Comparing Minority Norm and Loss Framing in SMS Message Reminders for Pay-as-you-go Solar Device Repayment

Max Eagle, Kurt Eulau, Denny Lehman

2022-08-03

Contents

Abstract . . . 2
Introduction . . . 2
Pre-Experiment Context . . . 2
Theory & Concept Under Investigation . . . 4
Experimental Design . . . 5
Subjects . . . 5
Timeline . . . 7
Pre-Experiment Analysis . . . 8
Power Analysis . . . 8
Covariate Balance Check . . . 9
Spillover Check . . . 10
Experiment Administration . . . 11
Results . . . 11
Test for Differential Compliance . . . 11
Test for Differential Attrition . . . 12
Linear Models . . . 13
Conclusion . . . 19
Appendix . . . 20
References . . . 21

Abstract

Over 570 million people in Sub-Saharan Africa live disconnected from the electrical grid and either forgo electricity or purchase gasoline generators and kerosene lanterns. Burning fossil fuels worsens climate change while costing more in the long term compared to solar home systems[1]. Many people buy solar devices via a pay-as-you-go (PAYGo) monthly repayment program with retailers and receive regular SMS text messages payment reminders. Our retail partner [redacted] historically sent loss framed reminders to clients. Our experiment sent both the loss framed message as well as minority norm framed message to a random sample of customers (n = 874) to discover which message most effectively decreased the number of days elapsed between the expected payment date and actual payment date. Our experiment failed to reject the null hypothesis that the two messages caused different repayment behavior for customers for our primary outcome of days past due from the scheduled payment date. The results from the study suggest that PAYGo companies should look elsewhere to improve customer repayment.

Introduction

How do text message reminders that implement minority norm framing rather than the previously utilized loss framing affect the timing of repayments for pay-as-you-go solar customers who buy at-home solar products on credit? We implemented an experiment in partnership with [redacted] to answer this question. Our outcome metric to measure the treatment effect is the customers' number of days past due at time of payment, though we also examined collection rates among treatment and control groups. Collection rate is the ratio of the amount a customer pays during the experiment window compared to how much they should have paid according to the terms of their repayment contract.

Pre-Experiment Context

[redacted] sells products like the [redacted], pictured below, to customers who do not have a connection to the electrical grid or who are connected but are supplied with an unstable power supply.

Figure 1: [redacted]

[1]Tracking SDG 7: The Energy Progress Report (2021) by the World Bank

The products displace technologies like kerosene lanterns or gas generators, which are both more expensive in the long term and emit carbon dioxide. One of the few downsides of solar home systems is that the upfront cost can put the products out of reach for some customers. Consequently, PAYGo solar retailers like [redacted] offer products on credit whereby customers pay a down payment and then commit to make regular payments until they have paid off their loan. Devices are Global System for Mobile (GSM) enabled and have a “disable” feature, which automatically disables devices after a failure to pay on time. Solar PAYGo companies are interested in decreasing the time elapsed between the expected payments and actual payments made by their customers. [redacted]

Prior to the experiment, [redacted] sent out both upcoming and late payment reminders to customers in order to encourage on-time payments. [redacted] sent upcoming payment reminders 1 day before the expected payment date and late payment reminders 1, 7, 14, 21, and 28 days after the expected payment date in the event that customers still have not paid.

Days past due	Description	Text message sent
≤ (-2)	Regular service	None
(-1)	One day before due date	Reminder message
0	Due date	None
+1	One day past due	Late payment message
+7	Seven days past due	Late payment message
+14	Two weeks past due	Late payment message
+21	Three weeks past due	Late payment message
+28	Four weeks past due	Late payment message
+29	Company ceases text message reminders	None

Figure 2: Payment Reminder Schedule

Additionally, messages for the highest priced solar home system (SHS) were sent in English and all other messages were sent in [redacted]. [redacted]. This messaging regime resulted in 4 different messages sent to customers.

	Reminder message	Late message
[redacted]		
English	“Hello,”	“Hello,”

Figure 3: Payment Reminder Languages

Prior to the experiment, all messages were loss framed, meaning they highlight the fact that the device shuts off automatically after a payment default and leaves the customer without electricity. A sample of an upcoming SMS text message reminder in English is shown below:

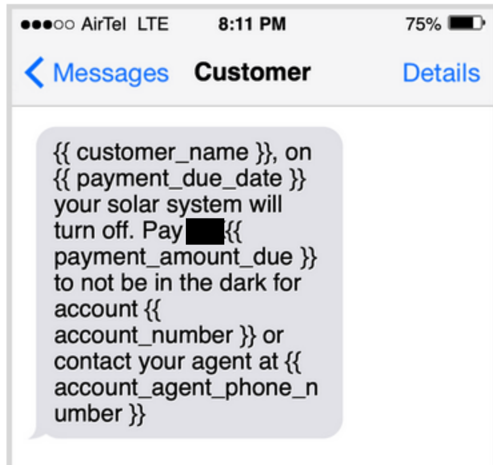


Figure 4: Sample Upcoming Payment Reminder

Theory & Concept Under Investigation

We believed that we could improve collections by changing the content of the text message reminders. In the absence of previous research in the PAYGo space to inform our experiment, we turn to microfinance studies that similarly looked to improve repayment through text messaging. Multiple experiments in microfinance in countries like Senegal or the Philippines demonstrated improvements only when messages make a social or cultural connection to the customer (Bursztyn et al. (2018), Behr and Jacob (2019), Karlan et al. (2016)). Behr and Jacob looked at ways to increase savings rates among microfinance savers in Senegal and found that including the name of a neighbor in text message reminders, implying that the neighbor had already made a deposit, increased the average amount saved. An experiment carried out by the UK tax authority, Her Majesty's Revenue and Custom, found that tax revenues increased by 10% when using minority normed letters to prompt tax payment (Luca and Bazerman 2020). While the outcome metrics in the aforementioned studies differ somewhat from PAYGo, the idea that emphasizing a person of interest's behavior in relation to the behavior of others has been shown to improve individuals' payment practices in a variety of contexts.

In that spirit, the treatment group will receive a text message, illustrated below, that uses minority norm framing to remind the customer that others have already made a payment.

