfavorable; thus, clustering might require infeasible sample size for delivering acceptable power.

Our experimental design addresses the challenges described above and takes advantage of a unique opportunity to study the impact of SMS training during a critical period while mitigating participant harm and plausibly satisfying SUTVA. Accordingly, we introduce a minimal perturbation of the timeline for the receipt of training, which is expected to result in a substantial manipulation of the opportunity cost of taking the training. Specifically, in our design a randomly selected subset of retailers was offered expedited training in November 2020, instead of the initially planned December 2020 timeline. Retailers offered the training earlier did not have the option to defer the training to December. This design allows us to estimate a lower limit of the December intent-to-treat effect, given certain additional requirements which we show are empirically justifiable.

The specifics of the design is discussed below using a potential outcomes framework that allows us to outline the assumptions and provides a better understanding of the methodological choices. The RCT was composed of three distinct arms based on the choices of the timing of SMS training and the offer of credit relief among a sample of 7,400 Jaza Duka retailers who had taken credit under the program. First, a random 30% of retailers from the sample received both SMS training and credit relief in November 2020 (arm 1: early training; early credit). Following this, another randomly picked 40% of the sample got early SMS training in November 2020, but with credit relief deferred to December 2020 (arm 2: early training; later credit). Finally, the remaining 30% of retailers received their SMS training in December 2020, whilst their credit relief was given in November 2020 (arm 3: later training; early credit).[6] Henceforth, we focus on the training manipulation by considering two estimation strategies. We either merge the first two arms and statistically account for when credit was offered, or compare arms 1 and

[6]Note that an additional 5,138 retailers were not eligible for credit relief and were randomly allocated to the November/December training arms in equal proportions. However, these retailers were not surveyed and their data is not included in this paper.

3, both providing early credit relief. We examine the possible implications of randomizing credit and training jointly when discussing the results.

2.1 Characterization and Justification of the Design

Consider a randomized design with 3 arms, Z_i, given by: CONT (perfect classical control), NOV (November training offer), and DEC (December training offer).[7] Note that in the context of our experiment the arms NOV and DEC are the ones that are implemented and the CONT arm would be hypothetical. Denote the endogenous training compliance by $D_i(Z_i)$. We set D_i = NOV, or D_i = DEC, if the individual decided to engage with the training in a respective month, and $D_i = 0$ otherwise. If the individual was assigned to control then $D_i = 0$, by definition. Further, according to the implementation of the design consumers cannot choose when they get trained, that is,

$$P(D_i = \text{DEC}|Z_i = \text{NOV}) = 0 \quad \text{and} \quad P(D_i = \text{NOV}|Z_i = \text{DEC}) = 0.$$

The design is general enough to allow for differential selection into training in a respective month, that is,

$$P(D_i = \text{DEC}|Z_i = \text{DEC}) \neq P(D_i = \text{NOV}|Z_i = \text{NOV}).$$

Recall as previously discussed one of the rationales for implementing our design is that settings such as ours a perfect classical control may violate the SUTVA assumption. Nevertheless, imagine for now that the hypothetical 3-arm design does not violate SUTVA and that unconfoundedness holds. Under SUTVA the potential outcomes for this experiment will be given by $Y_i(Z_i, D_i)$. In our case, Y_i is either the shop revenue or other business outcomes measured in March of the following year. Unconfoundedness will

[7] A perfect 3-arm experiment is conceptually equivalent to two separate 2-arm experiments – randomizing training offers in November and December and executed on two identical populations.

mean that the assignment of arms (different months of the training offer or control) only affects the outcome through the training and not directly. In other words, it ensures that the relationship between arm assignment and outcomes is solely mediated by training.[8] Mathematically, unconfoundedness is expressed as:

$$Y_i(Z_i, D_i) = Y_i(Z_i', D_i); \forall Z_i, Z_i'$$

It is important to note that unconfoundedness does not imply that the month of training does not matter. In fact, the outcome effectiveness of training in December can be different from that of training in November. This is reflected in the condition:

$$Y_i(Z_i, D_i = \text{NOV}) \neq Y_i(Z_i, D_i = \text{DEC}).$$

Under unconfoundedness the potential outcomes can be written as $Y_i(D_i)$. We can then define the ITT effect of the training offer given an observed covariate matrix X by the following equation:

$$\text{ITT}^{\text{Z}} = E[Y_i|Z_i = Z, X] - E[Y_i|Z_i = \text{CONT}, X]; \quad \forall Z \in \{\text{NOV,DEC}\}$$

Our objective is to estimate ITT^{DEC}. Given that the control is assumed to be exogenous, a simple comparison of the means between each treatment arm and the control would provide us with the ITT for training in both November and December. However, the classical control is not available in our study – we only observe the two treatment arms, namely NOV and DEC. To proceed further, consider the following proposition:

Proposition. *Define a treatment dummy variable*

$$M_i = \begin{cases} 1 & \text{if } Z_i = DEC \\ 0 & \text{if } Z_i = NOV \end{cases}$$

[8]One example when unconfoundedness could be violated is if Y_i was measured in December for the December arm, and in November in November arm. All our outcomes each arm are measured in March of the following year.

