Nudging Debtors to Pay their Debt:
A Randomized Control Trial

Pre-Analysis Plan for the Second Trial

Felix Holzmeister[†]   Jürgen Huber[†]   Michael Kirchler[†,‡]

[†] Department of Banking and Finance, University of Innsbruck
[‡] Centre of Finance, University of Gothenburg

Outline

We will conduct a randomized control trial with a partnered debt collection agency in Europe, with the goal to "nudge" debtors to pay their debts and/or to contact the agency to arrange a personalized payment plan. Arranging such a plan with a debt collection agency is the final step before the case is brought to an enforcement agency. Potential consequences would entail a "payment remark" with markedly lowered chances to get loans in the upcoming years and with downgraded conditions for various financial services. Hence, achieving a higher debt payment rate at this stage is highly important from an individual and a social perspective.

The randomized control trial (RCT) described in the current pre-registration is the second stage of a two-stage design. The first trial has been preregistered at the American Economic Association's registry for randomized controlled trials (AEARCTR-0005023).

In this second stage of the project, the debt collection agency will distribute about 41,000 hard-copy letters to debtors who currently owe money to the agency, calling on them to settle their debt (mainly consumption debt). While all clients received the same letter, but in different envelopes in the first step, clients will now receive letters with systematically varied contents in the same envelope in the second trial. In particular, we will add a factorial combination of social

and normative nudges (outlined in detail below) to the standard mailing which are hypothesized to increase response rates and repayments of clients' debts (plus an additional control treatment as outlined below). The envelope used in the second trial will be the most effective one in terms of the response rates and payments observed in the first trial.

The current pre-registration only refers to the second randomized control trial, i.e., only the variation of letter contents, while holding the envelope design constant. The trial will be conducted in February 2020. In particular, the debt collection agency will distribute approximately 41,000 hard-copy letters to debtors during three days (February 12 to 14, 2020), i.e., about 13,667 each day. The reason why the distribution of letters is spread out over three days is that the company needs to make sure that they will have sufficient resources to handle incoming phone calls and e-mails. The nine treatments will be randomly allocated to clients; likewise, the distribution of treatments across the three days on which the letters will be sent out will be randomized.

Interventions

We will implement nine different conditions/treatments based on factorial combinations of social and normative nudges (plus an additional control treatment). The letters in all treatments will be sent using the same envelope design, in particular the envelope that turned out to attract the highest response rate in the first trial (see Fig. 1; for further details, please refer to the pre-registration of the first trial at AEARCTR-0005023).[1]

In this randomized control trial, the information in the mailings sent to clients will be augmented by different nudges. The *Baseline* condition of the letter reads: "We want to help you to be free of your debts. You do currently have one or more cases with us with unpaid debts. On "My Page" you find a breakdown of your debts that today amounts to [amount]. You find "My Page" on [url]. Below you see information about how you can pay and contact us." In addition, the letter contains an overview about the different payment options and the note that "[if] you do not have the possibility to pay your debt, contact us so that we can review your case and find a solution that works for you."

One dimension of the factorial treatment design is to add the following *Social* nudge before the standard text of the letter (which holds true for the country where the RCT is conducted): "Approximately 1 out of 9 persons in [country] has just like you a debt to a debt collection company. Thus, you belong to a small part of the population that has a debt of this kind." The second dimension of the factorial design consists of three different *Normative* nudges using the following phrasing, added before the standard text of the letter (but below the social nudge, if applicable):[2]

---

[1] The envelope shows the logo of the debt collection agency in the top left corner, which has been canvased in Fig. 1 to preserve the anonymity of the partnered company. Likewise, text stating *"Important! Open me"* in the national language that would allow to infer the location of the partnered debt collection agency has been blurred in the bottom left corner of Fig. 1.

[2] Note that the enumeration of benefits/consequences will not be displayed inline as in the list below, but rather as bulleted (vertically arranged) list in the actual letters.

Figure 1: Envelope design for condition *"Important"*

- **#1 (P-I):** Pay your debt today and save a lot of money! It has the following benefits to you: (i) You avoid additional fees and interest; (ii) You avoid worsening your future economy; (iii) You avoid legal actions.

- **#2 (P-E):** Pay your debt and feel free! It has the following benefits to you: (i) You avoid additional fees and interest; (ii) You avoid worsening your future economy; (iii) You avoid legal actions.