[redacted]

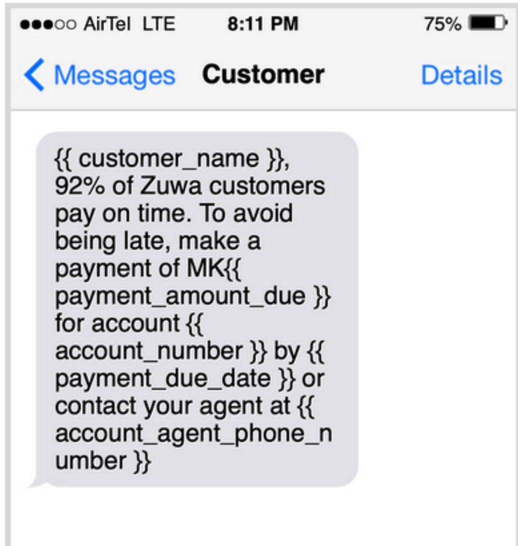


Figure 5: Sample Upcoming Payment Reminder for Treatment

Minority norm framing informs the customer that they will be in the minority of their peers by not making the company's preferred choice. Since being in this minority is often perceived as undesirable, the decision maker is nudged to make a different choice than they otherwise would have.

Before proceeding to review our experimental design, we do want to state that for the first 6 months of 2022, about 60% of customers pay on time while about 40% pay late. Minority norm messages in our study should therefore express that "60% of your peers pay on time" as the reminder message but we believe this to be unmotivating. In an effort to increase the intensity of the treatment, we instead sent reminder messages that informed customers that 92% pay on time up from the true 60% on time rate. Our goal here was to maximize the treatment effect, if present, by providing a stronger version of a minority norm message. We discussed other possible consequences of using a fabricated number, such as pressuring customers to make payments when they may be stretched thin financially and could use the money elsewhere. However, we arrived at the conclusion that customers who pay late in order to allocate spending to their needs would continue to do so. Additionally, [redacted] was supportive of using a more intense treatment in order to see if it had a positive benefit on their business operations. Leadership [redacted] also stated confidently that we would cause no harm to customers by using this higher dosage minority norm message.

Experimental Design

This experiment involved 874 [redacted] customers and took place between June 17, 2022 and July 24, 2022. Customers were blocked by various features, such as past repayment history, and then assigned to treatment and control groups. As customers approached a payment due date, they were sent reminder messages with the control or treatment wording. Customers who do not pay before the deadline enter a period of no service and are sent late payment messages with treatment or control wording. We capture the days past due as the outcome variable for our analysis to determine the treatment effect but also look at collection rate.

Subjects

We filter on the population of [redacted] customers whose payments are due within the month-long period of intervention. As per [redacted] instructions, we also constrained our target population to customers who average greater than 0 and less than 60 days late on repayment over the last 6 months. These constraints resulted in a sample of 874 accounts out of 2537 active accounts, with 438 randomly assigned to treatment and 436 assigned to control. Assignment to treatment and control relied on blocking on the average days past due. From [redacted] customer database, we leveraged covariate data on customer gender, age, and other variables for a covariate balance check and inclusion in various linear models.

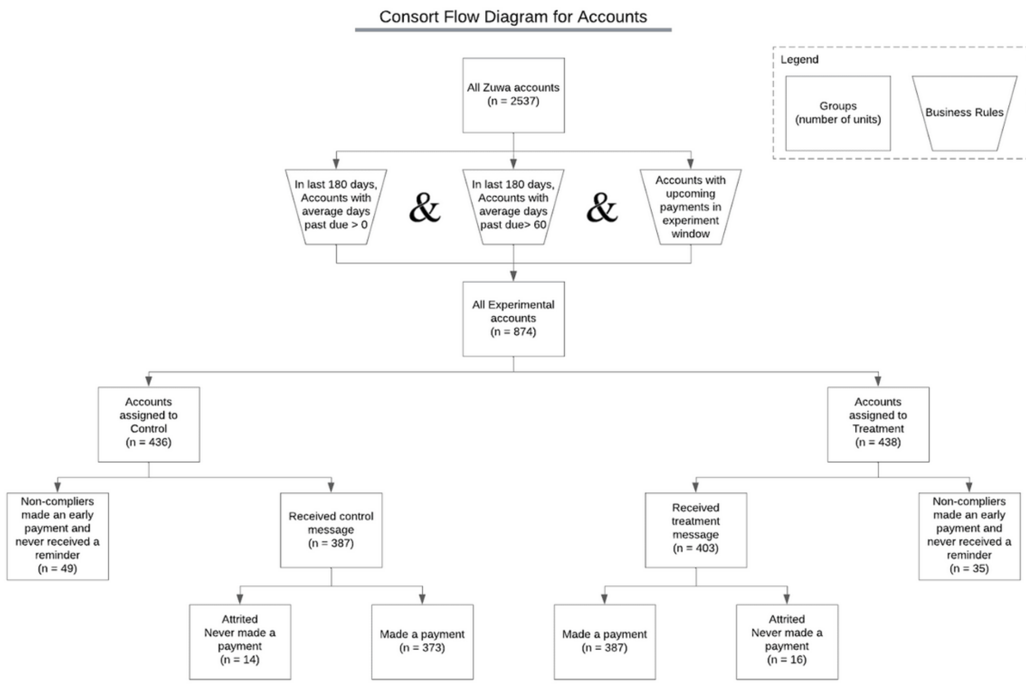


Figure 6: Consort Flow Diagram for Accounts

The nature of pay as you go solar customers varies and we identified six customer segmentations:

- **Early Payer**: An early payer pays well before their due date, meaning they do not receive a reminder message. These customers are assigned to treatment or control, but because they do not actually receive a message, they are treated as non-compilers and omitted from the experimental analysis. See section "Test for Differential Compliance" for more information.
- **Expected Payer**: An expected payer is assigned to treatment or control and receives the appropriate reminder message. They then pay before or at the due date and never enter a period of no service due to lateness. Their days past due value is never positive.
- **Late Payer**: Late payers receive a reminder message and pass their due date without making a payment. Their solar unit is remotely turned off, and they begin receiving the appropriate late messages from their assigned group. After some time, the late payer pays and re-enters normal service.
- **Frequent Payer**: A frequent payer pays a little bit at a time. Normally, they pay just enough to keep the lights on for an additional few days. Since the due date is the date when their account balance reaches zero, the frequent payer constantly resets their due date to a few days in the future. This behavior means that they receive many reminder messages in the experiment because of their multiple due dates.
- **Big Payer**: A big payer is assigned to treatment or control and then receives a reminder message as their due date approaches. This customer then makes a big payment, followed by a second payment. For example, if a customer pays for 90 days of light and the next day makes a second payment, the

system will register their days past due as -89 days. Big payers, therefore, can produce negative values.