If the option of engaging in training in November can only help or have no effect, but cannot harm, i.e., $ITT^{NOV} \geq 0$, then $ITT^{DEC} \geq E[Y_i|M_i = 1] - E[Y_i|M_i = 0]$.

Proof.

$$E[Y_i|M_i = 1] - E[Y_i|M_i = 0] =$$

$$E[Y_i|Z_i = \text{DEC}, X] - E[Y_i|Z_i = \text{CONT}, X] -$$

$$(E[Y_i|Z_i = \text{NOV}, X] - E[Y_i|Z_i = \text{CONT}, X]) =$$

$$\text{ITT}^{\text{DEC}} - \text{ITT}^{\text{NOV}} \leq \text{ITT}^{\text{DEC}}$$

$\square$

Since M_i is randomized, one can obtain consistent estimates $E[Y_i|M_i = 1] - E[Y_i|M_i = 0]$ by an OLS regression:

$$Y_i = \alpha M_i + \beta X_i + \epsilon_i, \tag{1}$$

where X_i are observable covariates orthogonal to M_i. In other words, if the no-harm assumption is satisfied, we can use the training in November as a quasi-control; that is, if we find that the December-trained retailers perform better than November-trained retailers, we can conclude that training was at least as effective as the gap providing a conservative lower bound of the training effect.

In what follows, we discuss the existing literature that shows evidence and provides justification for 'no-harm' assumption. The meta-analysis of McKenzie (2020) provides findings concerning the effectiveness of business training by taking into account 14 individual studies examining the effect of training on profits and another 17 that examines the impact on sales. Notably, none of the 31 studies discern a negative impact at a 5% level, although many confidence intervals at the individual study level are broad and include zero. The aggregated analysis proffers an overall estimated effect on profits of 10%, with a standard error of 3.06%, and an impact on sales of 4.7%, with a standard error of 2.3%. Both estimates refute the notion of a negative impact of business training at a 5% level.

Further evidence comes from the meta-analysis performed by Cho and Honorati (2014) which examines 37 impact evaluation studies containing 1,116 estimates for six types of outcomes pertaining to income/profits, labor market activities, and business knowledge. The type of interventions in the sample of studies can be classified as business training, financing, and business counseling, which align well with the Arifu SMS training program. The studies cover a broad range of 25 countries across all six continents, including Africa, which accounts for 17% of the studies. Experimental interventions account for 80% of the estimates. The studies use profits, business expansion, and business performance metrics (such as the size and revenue) as outcome variables. The overall results are that the effects of business training programs can vary; but they are rarely negative. Precisely, only 2.8% of studies detect negative effects at a 5% level, whereas 28.2% identify positive effects. The majority of the studies, approximately 68%, detect no effect. In the context of our design this is a useful finding: If training was indeed ineffective in November, the bound in our Proposition would be exact. In other words, November would then serve akin to a true control for the December treatment.

Support for the "no harm" requirement also comes from the findings pertaining to the impact on business practices. Conceptually, it seems obvious that training cannot hurt when the outcome Y_i are decisions to adopt superior business practices, for example bookkeeping. For training to be detrimental, we would need a sufficiently large number of retailers that previously engaged in bookkeeping to abandon it as a result of the business training – a possible, albeit unlikely scenario. Indeed, a mere 1.9% of estimates signal a negative impact for business practices at a 5% level, compared to as much as 31.6% of the estimates being positive indicating strong support for the "no-harm' assumption.

Turning to store profits as outcome variables, a negative average effect would require either risk-loving preferences leading to over-adoption or else some degree of irrationality. Indeed the existing literature points to under-adoption rather than over-adoption of business training. For instance, Karlan and Valdivia (2011) find that entrepreneurs who express the least interest in business training could benefit significantly from it. This leads us to expect attenuation from under-adoption rather than excessive adoption or any ex-post harm from risk taking. In the Appendix A we provide additional discussion of factors for negative impact and why they are not applicable in the context of our study. In the next section, we describe our data and furnish direct quantitative evidence corroborating the 'no-harm' assumption for November.