- **#3 (N-I):** To not pay your debt is very expensive for you! It has the following consequences for you: (i) Additional fees and interest will be added; (ii) You risk worsening your future economy; (iii) You risk legal actions.

Note that the different versions of the normative nudges vary along two dimensions: (i) While #1 and #2 are framed positively (in terms of benefits resulting from settling one's debt), #3 provides the same information framed negatively (in terms of negative consequences). (ii) Likewise, while #1 and #3 provide objective information, #2 calls upon the recipients emotions ("feel free!"). By this means, the treatment design can be summarized as a 2 (social) × 2 (framing) × 2 factorial design (see Table 1. Yet, we do not include the factorial combination of a negatively framed statement calling upon the client's emotions since the corresponding phrasing of the information contravenes a principle of the partnered debt collection agency regarding communication with their clients.

		Normative Nudge				
		no	positive		negative	
			informing	emotional	informing	emotional
Social Nudge	no	Baseline	P-I	P-E	N-I	—
	yes	Social	Soc. / P-I	Soc. / P-E	Soc. / N-I	—

Table 1: Treatment Overview

The standard text which is part of all treatments (see above) is a simplified version of the company's previous standard letter. In particular, together with the company, we decided to shorten the text, emphasize the most crucial information, and arrange the contents in a more clear-cut manner. To examine the effects associated with these changes, which are applicable to all treatments summarized in Table 1, we add an additional *Control* treatment using the old standard letter as used by the company for correspondence purposes prior to the first of our two randomized control trials. Thus, in total, our trial will involve nine different treatment conditions.

Power Calculations

The purpose of this power analysis is to evaluate the sensitivity of the design of the randomized control trial (as described above) and the statistical test used to determine whether the interventions show significant effects. In particular, we calculate the minimum sample size required to have an experimental design sensitive enough to consistently detect reasonably small hypothetical effect sizes.

Although the main analysis (as outlined in detail below) will be based on logit regressions of the binary response indicators on treatment dummies and their interactions, the power calculations are based on pairwise z-tests of proportions (as were the power calculations for the first RCT). Based on the most successful intervention in the first randomized control trial (i.e., the envelope design that will be used in this trial), we expect a response rate of 4.7% for the *Baseline* condition (for details on how responses are classified, please refer to the "Data Analysis" section below). In order to reliably detect a small effect of Cohen's $h = 0.10$ (relative to the *Baseline* condition) with probability greater than 90%, we require a minimum sample size of approximately 3,370 for each condition, assuming a two-sided criterion for detection that allows for a maximum type-I error rate of $\alpha = 0.005$ (see Fig. 2). Cohen's h is calculated as the absolute difference between the response rate's arcsine transformations, i.e. $h = |\phi_1 - \phi_2|$ with $\phi_i = 2 \cdot \arcsin \sqrt{p_i} \, \forall \, i = \{1, 2\}$. Given the assumed response rate of 4.7% for the baseline condition, $h = 0.10$ corresponds to a relative effect of 50%, i.e., an increase (decrease) from 4.7% to 7.1% (2.4%).