- **Never Payer**: A never payer is assigned to treatment or control and receives their reminder message. They then miss payment on the due date and begin receiving late payment messages throughout the experiment. Because these customers did not make a payment by the end of the study, we excluded them from all linear models after testing for differential attrition. See section "Test for Differential Attrition" for more information.

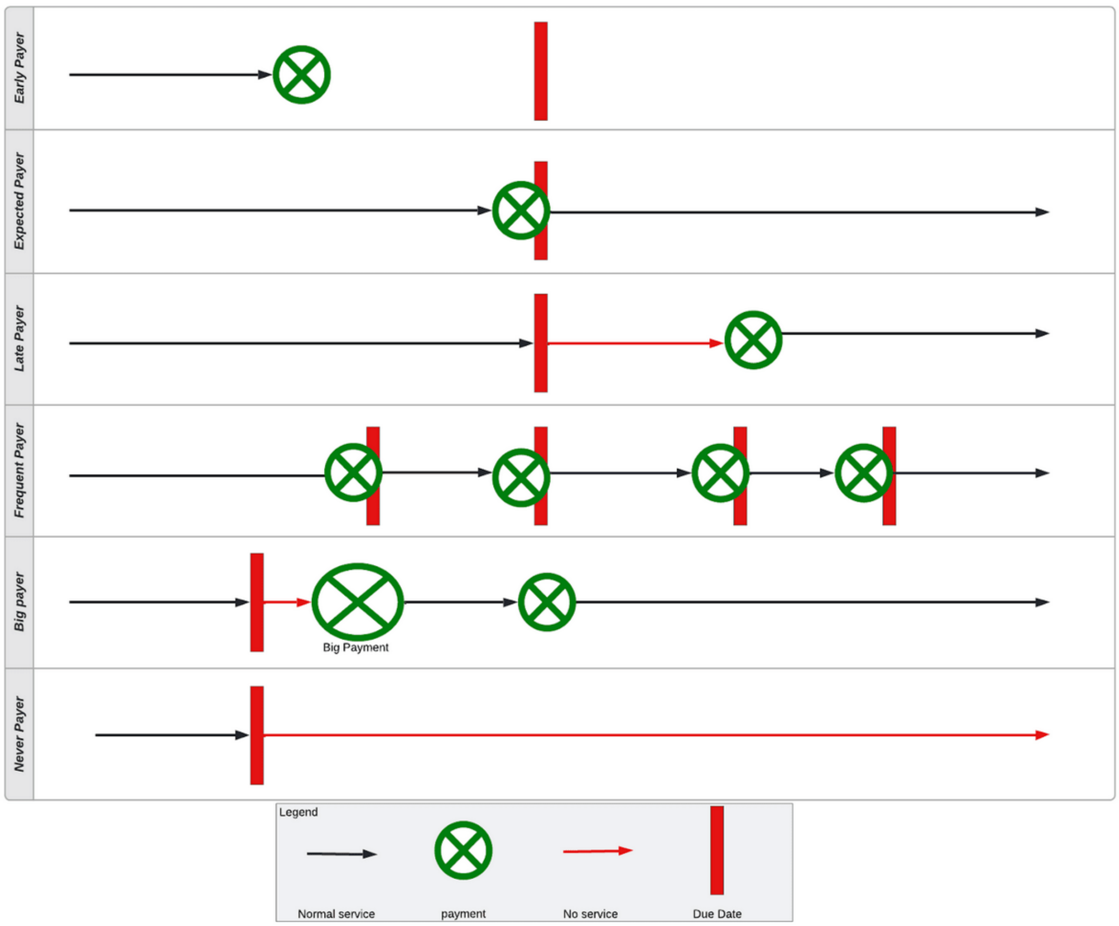


Figure 7: Payment Timeline Examples for Different Customer Segmentations

Timeline

This experiment took place between June 17, 2022 and July 24, 2022. Upcoming payment reminders were sent out from June 17 through July 17, and late payment reminders were sent out from July 18 through July 24. This extra week of late payment reminders allowed the experiment time to mature for customers who received upcoming payment reminders in the final week and days of the experiment.

With one control and one treatment group, we sent up to eight different messages per day, as shown by the table below:

Group	Reminder Type	Language	Group	Reminder Type	Language
Control	Upcoming	Chichewa	Treatment	Upcoming	Chichewa
		English			English
	Late	Chichewa		Late	Chichewa
		English			English

Figure 8: Possible Daily Message Combinations

Upcoming payment reminders are still sent 1 day before the expected payment date and late payment reminders are still sent 1, 7, 14, 21, and 28 days after the expected payment date.

The upcoming and late payment messages in English for both control and treatment are shown below:

Upcoming Control	{{ customer_name }}, on {{ payment_due_date }} your solar system will turn off. Pay [redacted]{{ payment_amount_due }} to not be in the dark for account {{ account_number }} or contact your agent at {{ account_agent_phone_number }}
Upcoming Treatment	{{ customer_name }}, 92% of [redacted] customers pay on time. To avoid being late, make a payment of [redacted]{{ payment_amount_due }} for account {{ account_number }} by {{ payment_due_date }} or contact your agent at {{ account_agent_phone_number }}
Late Control	{{ customer_name }}, your payment of [redacted]{{ payment_amount_due }} for account {{ account_number }} is overdue by {{ account_days_overdue }} days! Please pay at least [redacted]{{ payment_amount_due }}
Late Treatment	{{ customer_name }}, 92% of [redacted] customers paid on time. You are late by {{ account_days_overdue }} days! Please pay [redacted]{{ payment_amount_due }} as soon as possible to account {{ account_number }}