3 Data

We conducted two telephone surveys of 1,500 stores proportionally distributed across the three experimental arms. The first survey was conducted in March 2021, and the second was conducted in July 2021. The first survey population was randomly drawn from the available pool of subjects until we reached a target of 500 respondents per arm. The surveyors tried to connect with a potential subject up to three times. If the potential subject did not pick up the phone, they were randomly replaced. Because the targets were specified per arm, it is important to test for selective attrition. The attrition rate was indeed quite low, approximately 9% in the December arm and 8% in the November arm and the numbers are not statistically different at 5% level. To further explore the possibility of selective attrition, we perform randomization checks using the completed survey responses later in this section.

The first survey was approximately 30 minutes and solicited basic demographic information, proxies for treatment intensity, and an array of outcome measures. The second survey which aimed to measure the longer-run impact of the treatment was shorter (15 minutes) and contained only basic demographic information and a smaller subset of outcome measures. All respondents that completed the initial survey completed the second survey.

Table 1 contains summary stats. More than 95%, or 1,434 stores, completed the survey after picking up the phone. The top panel of the table presents chosen summary statistics from the initial survey. First, we gauge if the retailers remember receiving Arifu's SMS training by asking, *"Did you receive SMS training as part of Jaza Duka?."* We find that 88% of respondents said that they received SMS training. Based on Arifu's records, approximately a similar proportion of subjects interacted with the SMS chatbot. Next, to obtain a self-reported measure of training efficacy, we asked the respondents the following question "*What types of assistance (financial or non-financial) do you think is the most important for you to survive in the business? Please list all that apply*". The possible options were: (i) More loans, (ii) Rent or bill help, (iii) Wage subsidy, (iv) Business training, (v) Lower wholesale price, and (vi) Other (open-ended response). About 27% of respondents included business training as the most important type of assistance. Lastly, to all those respondents that report receiving SMS training, we asked the following question: "*How have your business practices been impacted by training?*" Nearly 78% (which is 88% of those that responded "yes" to the first SMS training question) reported an impact of the training on their business practices. Overall, these aggregate statistics provide suggestive evidence of the respondent's beliefs about the impact and the effectiveness of the SMS training.

The average monthly revenue of retailers in the sample was 137,213 Shillings, which as of the January 2022 exchange rate, amounts to approximately 1,200 USD. We asked if the store keeps written financial records and whether the retailer understands the concept of a loan interest. Nearly 80% of stores mention that they keep written financial records, and 80% of the retailers mention that they understand the concept of a loan interest.

One of the frictions that retailers face in using Jaza Duka is that they may have attention costs of monitoring the credit card's balance and the monthly payment amounts. Therefore, a part of Arifu's SMS training was designed to make the trainees more attentive to checking their balances and payments. To obtain insight into this possible attention friction, we asked, "When was the last time you checked your Jaza Duka balance." The answers are reported in Table 2. We asked this question only to retailers that recalled using Jaza Duka (i.e., 604 stores). More than 35% did not check their balance for more than a week despite using Jaza Duka. We constructed a checking frequency score by coding the responses from 0 to 6, where "earlier today" is 6 and never is 0. The average frequency score as reported in Table 1 is 3.61.

The Jaza Duka credit line and the relief package, which included the SMS training, aimed to alleviate the liquidity problems retailers faced due to the adverse impact of the pandemic on sales. To estimate the extent of the liquidity gap, we asked "*Did you have to sell any of your belongings to pay back loans in the last 3 months?*" Approximately 9% of respondents answered "yes" to this question, suggesting a liquidity shortage.

The shorter follow-up survey in July 2021 was designed to estimate the longer-run impact of our intervention on key business performance measures such as revenue and profits. The four measures are reported in the bottom panel of Table 1. The average revenue increased from 137,000 to 190,000 Shillings. This increase is primarily due to seasonality and the easing of COVID-19 restrictions. We solicited both daily and monthly revenues and profits to decrease the possible measurement errors of the self-reported measures.[9] Store owners may account for fixed costs differently when estimating daily

[9]Overall, daily numbers are consistent with monthly numbers after accounting for the fact that most

and monthly profits. In particular, we conjecture that fixed costs are less likely to be included in the estimates of daily profits because of the monthly cost payment cycles. In contrast, fixed costs are more likely to be included in the estimates of monthly profits. Indeed, we find that the average margin of the store is approximately 11% and 9% when using daily and monthly numbers, respectively. Lastly, we obtained a pre-experimental measure of credit purchase volume in November 2020. This measure, as discussed below, would serve as one of the randomization checks and it is reported in the third panel of Table 1.