Yet, the available sample allows us to considerably increase the sample size per condition. In total, we will send approximately 41,000 hard-copy letters, randomly assigned to one of nine conditions. Thus, the sample size per treatment will be approximately 4,550 per treatment such that
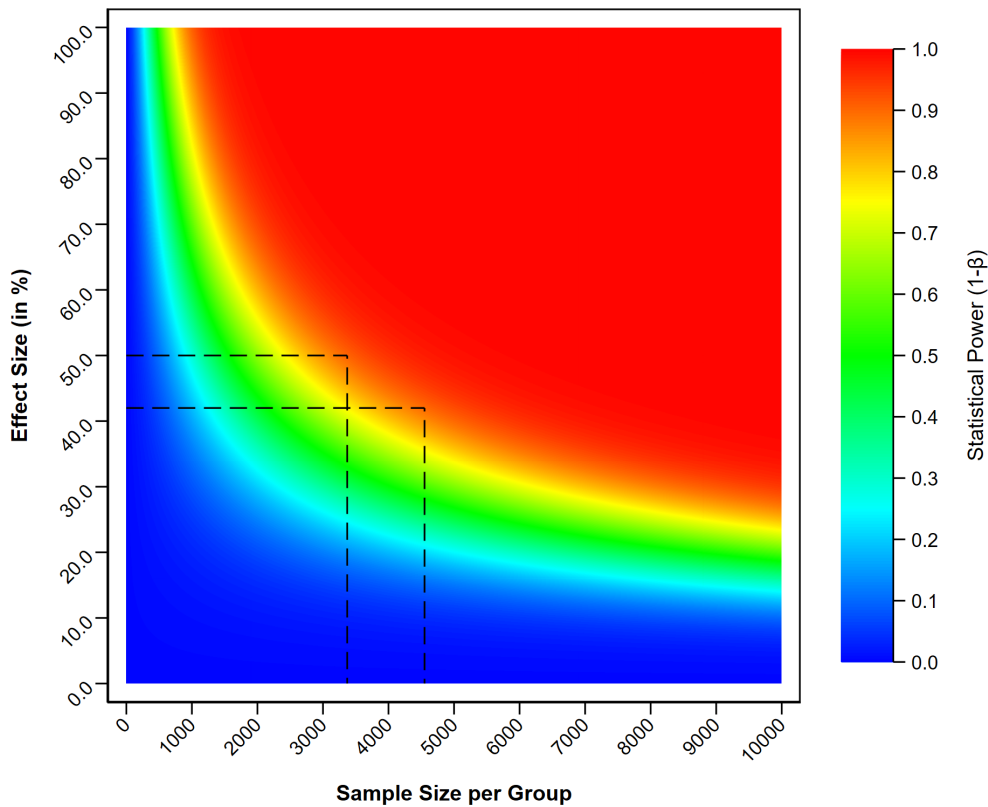
Figure 2: Power contour plot. The analysis assumes a response rate of 4.7% in the control condition and a two-sided significance criterion of $\alpha = 0.5\%$.

pairwise comparisons between the treatment conditions and the *Baseline* condition using two-sided z-tests of proportion will have 90% statistical power to detect effects as small as $h = 0.086$ (which, given the assumed base rate of 4.7%, corresponds to a relative effect of approximately 42%, i.e., an increase/decrease from 4.7% to 6.7% (2.7%)), conditional on a type-I error rate of $\alpha = 0.005$ (see Fig. 2).

Since the response rate can only be reasonably estimated for the *Baseline* condition (which is identical to the most successful intervention in the first-stage RCT), an a priori power analysis for comparisons between the treatment conditions is not possible (but only relative to the baseline condition). Yet, since the interventions (except for the additional control treatment) are assumed to induce an increase in response rates, the statistical power of proportion tests between the treatment conditions is expected to be even higher than for comparisons between the treatment and *Baseline* conditions. Likewise, as our main analysis using logit regressions implies a more parsimonious data analysis, the power calculations above are considered conservative benchmarks.

Data

The data as it will be provided by the debt collection agency can be categorized in four classes: (i) information about the debtors, (ii) information about the debt, (iii) the debtors' activity log, and (iv) the debtors' payment history. All data will be pseudonymous; only the debt collection agency will have the unique key to trace back the data provided to us to the non-anonymous data records of the company.

On the debtor level, we will receive information on the clients' birth year, their gender, and their yearly income (in terms of income classes). Debt-level information will detail the current size of the debt,[3] the data when it was acquired by the debt collection agency, as well as the original creditor, from which the debt collection agency acquired the liabilities. Payment history information will detail all payments (i.e., the amount and the date of payment) recorded on the debtor level.

To examine the effectiveness of our nudging interventions, the most crucial information is contained in the clients' activity log. In particular, the log file includes timestamps (exact to the second) for any type of communication which occurs between the debt collection agency and a particular client. The types of communication include "sent letter," "received letter," "outgoing call," "incoming call," "outgoing e-mail," "incoming e-mail," "sent short message (SMS)," "payment plan related activity," and "manual note." The data, thus, not only allows for identifying responses and payments, but also to control for a debtors' payment history, the overall size of the debt, and demographic information.

The data described above will be provided for all 41,000 debtors who will be contacted by the debt collection agency. In addition, the agency will provide us with a sample (including the same variables) of 4,500 additional debtors who will *not* be contacted in the randomized control trial.

Data Analyses

In a first step, we will identify debtors' responses to the mailing based on any of the communication means offered in the letter, i.e., phone calls, e-mails, and hard-copy letters, within four weeks starting from the day of the outgoing mail (i.e., until March 11, 12, or 13, depending on whether the particular letter was sent on February 12, 13, or 14). Based on the information about communication attempts initiated by the debtor, we will generate a binary indicator variable (response: yes/no) that will serve as the dependent variable.