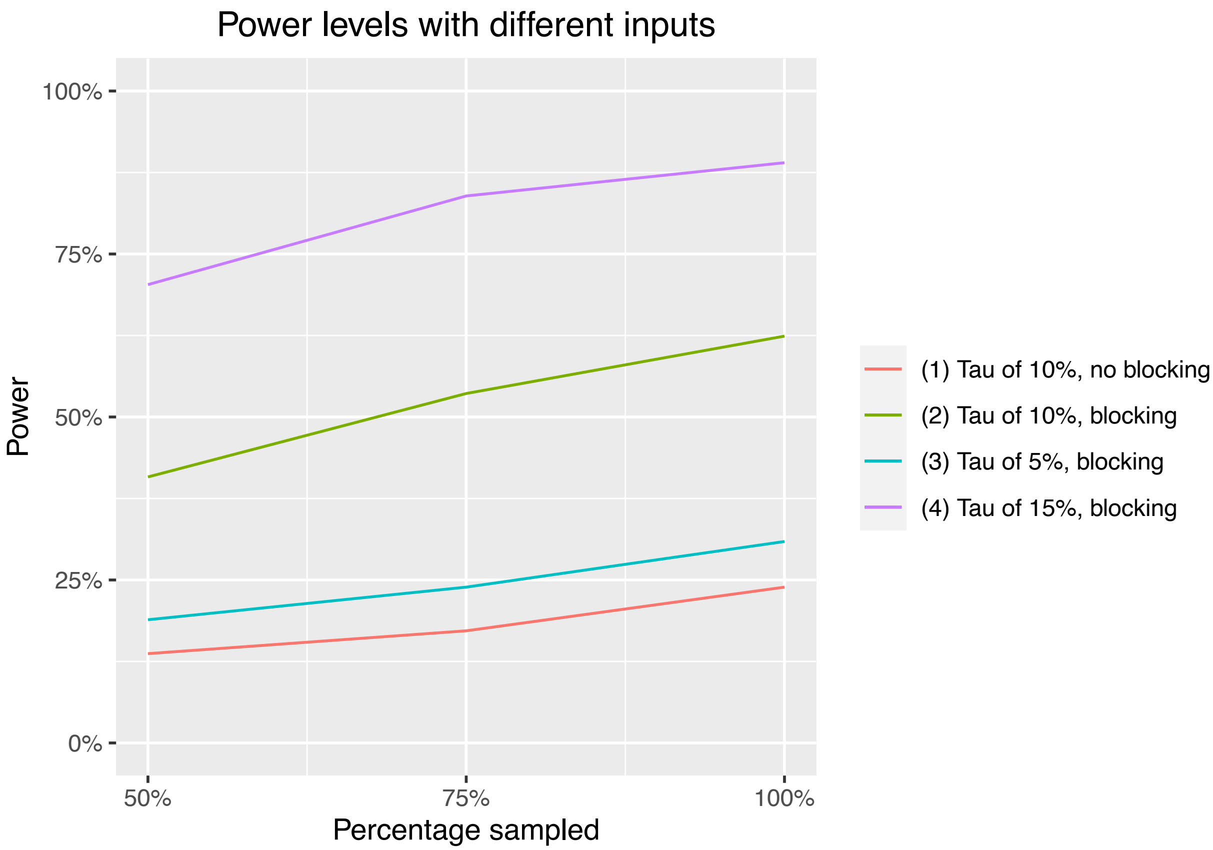
Figure 9: Control and Treatment Messages

[redacted]. We omit the four other messages for brevity. See Appendix for details.

Pre-Experiment Analysis

Power Analysis

Before conducting the experiment, we first calculated the statistical power of our experiment under different sample sizes, treatment effect sizes, and blocking conditions. While blocking with the sample size we expected, we found that a treatment effect of a 10% reduction in the number of days past due resulted in a power of about 62.4%. While this value is lower than the ideal benchmark of 80% or higher, this still gives us reasonable confidence that we can detect a real effect.



Covariate Balance Check

After randomly assigning accounts to treatment and control groups, we conducted a covariate balance check to identify any possible execution problems in the randomization process.

```
model_random <- accounts[ , lm(treat ~ 1)]

model_covariate <- accounts[ , lm(treat ~ unlock_price + customer_gender + customer_age +
                              missing_age + message_is_english)]

balance_check_rse <- waldtest(model_random, model_covariate, test = "F", vcov = vcovHC)
```

To check the integrity of our random assignment we first conducted an F-test to see if assignment to treatment and control groups can be better explained by covariates than by randomization. With an alpha level of 0.05 and a p-value of 0.977, we fail to reject the null hypothesis that the random model better explains assignment than the covariates model.

Note that this F-test used robust standard errors and thus could not include as many covariates due to challenges with generating an invertible matrix. Had we used classical standard errors, we could have included many more covariates.

```
model_random <- accounts[ , lm(treat ~ 1)]

model_covariate <- accounts[ , lm(treat ~ group_name + unlock_price + responsible_user +
                              customer_gender + customer_age + missing_age +
                              customer_region + customer_occupation +
```

```
                                        missing_occupation + message_is_english)]

balance_check_cse <- anova(model_random, model_covariate, test = "F")
```

With a p-value of 0.268 in this case, we reach the same conclusion that the random model better explains the assignment than the covariates model, suggesting that there are not any large imbalances in the treatment and control groups with respect to the available covariates.

Second, we include a table showing differences between treatment and control for a select set of covariates. For each covariate, we fail to reject the null hypothesis that there is a difference between treatment and control given that the standard deviation is larger than the difference, meaning that we have a t-value of less than one. These averages and lack of statistical significance provide further evidence that there are not any large imbalances in the covariates as a result of randomization.