In the online Appendix B we report the randomization checks. We perform randomization checks using individual characteristics that are not affected by the timing manipulation, such as gender, age, and education. In particular, $1,433$ respondents answered questions about their gender, while $1,432$ answered their age and education questions. None of these characteristics vary in a statistically significant way between the November and December arms. We also obtained the data from Mastercard on actual credit purchases in the first week of November 2020, which is before the any training roll-out. We compare these purchases across December and November for the population that responded to the survey and do not have statistically significant differences. These results suggest that the randomization was implemented correctly.

Next, we provide support for the no-harm hypothesis by providing evidence consistent with a non-negative impact of training in November. Using data from the March 2021 survey we can examine the perceived ex-post value of the training among the complying retailers. Among the subjects offered SMS training in November, 90% of those who reported receiving any business training considered this training to be crucial to their business survival. Also relevant is that 88% of those who reported receiving the SMS

stores are closed on weekends and for significant stock-outs. Specifically, the daily numbers are approximately 17 times smaller than monthly.

training perceived a significant impact of this training on their business practice. These reported evidence support the validity of the no-harm hypotheses for our study.

In the next section, we present a stylized, testable model that elucidates why we expect distinct ITTs in November and December. The model offers insight into the mechanisms operating behind our timing manipulation.

4 Stylized Model and Rationales

Within this section, we present a stylized model of the process of training adoption, emphasizing the contrasting dynamics between November and December. The model indicates verifiable forecasts regarding the intensive and extensive margins of training adoption across our two experimental groups.

Consider an retail entrepreneur, denoted as i, who faces a decision regarding the adoption of training in either November (NOV) or December (DEC) depending on a random assignment Z_i. In accordance with our design, the subject cannot choose the month of training. The entrepreneur's utility function for adopting the training is denoted as $U_i(D_i)$. This utility function represents the additional value or benefit that the entrepreneur gains by undergoing training compared to the current value of operating their shop without training.

To be more specific, $U_i(D_i)$ captures the incremental value of training over the entrepreneur's existing status quo situation. This can be measured, for example, by comparing the expected discounted stream of profits with and without training, starting from a certain month (in this case Z_i) and projecting into a relevant future horizon. It is important to note that the utility U_i is related to, but may not be necessarily equal to, the outcome variable Y_i introduced in the previous section. This is because Y_i is measured in a specific month (in our case March 2021), while U_i is measured starting from a specific month (in our case Z_i) and extending into the relevant future horizon.

The utility from training adoption can differ across individuals for various reasons. These differences can stem from variations in discount factors, the added value of skills obtained through training, the size of the entrepreneur's shop, or on in general the relative benefits of incorporating new business practices. Next, we introduce the concept of the opportunity costs of adopting training, denoted as $c_i(D_i)$. These costs are also likely to differ among individuals due to variations in time constraints and the perceived value of alternative tasks that could be performed instead of training.

Considering the assignment to a specific training arm, denoted as Z_i, an individual will decide to adopt training if the utility derived from training, $U_i(D_i)$, exceeds the associated cost, $c_i(D_i)$. Specifically, if Z_i corresponds to November ($Z_i = \text{NOV}$), the individual will adopt training if $U_i(\text{NOV}) > c_i(\text{NOV})$. On the other hand, if Z_i corresponds to December ($Z_i = \text{DEC}$), the individual will adopt training if $U_i(\text{DEC}) > c_i(\text{DEC})$.

To facilitate a deeper understanding of the mechanisms underlying training adoption in November and December, we introduce a set of assumptions regarding the economics of training adoption across these two months. It is important to note that these additional assumptions are not necessary for estimating treatment effects. Instead, they serve to provide a clearer description of the mechanisms involved in our experimental manipulation and the resulting differences in the IITs detailed in the next section. The model also provides insights into why we cannot assert an equivalent to our main Proposition for the Average Treatment Effect on the Treated (ATT).

Based on the earlier discussion, we adopt the concept of Monotone Treatment Response (MTR) introduced by Manski (1997) which in our context implies that the impact of training in December is at least as large as the impact in November. Formally, we express this assumption as follows

$$U_i(D_i = \text{DEC}) \geq U_i(D_i = \text{NOV}).$$

Additionally, suppose we assume that individuals face a lower opportunity cost of engaging in training in December compared to November – i.e., $0 \leq c_i(\text{DEC}) \leq c_i(\text{NOV})$.