As the main analysis, we will conduct a logistic regression of the response indicator on the binary treatment variables (*Social*, *positive/negative*, *informing/emotional*), all interaction terms, and a dummy variable indicating the *Control* condition. In a second model specification, we will add the following explanatory variables (in addition to the treatment indicators) to the model: (i) the size of the debt, (ii) an indicator variable capturing whether or not previous payments have been

---

[3] Depending on the distribution of the data, we may take the log of the total amount due rather than its level values.

recorded, (iii) the time how long the debt has been on the agencies book, (iv) the interaction term of (i) and (ii), and (v) the interaction term of (i) and (iii). In a third model, we will add (vi) gender, (vii) age, and (viii) dichotomous indicators for the (ordinal) income categories to the equation.

The second line of main analyses will focus on debtors' repayments of their debts. We will identify debtors' payments within four weeks starting from the day after the outgoing mail (i.e., March 12, 13, or 14, depending on whether the particular letter was sent on February 12, 13, or 14), as the letter will be received the day after it has been forwarded at the earliest. Based on the payment records, we will define three dependent variables: (i) a binary variable indicating whether or not at least one payment has been recorded within four weeks, and (ii) the relative amount repaid within four weeks, i.e., the sum of payments relative to the total amount of debt.

For the repayment indicator variable we will conduct the same analysis as for the response data described above. In particular, we will conduct a logistic regression of the repayment indicator on the binary treatment variables, all interaction terms, and a dummy variable indicating the *Control* condition. In a second model, we will again add the following covariates to the model: (i) the size of the debt, (ii) an indicator variable capturing whether or not payments have been recorded prior to the mailing, (iii) the time how long the debt has been on the agencies book, (iv) the interaction term of (i) and (ii), and (v) the interaction term of (i) and (iii). Again, in a third model, we will add (vi) gender, (vii) age, and (viii) dichotomous indicators for the (ordinal) income categories to the equation.

For the amount repaid during the four week period starting from the day of the sent letter (relative to the overall debt as recorded on the day of the outgoing mail), we will fit fractional response models (using a logit link) for the conditional mean. The log-likelihood function for the fractional model is of the form

$$\ln L = \sum_{j=1}^{n} y_i \ln \left\{ \frac{\exp(\mathbf{x}_j'\beta)}{1 + \exp(\mathbf{x}_j'\beta)} \right\} + (1 - y_i) \ln \left\{ 1 - \frac{\exp(\mathbf{x}_j'\beta)}{1 + \exp(\mathbf{x}_j'\beta)} \right\} \tag{1}$$

where n is the sample size and y_i is the dependent variable, $\mathbf{x}_j$ are the covariates for individual j, and Φ is the standard normal cumulative density function. $\ln L$ is maximized using the Newton-Raphson algorithm. Specifically, we will examine the impact of the binary treatment variables, all interaction terms, and a dummy variable indicating the *Control* condition. In addition we will again estimate models adding the same set of covariates as described above, i.e., (i) the size of the debt, (ii) an indicator variable capturing whether or not payments have been recorded prior to the mailing, (iii) the time how long the debt has been on the agencies book, (iv) the interaction term of (i) and (ii), (v) the interaction term of (i) and (iii), (vi) gender, (vii) age, and (viii) dichotomous indicators for the (ordinal) income categories to the equation.

Since debtors are contacted by the debt collection agency on a regular basis, some debtors may repay (parts of) their debt or contact the agency even without being reminded of their debt. In order to identify the effect of reminding debtors via hard-copy letters, we will compare the response and payment data of the additional sample of 4,500 debtors who will not be contacted with those of the *Baseline* and *Control* condition using z-tests of proportions.

In addition, we will conduct pairwise z-tests for proportions, comparing the treatment interventions (including the *Control*) to the *Baseline* condition, since these tests form the basis for the power calculations. For the sake of completeness, we will also report pairwise proportion tests comparing all treatment interventions with one another. The z-tests for proportions will be conducted as outlined in detail in the pre-registration of the first RCT (please refer to AEARCTR-0005023 for details).

For all analyses, we will assume a $\alpha = 0.005$ significance threshold; results with $0.005 < p < 0.050$ will be considered suggestive evidence. For all regression analyses (i.e., both linear and non-linear models), we will use robust standard errors based on the Huber/White sandwich estimator.