Average Customer Characteristics

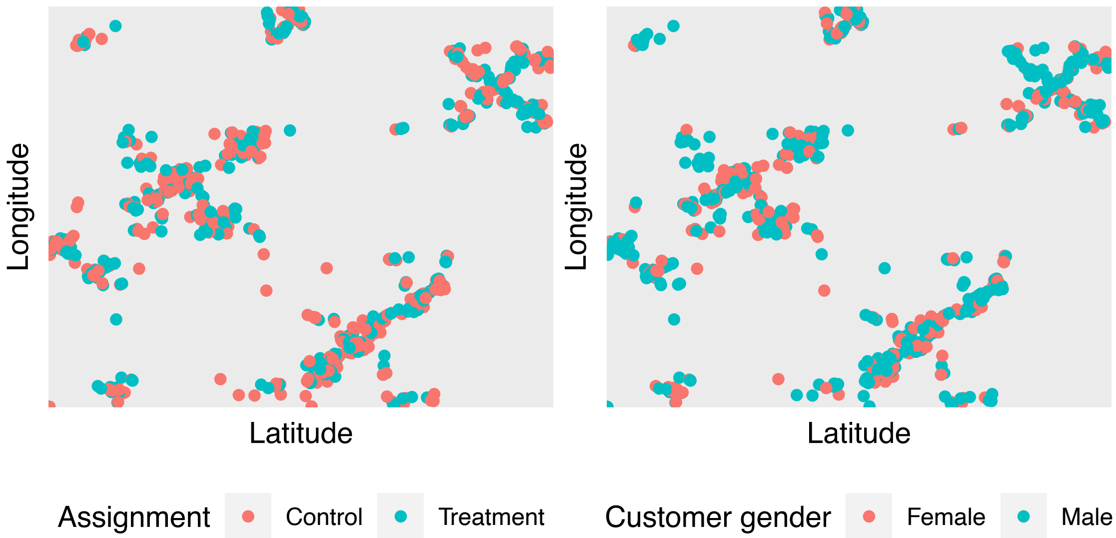
Characteristic	Control	Treatment	Difference
Unlock price	2.546×10^5	2.569×10^5	2329.017
			(9503.189)
Male	0.674	0.655	-0.019
			(0.032)
Female	0.326	0.345	0.019
			(0.032)
Age	40.356	40.253	-0.102
			(0.868)
English message	0.021	0.021	-9.426×10^{-5}
			(0.01)

Spillover Check

In order to check for possible spillover, we leverage latitude and longitude data to see if any accounts share the same location. [redacted] assured us that customers should not have multiple accounts, so that mitigated the risk of spillover in the event that a customer has one account assigned to treatment and another to control. However, there may be some possibility that different customers cohabitate such that those assigned to control may be exposed to treatment and vice versa.

We found that 34 accounts share a location that has at least one account assigned to treatment and at least one assigned to control. After discussing this with [redacted], we learned that these accounts may share location as a result of a clerical error (e.g, the sales rep entered the wrong registration location) or a measurement error (e.g., registration location has a margin of error of 40 to 450 meters depending on the device). Another common purported cause of duplicate geolocation information was the fact that sales agents sometimes activate devices at the retail store, causing some devices to all have a latitude and longitude associated with the [redacted] regional store rather than the customer's residence. Out of an abundance of caution, we assume that these accounts may be exposed to treatment and control and thus model for spillover with a dummy variable and interaction term in our analysis.

To visualize the data geospatially, we plot accounts by their latitude and longitude values and color them by assignment to treatment and control and select covariates. These plots illustrate that the accounts have a balanced geographic distribution.



Experiment Administration

In order to send the eight message varieties necessary for the experiment, we created a script to ingest and analyze daily payment information and then manually created bulk SMS message jobs in [redacted] customer resource management (CRM) system. Our script examined all [redacted] accounts that were scheduled for either an upcoming payment reminder or a late payment reminder. For those in the experiment, the appropriate treatment or control reminder was sent, and for those not in the experiment (e.g., accounts created recently or existing accounts that were overdue), the control reminder was sent. The script also determined the language of the message by examining the account's product and price plan. This all worked very well, but we ran into issues as the CRM's text messaging feature did not always deliver messages when message jobs were scheduled. For instance, about 27% of accounts received messages late enough that they no longer had a reminder one day before the payment date. To address this delayed message delivery issue, we include a binary variable to denote if a message was sent in a timely manner or not.

Another concern we encountered after the administration of the experiment was that several customers paid multiple times. The payment plan outlined in customer contracts is for regular payments every thirty days, much like a monthly installment in order to pay off the device. Many individuals, however, paid up to nine times during the course of the experiment. These subjects essentially paid the minimum amount allowable to receive light for a couple of days and then paid again when their lights were disabled via GSM. To model this behavior, we took the average days past due for each client across all payments they made.

Results

Before fitting linear models to analyze our treatment effect, we first conduct two tests to ensure that our approach is valid.

Test for Differential Compliance

One is examining noncompliance between accounts in treatment and control. Although we assigned 874 accounts to the experiment, there was no guarantee that these accounts would receive upcoming payment reminders. Our process involved identifying accounts with an upcoming payment during the month-long period of intervention. However, some of these accounts might have paid early, thus removing them from the population of accounts that would receive an upcoming payment reminder. Of the total accounts, 84 followed this behavior, meaning that they did not receive any message.

Typically, noncompliance is thought of in terms of Compliers, Never-Takers, Always-Takers, and Defiers. Our case of noncompliance is different in that accounts do not fall into these groupings but rather either "complied" by not paying early and receiving an upcoming payment reminder or "did not comply" by paying early and never receiving any message. We acknowledge that this problem may have been avoided by randomizing accounts among those set to receive an upcoming payment reminder on any given day, but we did not necessarily anticipate this problem in our design. Nonetheless, our randomization process offered benefits in terms of knowing the size of treatment and control groups before the experiment and allowing for an easier time blocking on baseline average days past due.

In order to test if the compliance or take-up rate differed between treatment and control, we regress a boolean variable encoding if the experiment was administered on a boolean for treatment and our blocking covariate. The null hypothesis is that the coefficient for treatment is zero, signifying that assignment to treatment is not predictive of whether the experiment was administered or not. We fail to reject the null hypothesis as there is sufficient evidence that the take-up rate is the same between groups. Given this result, we can estimate our treatment effect using the complier average causal effect (CACE) by filtering out non-compliers.