The assumptions imply that the population of individuals who self-select to engage in training will differ between November and December. Specifically, we will obtain that the set of adopters is (weakly) larger in December compared to November, expressed as:

$$A(\text{NOV}) = \{i : U_i(\text{NOV}) > c_i(\text{NOV}\} \subseteq \{i : U_i(\text{DEC}) > c_i(\text{DEC})\} = A(\text{DEC}). \tag{2}$$

Since the assignment of individuals to a specific training month is randomized (i.e., the distribution of i does not depend on Z), we can derive the following inequality:

$$P(D_i = \text{NOV}|Z_i = \text{NOV}) \leq P(D_i = \text{DEC}|Z_i = \text{DEC}). \tag{3}$$

Furthermore, we can compare the average value of the training offer between the two treatment arms:

$$\begin{aligned} E[U_i(D_i)|Z_i = \text{DEC}] = \int_{A(\text{DEC})} U_i(\text{DEC})dF(i) \geq \\ \int_{A(\text{NOV})} U_i(\text{DEC})dF(i) \geq \int_{A(\text{NOV})} U_i(\text{NOV})dF(i) = E[U_i(D_i)|Z_i = \text{NOV}] \end{aligned} \tag{4}$$

where $F(i)$ represents the cumulative distribution function of individuals. The first and last steps (equalities) follow from the random assignment to the treatment arm, Z_i (i.e., $F(i)$ does not depend on Z_i). The second step (inequality) follows from (2), and the fact that integrating positive function over a sub-set yields a weakly smaller value. The third step (inequality) is implied by the MTR.

The above equations (3) and (4) are implications of the model which can be empirically tested. To test Equation (3), note that in both the November and December groups, all subjects received an invitation to the SMS training. Based on information from Arifu's, most respondents interacted with the SMS training, at least initially for the first few messages. However, for the interaction to be considered "business training," the respondents would need to complete the initial steps and reach the business-related content. Referring to Table 3, we observe that there is no significant difference in the fraction of individuals who reported interacting with Arifu's SMS chatbot between the November and December groups. However, there is a 2.7 percentage point (pp) increase (which is significant at the 5% level) in the number of subjects reporting receiving any actual "business training" in the December group compared to the November group. These results are not surprising since most retailers remember receiving the invitation to the SMS training, but this would not be regarded as business training if they did not reach or barely started engaging with the training content. Nevertheless, the regression results regarding the increase in the number of subjects receiving business training indicate a higher adoption of training in December, which supports the stylized model.

To test the implication of Equation (4) (that the value of the training offer is larger in December), we measured the retailers' perceptions of the usefulness of training, U_i. Our inquiry into the shop owners' perception of the influence of the SMS training on their business practices unveiled that 4.3 pp more subjects in the December group reported a noticeable impact on their business practices. Moreover, when we inquired about the type of COVID assistance they found most useful, nearly 12pp more subjects in the December group identified business training as the most beneficial form of assistance they received.[10] These findings further reinforce the efficacy of our timing manipulation in

[10]The exact question was: "What types of assistance (financial or non-financial) do you think is the most important for you to survive in the business? Please list all that apply." The options are: 1. More loans, 2. Rent or bill help 3. Wage subsidy, 4. Business training, 5. Lower wholesale price 7. Other (open question). The focal outcome is a dummy variable equal to 1 if option 4. was one of the chosen options.

successfully influencing the factors driving engagement in the training program, aligning well with the predictions of our stylized model.

It is worth emphasizing that this simple model permits us to deduce the rank order of the "offers of training," or $E[U|Z]$ (ITT in utility terms), across months. However, the rank order of the value of training for those who received it, or $E[U|D]$ (ATT in utility terms), remains undetermined. This ambiguity stems from the fact that the training value for trained retailers is dependent on the distribution of treatment effects. For example, the effect on those treated could be more pronounced in November if, due to a higher opportunity cost $c_i(NOV)$, the training is adopted exclusively by those with the most potential for gain (highest types). For the same reason, we cannot assert an equivalent to our main Proposition for the ATTs. Accordingly, in the ensuing section, we estimate and discuss ITT in December.

5 Results

As mentioned earlier, the experiment consisted of three arms that manipulated the timing of deployment of training and credit relief: i.e., (training now; credit now), (training now; credit later), and (training later; credit now). Given the requirements of Mastercard, we did not implement the training randomization independently from credit relief randomization. Therefore, we estimate the ITT of deploying training during holidays in December 2020, keeping the credit deployment constant. We first present results based on estimates that control for credit deployment and then show robustness by re-estimating our main results by comparing the (training now; credit now) and (training later; credit now) arms.

Table 4 contains the main results. The point estimates suggest that monthly store revenue increases by more than 90,000 KSh, constituting 60% of the average revenue.

The point estimates are somewhat noisy, albeit statistically significant at a 5% level. The noisy estimates are likely caused by measurement errors embedded in the shop owner's reported monthly revenue estimates. A conservative measure, a lower bound of 95% confidence interval, would estimate a 1.5% impact on monthly revenue.[11] However, beyond statistical conservatism, this measure is also an understatement of the impact of training because it uses a quasi-control group that also received the SMS training but not during the December holidays.