	Dependent variable:
	Experiment administered
Intercept	0.885***
	(0.017)
Treatment	0.032
	(0.020)
Avg. days past due: 8-29 days	0.040
	(0.024)
Avg. days past due: 30-59 days	-0.128
	(0.095)
Observations	874
R2	0.011
Adjusted R2	0.007
Residual Std. Error	0.294 (df = 870)
Note:	*p<0.05; **p<0.01; ***p<0.001

Test for Differential Attrition

Another test is examining attrition between accounts in treatment and control. Of the total accounts, 31 did not make a payment by the end of the experiment, meaning that we could not measure their outcome at this time. Some of these accounts might pay a few days after we stopped collecting data; some might pay weeks later. Consequently, imputing a value would be tricky given the high degree of uncertainty. We argue that not having these attrited accounts results in a possible underestimate for our experiment's treatment effect. Among the attrited accounts, the average days past due at experiment end was roughly 16 compared to a value of about 1 for all other accounts. We keep this in mind as we move forward with fitting linear models in the next section.

Nonetheless, we test for the presence of differential attrition to examine if there is non random attrition occurring that might result in a biased estimate. We do this by regressing a boolean variable encoding if the outcome variable was observed on a boolean for treatment and our blocking covariate. The null hypothesis is that the coefficient for treatment is zero, signifying that treatment is not predictive of attrition. We fail to reject the null hypothesis as there is sufficient evidence that the attrition rate is the same.

	Dependent variable:
	Outcome observed
Intercept	0.013
	(0.007)
Treatment	0.007
	(0.012)
Avg. days past due: 8-29 days	0.094***
	(0.028)
Avg. days past due: 30-59 days	0.165
	(0.086)
Observations	874
R2	0.050
Adjusted R2	0.047
Residual Std. Error	0.181 (df = 870)
Note:	*p<0.05; **p<0.01; ***p<0.001

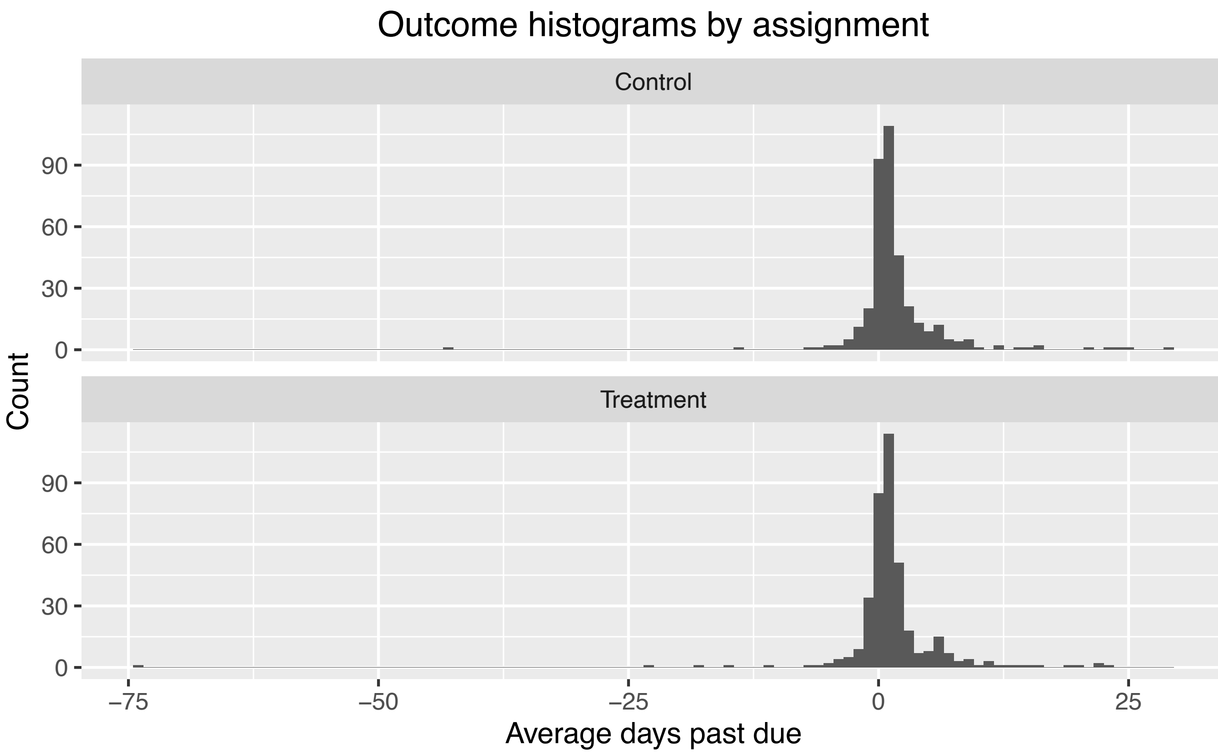
Linear Models

Because of the noncompliance identified in the experiment, we cannot generate an estimate for the average treatment effect (ATE). However, we can estimate the Complier Average Causal Effect (CACE) by filtering out non-compliers. Similar to the ATE, estimating the CACE requires three assumptions: random assignment, excludability, and non-interference.

How well does our experiment fulfill these assumptions?

1. Random Assignment: As discussed in the experimental design section, we randomly assigned accounts to treatment and control blocking on baseline average days past due, and we conducted covariate balance checks to identify any possible execution problems in the randomization process.
2. Excludability: Excludability refers to the assumption that receipt of treatment is the only causal agent. There may be a concern here as receiving a minority norm upcoming payment reminder is not the only means of changing days past due. Paying early, paying more than the required amount, or paying repeatedly resets the payment date into the future, so accounts in treatment and control with atypical repayment behavior can affect days past due via a backdoor channel. We test this by regressing average days past due on a boolean for treatment, our blocking covariate, and two additional right hand side terms that capture repeated payments and overpaying, respectively. Both of these terms have statistically significant coefficients, indicating that we appear to have violated excludability. While this is a concern, we plotted histograms for average days past due for control and treatment and found that the distributions are tightly centered around zero with relatively few extreme negative values. For control, there were 16 accounts with average days past due of less than -2, and for treatment, there were 20 accounts. This gives us some confidence that the extent of the problem is small and thus does not compromise the validity of the experiment's entire resulting dataset.

	Dependent variable:
	Avg. days past due
Intercept	4.710***
	(0.670)
Treatment	-0.217
	(0.356)
Avg. days past due: 8-29 days	2.480***
	(0.593)
Avg. days past due: 30-59 days	1.600
	(1.850)
Number of upcoming messages	-0.207*
	(0.097)
Collection rate	-2.680***
	(0.582)
Observations	760
R2	0.122
Adjusted R2	0.116
Residual Std. Error	4.840 (df = 754)
Note:	*p<0.05; **p<0.01; ***p<0.001



3. Non-interference: Non-interference refers to the assumption that the outcome of one subject is unaffected by the treatment assignment of another subject. As discussed in the pre-experiment analysis, there may be a slight concern of spillover for a small subset of accounts given reportedly shared device latitudes and longitudes, so we include a dummy variable and interaction term in one of our models to address this possibility analytically.