We then look at the impact of training on less noisy and intermediate measures, such as the maintenance of written financial records and the understanding of loan interest. Nearly 7 pp. more shop owners report adopting financial record keeping after receiving the training during the December period (vs. receiving training in the non-holiday November period). In other words, we observe increases in the prevalence of financial records by approximately 10%. We observe similar effects on confidence about understanding financial concepts, such as an interest rate.

Next, we measure the ITT of training on financial attention and the checking of the Jaza Duka balance. We find that the effect is significant and increases the attention score by 0.5. Lastly, we examine the impact of the training on the need to liquidate assets. Here, the impact is the most dramatic. Training during the holidays led to a more than 8 p.p. decrease in the need to liquidate property to pay for debts. This decrease is particularly relevant since we delivered the training in the middle of the COVID-19 pandemic. The substantial decrease in asset liquidation need indicates that the seemingly light-touch SMS

[11]The online appendix reports confidence intervals for the main outcomes. It also contains further analysis aimed at obtaining more precise estimates. First, we considered trimming of the largest values of the revenue variable at 1st percentile, 5th percentile and 10th percentile. We obtain more precise estimates with similar magnitudes of the main effects. We also re-estimated the model using quantile regression, which should be less susceptible to outliers. Finally we applied several parametric models, including Poisson, and Tobit. Using the alternative models delivers the same results as with trimming – namely, better precision and similar magnitudes of the main effects.

training intervention can be useful in building the micro-retail entrepreneurs' financial resilience to large demand shocks.

Table 5 contains estimates of longer-run effects of training.[12] We find that 6 months after training deployment, it still affects business performance and increases daily and monthly profits by 20-30%. The estimates are noisy, which is expected when measuring longer-run impact, albeit they are significant at 10% level.

To show robustness of the main results, we also compared the estimates of the (training now; credit now) with the (training later; credit now) arm. As shown in Table 6 the main results are qualitatively similar to what we describe above in this section. This suggests that the interaction between the training and the credit interventions in this program was minor.

Recapping the results, we find a positive impact of SMS training on intermediate inputs (business practices), such as bookkeeping and on knowledge and the understanding of financial concepts. We also find an impact on the final output measures such as revenue and longer-run profits. However, it remains an empirical question if the intermediate inputs mediate the effect on output or if both are affected by the training simultaneously. In other words, might it be that the effect of training on the intermediate inputs serves as a mechanism to improve the final outcomes measures. To investigate this mechanism, we examine below the impact of training based on the prior knowledge of entrepreneurs.

We use formal education level as a proxy for prior knowledge. According to the survey, less than 1% of subjects have less than the 6th grade of formal education, 77% have between 6th and 12th grades, 2% have post-high-school technical training, and 19% have a college degree. First, we check if formal education correlates with business practices in the

[12]Beyond the estimation of longer-run effects themselves, we include this analysis because we want to rule out recency effects – namely, training deployment closer to the first survey may cause memory or salience effects. Such effects may be potentially at play for revenue elicitation, but should be less important for more objective measures such as the prevalence of written records or property liquidation.

November group. If training had negligible impact in November, analyzing the November group provides baseline association of education and adoption of business practices. We find that 30% of owners without a college education do not have written records, compared to 20% for the university educated. Similarly, 22% in the non-college group "do not fully understand interest," compared to 15% in the university group. Next, we show that the effect of training depends on prior education. Table 7 shows that training works to improve profits only for entrepreneurs without a college education. This heterogeneous effect suggests that the knowledge gap is a mediating factor for the impact of training.

Finally, we also considered the effect of shop ownership as a possible mechanism that might mediate the impact of training. As seen in Table ?? affects the profit outcomes (business practice adoption outcomes??) of entrepreneurs who report that they are shop owners and it is ineffective for non-owners. This suggests that ownership rights are important for entrepreneurs to use business training inputs for effective decision making that improves the adoption of business practices and profit outcomes.

6 Discussion and Conclusion

We have shown that SMS technology can provide the basis for automated training that is easily scalable, widely and readily deployable, and most importantly, effective in changing behavior and outcomes in economically meaningful ways.[13] Our massively scalable SMS training technology costs $2.30 per person and delivers between 2.4% and 60% increase (at 95% CI) in monthly profits. McKenzie (2020) reports that an average in-person training program costs $177 and delivers between 4.1% and 16.1% increase in profits. Even the most conservative comparison of return on investment (ROI) between conventional and

[13]In the context of the Covid 19 pandemic, manipulating the timing of training delivered a larger treatment effect on the revenue and financial resilience measures than manipulating the timing of monetary credit relief.