Although there are some concerns about two of the assumptions, primarily regarding excludability, we continue forward with fitting linear models and take utmost caution in interpreting any statistically significant results. Below we outline three major families of models:

1. Models with blocking covariates; blocking covariates and fixed effects; blocking covariates, fixed effects, and other relevant covariates
2. Models with days past due as the outcome variable; and collection rate as the outcome variable
3. Models with dummies and hetergenous treatement effects for spillover; late messages; and multiple messages

In our baseline model, we regress average days past due on a boolean for treatment and our blocking covariate. The null hypothesis is that the coefficient for treatment is zero, signifying that turning on treatment does not result in a difference in days past due. We fail to reject the null hypothesis as there is sufficient evidence that average days past due is the same between treatment and control. We reach the same conclusion after adding daily fixed effects, which refer to the day of the week of the first upcoming payment reminder, *and* other relevant covariates such as unlock price, customer gender, customer age, and whether the message was in English. These additional terms did not change the coefficient for treatment in a meaningful way nor did they increase the estimate's precision either.

	Dependent variable:		
	Avg. days past due		
	(1)	(2)	(3)
Intercept	1.200***	1.210***	1.410*
	(0.227)	(0.220)	(0.648)
Treatment	-0.305	-0.300	-0.307
	(0.367)	(0.368)	(0.372)
Avg. days past due: 8-29 days	2.790***	2.730***	2.630***
	(0.637)	(0.645)	(0.639)
Avg. days past due: 30-59 days	1.990	1.970	2.060
	(1.870)	(1.920)	(1.920)
Unlock price			-0.00000
			(0.00000)
Male			-0.049
			(0.424)
Age			0.003
			(0.014)
Missing age			2.620
			(1.650)
Message in English			0.832
			(1.040)
Daily fixed effects	No	Yes	Yes
Observations	760	760	760
R2	0.040	0.042	0.047
Adjusted R2	0.036	0.030	0.029
Residual Std. Error	5.050 (df = 756)	5.070 (df = 750)	5.070 (df = 745)
Note:			*p<0.05; **p<0.01; ***p<0.001

Another way to gauge customer repayment is collection rate. Collection rate measures how much individuals paid rather than when solar home system owners paid. When we re-ran the regression to measure the effect of minority norm versus loss framed text message reminders, we similarly fail to find evidence that would allow us reject the null hypothesis that the two message types cause a difference in customer repayment. Between these two models, there is evidence that the treatment of sending minority norm upcoming payment reminders did not seem to have an effect upon either days past due or the collection rate.

	Dependent variable:	
	Avg. days past due	Collection rate
	(1)	(2)
Intercept	1.200***	1.160***
	(0.227)	(0.025)
Treatment	-0.305	0.038
	(0.367)	(0.040)
Avg. days past due: 8-29 days	2.790***	-0.083
	(0.637)	(0.048)
Avg. days past due: 30-59 days	1.990	-0.128
	(1.870)	(0.091)
Observations	760	760
R2	0.040	0.005
Adjusted R2	0.036	0.001
Residual Std. Error (df = 756)	5.050	0.551
Note:	*p<0.05; **p<0.01; ***p<0.001	

We conclude by fitting several linear models with various heterogeneous treatment effect terms to account for some of the challenges we faced in the experiment:

- possible spillover for accounts that share a latitude and longitude value;
- accounts that received their first upcoming payment reminder message late;
- accounts that received any upcoming payment reminder message late;
- and accounts that received multiple upcoming payment payment reminder messages because they made small partial payments ahead of their initial payment date.

None of these models, however, result in meaningful changes to the treatment coefficient or increase the estimate's precision. We failed to detect a treatment effect in the baseline model and proceeded to control for a variety of other covariates and dummy variables to account for eccentricities in the data. Even with these other right hand side terms to help increase precision in our estimator, we still failed to detect a statistically significant treatment effect. Consequently, we have reasonable confidence that there is in fact no effect of sending minority norm text messages compared to loss framed text messages reminders.

One additional and notable observation is that the number of upcoming payment reminder messages is statistically significant. All else equal, accounts that received one additional message are associated with a decrease in average days past due by about one quarter of a day. While this may seem interesting, it is an artifact of how PAYGo repayment plans are structured. Accounts that pay multiple times often pay early by a partial amount to push out their payment date by a few days. This means that these accounts have negative days past due as a result of their behavior paying early. We caution against incorrectly interpreting this casually that sending more messages and/or paying multiple times causes accounts to have lower days past due.

	Dependent variable:				
	Avg. days past due				
	(1)	(2)	(3)	(4)	(5)
Intercept	1.200***	1.200***	1.160***	1.220***	1.510***
	(0.227)	(0.234)	(0.281)	(0.301)	(0.391)
Treatment	-0.305	-0.390	-0.515	-0.529	-0.414
	(0.367)	(0.377)	(0.439)	(0.464)	(0.605)
Possible spillover		-0.146			
		(0.689)			
First upcoming message sent late			0.145		
			(0.506)		
Any upcoming message sent late				-0.060	
				(0.459)	
Received multiple upcoming messages					0.325
					(0.527)
Number of upcoming messages received					-0.249*
					(0.098)
Avg. days past due: 8-29 days	2.790***	2.790***	2.770***	2.770***	2.720***
	(0.637)	(0.640)	(0.635)	(0.637)	(0.642)
Avg. days past due: 30-59 days	1.990	2.020	2.020	2.030	1.970
	(1.870)	(1.870)	(1.830)	(1.850)	(1.870)
Treatment x Possible spillover		2.540			
		(1.500)			
Treatment x First upcoming message sent late			0.795		
			(0.784)		
Treatment x Any upcoming message sent late				0.767	
				(0.735)	
Treatment x Received multiple upcoming messages					0.253
					(0.664)
Observations	760	760	760	760	760
R2	0.040	0.043	0.043	0.042	0.041
Adjusted R2	0.036	0.037	0.037	0.035	0.033
Note:	*p<0.05; **p<0.01; ***p<0.001				

Conclusion

We set out to determine if minority norm framed message reminders pushed customers towards lower days past due compared to the previously used loss framed messages. The experiment failed to show any differences in days past due between the treatment and control groups. The lack of effect held true even when switching the outcome variable from days past due to collection rate. While we analyzed many different models and found no effect, we do believe that more work can be done in this area.

One area of further investigation is exploiting the power of social connection. Past research, such as that conducted by Karlan et al. (2012), showed improvement in micro-finance repayment only when leveraging social connection in repayment reminders. Given the potentially impersonal nature of text messages, one avenue for future exploration could be to experiment with different forms of phone call reminders to customers by their assigned customer care officer.

Of the three major assumptions necessary for valid causal inference, we met random assignment, likely met spillover, and failed to meet excludability. We modeled for spillover out of an abundance of caution, though our concerns were alleviated by [redacted] assessment that customers were not actually co-located. We did find significant excludability concerns given the method of calculating our outcome variable, days past due. Rather than days past due being affected exclusively by the treatment message, customers could also have atypical payment behavior that affects the outcome variable through a backdoor channel.

Another concern is how some customers made several small payments and received multiple upcoming payment reminders, either in treatment or control. These additional messages not only increased the dosage of the treatment for some subjects but also lent greater weight to their data points in the experiment. We controlled for this phenomenon by averaging the days past due values for these accounts and summing the total collection rate. While this was one approach to deal with the issue, we still may have a biased estimate. In future research, we would address these concerns during experimental design, but due to time and business constraints, we were unable to perform this analysis in the initial stages.

Finally, we would recommend an industry shift in PAYGo accounting practices before performing additional studies in this population. The days past due metric contains inherent accounting flaws that allow for excludability in a causal study. Customers can be late to pay, make a relatively large payment, and end up with negative days past due despite not paying on time initially. This backdoor channel fundamentally does not allow for causal inference on the effect of different reminder messages to make customers pay on time. Unfortunately, this is the current standard for PAYGo business models, so the accounting of days past due would have to change to allow for meaningful experimentation.

To conclude, we found no effect of utilizing minority norm framed messages on days past due. However, we think that these results may be inconclusive in helping PAYGo companies improve customer repayment. We are optimistic about additional research that involves more creative approaches. In particular we recommend the following:

1. Select a different outcome variable such as collection rate
2. Prioritize certain customer subpopulations based on this new outcome variable
3. Consider experiments using alternative forms of communication, such as personal calls by customer care officers, because text messages appear inconsequential

Text message campaigns are not the only way to improve customer repayment and an industry shift in best practices could lead to a brighter future[redacted].

Appendix

Upcoming treat EN	{{ customer_name }}, 92% of Zuwa customers pay on time. To avoid being late, make a payment of MK{{ payment_amount_due }} for account {{ account_number }} by {{ payment_due_date }} or contact your agent at {{ account_agent_phone_number }}
Upcoming treat CH	{{ customer_name }}, makasitomala a Zuwa 92 mwa 100 alionse amalipira mu nthawi yake. Musakhale otsalira. Lipirani pa akaunti iyi {{ account_number }} ndalama yokwana MK{{ payment_amount_due }}. Pasanafike pa {{ payment_due_date }}. Mukafuna chithandizo imbani pa {{ account_agent_phone_number }}
Upcoming control EN	{{ customer_name }}, on {{ payment_due_date }} your solar system will turn off. Pay MK{{ payment_amount_due }} to not be in the dark for account {{ account_number }} or contact your agent at {{ account_agent_phone_number }}
Upcoming control CH	{{ customer_name }}, Magetsi anu a Zuwa pa akaunti {{ account_number }} azima pa {{ payment_due_date }}. Lipirani MK{{ payment_amount_due }} kuopa kukhala mum'dima. Mukafuna chithandizo imbani {{ account_agent_phone_number }}
Late treat EN	{{ customer_name }}, 92% of Zuwa customers paid on time. You are late by {{ account_days_overdue }} days! Please pay MK{{ payment_amount_due }} as soon as possible to account {{ account_number }}
Late treat CH	{{ customer_name }}, 92 mwa 100 yamakasitomala a Zuwa amalipira munthawi yake. Musakhale m'modzi mwa ochepa omwe amalipira mochedwa polipira MK{{ payment_amount_due }} ku akaunti {{ account_number }}. Pakadali pano mwatsalira ndi masiku {{ account_days_overdue }}
Late control EN	{{ customer_name }}, your payment of MK{{ payment_amount_due }} for account {{ account_number }} is overdue by {{ account_days_overdue }} days! Please pay at least MK{{ payment_amount_due }}
Late control CH	{{ customer_name }}, Lipirani magetsi a Zuwa pa akaunti {{ account_number }} kuopa kukhala mu Mdima ngat lero, mwakhala osalipira kwa masiku {{ account_days_overdue }}. Ndalama zoyenera kulipira mwezi uno ndi MK{{ payment_amount_due }}

Figure 10: All Control and Treatment Messages

References

Bursztyn, Leonardo and Fiorin, Stefano and Gottlieb, Daniel and Kanz, Martin, Moral Incentives: Experimental Evidence from Repayments of an Islamic Credit Card (September 23, 2015). World Bank Policy Research Working Paper No. 7420, Available at SSRN: https://ssrn.com/abstract=2664878

Dean Karlan, Melanie Morten, and Jonathan Zinman, A Personal Touch: Text Messaging for Loan Repayment (March 2012). National Bureau of Economic Research Working Paper No. 17952, Available at: https://www.nber.org/papers/w17952

Jorge Rodrigues Jacob and Patrick Behr (2019) ,"Social Capital and Financial Inclusion: Evidence From a Randomized Field Experiment", in NA - Advances in Consumer Research Volume 47, eds. Rajesh Bagchi, Lauren Block, and Leonard Lee, Duluth, MN : Association for Consumer Research, Pages: 658-659.

Luca, Michael, and Max H. Bazerman. The Power of Experiments: Decision Making in a Data-Driven World. Cambridge, Massachusetts ;: The MIT Press, 2020. Print.