SMS training indicates that SMS training delivers an order of magnitude larger ROI. In particular, even the lower bound of SMS training ROI in our study is a 1.04% increase in profits per $1 of investment, compared to an upper bound of 0.09% increase in profits delivered by conventional training at the same cost.

The significantly greater ROI of automated digital delivery enables cost-effective deployment to training to a larger population. As demonstrated by our study, through the automated SMS technology, the training of retailers in the entire country can become practically feasible and economically justifiable. Automated digital delivery is also promising in terms of facilitating continuous learning. A potential drawback of traditional in-person training programs is their eventual one-shot nature and the lack of continuity due to low ROI. Learning and education are often continuous processes. While intensive in-person classroom training may promote immediate learning of business concepts, the long-term retention and implementation of what is learned may require continuous delivery of relevant knowledge over time.

Finally, our findings on the impact of SMS training on business practices as well as on sales outcome may highlight the importance of active learning for training programs. As McKenzie (2020) points out even traditional classroom training programs attempt to incorporate active learning mechanisms such as having participants do exercises or games to increase engagement. The automated SMS chatbox technology can very naturally promote active learning through several channels. The technology is interactive such that the retailer can respond to scenarios based on which additional relevant content can be served. The retailer can also choose what and when to learn and this leads to a more efficient needs-based uptake.

Our study was designed to examine the training was deployed during a period of need, specifically during the COVID-19 pandemic, when many of the study participants were in financial distress. We believe this factor strengthens the study's relevance, as identifying effective interventions for periods of urgent need is crucial for public policy development. Nevertheless, these circumstances render our results contingent upon the no-harm assumption. While, in our case, we have shown that this assumption is empirically testable, there may be other situations with new technologies where impact of the intervention may have to a tested in the absence of the assumption.

Funding and Competing Interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the research or materials discussed in this manuscript. Authors B and C received research support from Mastercard for the project.

References

ABOWD, J. M., J. HALTIWANGER, R. JARMIN, J. LANE, P. LENGERMANN, K. McCUE, K. MCKINNEY, AND K. SANDUSKY (2005): "The relation among human capital, productivity, and market value: Building up from micro evidence," in *Measuring capital in the new economy*, pp. 153–204. University of Chicago Press.

ANDERSON, S. J., L. IACOVONE, S. KANKANHALLI, AND S. NARAYANAN (2022): "Modernizing Retailers in an Emerging Market: Investigating Externally Focused and Internally Focused Approaches," *Journal of Marketing Research*, p. 00222437211058437.

ANDERSON, S. J., AND D. MCKENZIE (2022): "Improving Business Practices and the Boundary of the Entrepreneur: A Randomized Experiment Comparing Training, Consulting, Insourcing, and Outsourcing," *Journal of Political Economy*, 130(1), 157–209.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013): "Does management matter? Evidence from India," *The Quarterly Journal of Economics*, 128(1), 1–51.

Bloom, N., and J. Van Reenen (2007): "Measuring and Explaining Management Practices Across Firms and Countries*," *The Quarterly Journal of Economics*, 122(4), 1351–1408.

Cho, Y. J., and M. Honorati (2014): "Entrepreneurship programs in developing countries: A meta regression analysis," *World Development*, 64, 490–508.

Cole, S., and A. N. Fernando (2021): "Mobile'izing Agricultural Advice Technology Adoption Diffusion and Sustainability," *Economic Journal*, 131(1), 192–219.

Karlan, D., and M. Valdivia (2011): "Teaching entrepreneurship: Impact of business training on microfinance clients and institutions," *Review of Economics and Statistics*, 93(2), 510–527.

Manski, C. F. (1997): "Monotone Treatment Response," *Econometrica*, 65(6), 1311–1334.

McKenzie, D. (2020): *Small Business Training to Improve Management Practices in Developing Countries*. World Bank, Washington, DC.

McKenzie, D., and A. Paffhausen (2019): "Small Firms Death in Developing Countries," *Review of Economics and Statistics*, 101(4), 645–657.

McKenzie, D., and C. Woodruff (2014): "What are we learning from business training and entrepreneurship evaluations around the developing world?," *The World Bank Research Observer*, 29(1), 48–82.

——— (2017): "Business practices in small firms in developing countries," *Management Science*, 63(9), 2967–2981.

Syverson, C. (2011): "What determines productivity?," *Journal of Economic literature*, 49(2), 326–65.

A Discussion of no-harm assumption

This appendix discusses some additional factors that could potentially lead to a negative training impact and the reasons why they would not be an issue in the context of our study.

Limited applicability Business training programs often rely on models and practices developed in developed countries, which may not be directly applicable to the specific contexts of developing countries. Our training is developed by Kenyan company that relies on local content producers and has significant experience in the Kenyan market. As compared to other typical training programs in research studies, our program should suffer less from limited applicability.

Increased competition and market saturation It is certainly possible that our training could generate competitive advantage for entrepreneurs that adopt the training. Importantly, our proposition does not require business stealing to be zero, since we allow for negative effect for some individuals. We rule out overall industry contraction.

Resource constraints Developing countries often face significant resource constraints, including limited access to capital potentially discouraging further entrepreneurial efforts. In our case, the training was studied in the context of deploying a credit product, so relative impact of resource constraints should be lower than other studies.

B Supplementary materials

	Mean	Std. dev.
Training during holidays	0.32	0.47
Earlier credit relief	0.66	0.47
What is your age?	37.81	14.56
Received SMS business training	0.88	0.32
Business training is the most useful	.27	.45
SMS training impacted business practice	.78	.42
Monthly revenue	137,212.96	508,070.65
Written financial records	0.76	0.43
Understands interest	0.80	0.40
Credit balance checking frequency	3.61	1.50
Liquidate property to pay back loans	0.09	0.28
Daily revenue longer-run	10,998.92	3,1641.84
Daily profit longer-run	1,227.77	2,596.08
Monthly revenue longer-run	190,346.70	758,187.13
Monthly profit longer-run	18,361.76	33,486.72
Credit purchase first week of November, 2020	599.66	1855.76
N	1,434	

Table 1: Summary stats

When was the last time you checked Jaza Duka balance	No	Pct.
Earlier today	52	8.61%
Yesterday	99	16.39%
Earlier this week	227	37.58%
More than one week ago	116	19.21%
A month ago	40	6.62%
More than one month ago	37	6.13%
Never	21	3.48%
Total	604	

Table 2: When was the last time you checked Jaza Duka balance

	(1) Received any business training	(2) Business training is the most useful	(3) Received SMS business training	(4) SMS training impacted business practice
Training during holidays	0.0267**	0.117***	0.0136	0.0432*
	(0.0114)	(0.0284)	(0.0211)	(0.0223)
Gender FE	yes	yes	yes	yes
N	1433	1433	1432	1261

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Manipulation checks

	(1) Monthly revenue	(2) Written financial records	(3) Understands interest	(4) Credit balance checking frequency	(5) Liquidate property to pay back loans
Training during holidays	90034.3**	0.0699**	0.0748***	0.489***	-0.0836***
	(45143.0)	(0.0276)	(0.0258)	(0.141)	(0.0311)
Earlier credit relief	-2722.7	0.0558**	-0.0653**	-0.221	0.00689
	(44885.5)	(0.0272)	(0.0254)	(0.142)	(0.0321)
Gender FE	yes	yes	yes	yes	yes
N	1015	1432	1433	592	476

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Average treatment effects of deploying SMS training during holidays.

	(1) Daily revenue longer-run	(2) Daily profit longer-run	(3) Monthly revenue longer-run	(4) Monthly profit longer-run
Training during holidays	2006.3	356.0*	-45920.0	5677.8**
	(2388.7)	(204.1)	(58520.3)	(2667.8)
Earlier credit relief	-1395.1	-107.7	66457.2	-700.3
	(2363.7)	(201.7)	(57972.1)	(2649.1)
Gender FE	yes	yes	yes	yes
N	1064	1015	1042	988

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Average treatment effects of deploying SMS training during holidays, longer-run.

	(1) Monthly revenue	(2) Written financial records	(3) Understands interest	(4) Credit balance checking frequency	(5) Liquidate property to pay back loans
Training during holidays	90037.1*	0.0698***	0.0747***	0.489***	-0.0835***
	(46544.5)	(0.0264)	(0.0262)	(0.142)	(0.0289)
Gender FE	yes	yes	yes	yes	yes
N	684	940	941	401	324

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Effect only for the group that received credit relief in November; i.e., the comparison between "Credit November, Training November," and "Credit November, Training December."

	(1) Daily profit longer-run	(2) Monthly profit longer-run
College or University	439.7	2566.3
	[-142.9,1022.4]	[-5162.6,10295.3]
Training during holidays	447.1**	6242.7**
	[31.57,862.7]	[707.1,11778.2]
College or University × Training during holidays	-940.8***	-7534.8*
	[-1610.1,-271.5]	[-16403.8,1334.2]
Earlier credit relief	-104.8	-861.3
	[-520.4,310.7]	[-6424.1,4701.6]
College or University × Earlier credit relief	389.2	4026.8
	[-300.4,1078.9]	[-5093.5,13147.2]
N	1404	1369

95% confidence intervals in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Training effect by education level.