Nudging Businesses to Pay Their Taxes: Does Timing Matter?

Christian Gillitzer

School of Economics, University of Sydney

Mathias Sinning

ANU Crawford School of Public Policy, PKU NSD/CCER, RWI, IZA

19 November 2019

Abstract. This paper provides empirical evidence on the implications of the timing of reminders by studying the effect of varying the timing of reminder letters to taxpayers on their payment behavior. The collection of unpaid tax debts constitutes a considerable challenge for tax authorities. We discuss potential mechanisms through which reminders may affect taxpayers' behavior and study the payment behavior of business taxpayers in a field experiment in Australia. We find that a simple reminder letter increases the probability of payment by about 25 percentage points relative to a control group that does not receive a letter from the tax authority. However, variation over a three-week period in the timing of the reminder letter has no effect on the probability of payment within seven weeks of the due date. Our findings indicate that sending reminders early results in faster payment of debts with no effect on the ultimate probability of payment.

JEL-Classification: C93, H25, H26

Keywords: Tax Compliance, Business Taxation, Natural Field Experiment, Behavioral Insights

1   Introduction

Timing is an important aspect of policy design that is often overlooked by policymakers. Prompting people at different times can have drastically different levels of success (BIT, 2014). In particular, policies that rely on people to take action do not only depend on economic or time constraints but also on behavioral factors affecting task completion (Mullainathan and Shafir, 2014). Reminders have been linked to improved individual behavior in various contexts, including savings (Karlan et al., 2016) and loan repayments (Cadena and Schoar, 2011; Karlan et al., 2015). Less is known about the extent to which the timing of reminders affects policy outcomes, although the effectiveness of early and late reminders may differ considerably.

The aim of this paper is to provide evidence on the implications of the timing of reminders by studying how varying the timing of reminder letters to taxpayers affect their tax payment behavior. There is now a large body of experimental research investigating how deterrence and non-deterrence methods affect taxpayers' income reports, but there is comparatively little research on the timely payment of tax debts.[1] We present results from a field experiment that was conducted in collaboration with the Australian Taxation Office (ATO) to learn about the payment behavior of small business taxpayers. Such an analysis is important because the volume of outstanding tax debts is large. For the US, Internal Revenue Service (IRS) tax gap estimates for the 2008-2010 tax years show average underpayment of USD39bn, which is about 10 percent of the gross tax gap (IRS, 2016). The bulk of unpaid debt is owed by individual taxpayers and unincorporated businesses. In Australia, 30 percent of small businesses did not pay their tax liabilities on time in the 2017 financial year and together owed around 67 percent of total collectible tax debt (ATO, 2017a).

We provide an overview of the tax compliance literature and the literature on reminders. While our main contribution is empirical, we deduce potential mechanisms through which reminders may affect tax compliance. The main mechanisms discussed in the literature are payment constraints, imperfect memory and procrastination. From a

---

[1]See Slemrod (2018), Mascagni (2018) and Hallsworth (2014) for a comprehensive review of the literature using RCTs and lab experiments to study tax compliance.

theoretical perspective, the effect of the timing of reminders on tax payments is ambiguous because early reminders may be more or less effective than late reminders (Ericson, 2017). Early reminders are less likely to be useful because some people may not pay their debt initially (e.g. because of payment constraints) and then forget about it over time. At the same time, late reminders may be less effective because they leave less time to pay the debt. While our trial design does not allow us to draw inferences about procrastination, we are able to link our findings to payment constraints and imperfect memory.

We contribute to the literature in several ways. Firstly, our field experiment is novel because it varies the time elapsed between the tax payment deadline and the reminder letter being sent. Deciding when to send reminders is an issue facing all debt collectors. Despite its importance, we provide the first scientific evidence on this choice in a large-scale real-world setting. Secondly, we explain how our findings are linked to imperfect memory and payment constraints. Empirical evidence on the mechanisms through which reminders affect individual behavior is scarce. Thirdly, our conclusions regarding the optimal timing of reminders are based on results that exhibit a high degree of both internal and external validity. Our trial studies the population of business taxpayers in Australia with a history of compliant payment behavior who have missed their tax deadline. Debt cases were randomly allocated to receive a reminder letter either one, two, or three weeks following the deadline; a control group did not receive a letter for the seven-week duration of the trial.

The results from the field experiment reveal a large effect of a reminder letter on payment behavior. The probability of an overdue debt being paid by the end of the seven-week trial was approximately 25 percentage points greater for cases receiving a reminder letter relative to cases in the control group, which did not receive a reminder letter. However, irrespective of the date at which reminder letters were received, overdue debts were paid at the same rate in each of the treatment groups. Earlier receipt of a reminder letter resulted in debts being paid sooner, but the share of debts paid in each group that received a reminder letter was the same at the end of the seven-week trial period. Accordingly, we find no benefit from delaying the date at which the reminder letter is sent. The only meaningful source of heterogeneity in tax debt payments is related to

2

the amount of outstanding debt. Receipt of a reminder letter increased the probability of payment for debts up to about AUS\$7,500, but had no discernible effect for debts above AUS\$7,500.[2] Even though reminder letters did not increase the probability of payment for high value debts, the probability of payment was already substantially greater for large than small debt cases. Our findings suggest that tax debt payments were affected by imperfect memory and payment constraints.

The remainder of this paper is structured as follows. Section 2 provides an overview of the literature and discusses mechanisms through which reminders may affect tax compliance. Section 3 describes the trial setting. Section 4 discusses trial design, empirical strategy and data. The results are discussed in Section 5. Section 6 concludes.

2   Related literature

Based on the seminal work of Thaler and Sunstein (2008), a recent strand of the behavioral economics literature focuses on the design of cost-effective interventions that "nudge" people to make better decisions without limiting their choices to improve their health, wealth and happiness. Nudging aims to target the psychological roots of human behavior to address cognitive biases. Important examples include present bias (O'Donoghue and Rabin, 1999), loss aversion (Kahneman and Tversky, 1979; Gächter et al., 2009), framing (Kahneman and Tversky, 1979; Levin et al., 1998), anchoring (Tversky and Kahneman, 1974; Wansink et al., 1998; Scott and Lizieri, 2012) and social norms (Ariely, 2008; Cialdini, 2008; Allcott, 2011). The work in this field has made extensive use of randomized controlled trials (RCTs) to detect cognitive biases and to nudge people to make better decisions.

Our paper adds to a growing literature that employs four design principles proposed by the UK's Behavioural Insights Team based on their own work and the literature: "(...) make it Easy, Attractive, Social and Timely (EAST)" (BIT, 2014, p. 4). We focus on studying the role of making nudges timely. Moreover, by exploring how varying the timing of reminders to taxpayers affects their tax payment behavior, we contribute directly to

---

[2]AUS\$=Australian dollar. One Australian dollar is equal to 0.69 US dollars (10 September 2019).

the literature on tax compliance.

The following subsections provide an overview of the tax compliance literature and place our experiment within the theoretical and empirical literature on reminders. A final subsection discusses the expected mechanisms through which reminders may affect tax compliance.

2.1 Tax compliance literature

Much of the empirical literature on tax compliance has studied the accuracy of income reports. For individual taxpayers, Slemrod et al. (2001) and Kleven et al. (2011) find that audit threat letters increase reported income. For business taxpayers, Hasseldine et al. (2007), Ortega and Sanguinetti (2013) and Harju et al. (2014) also find evidence that an increased threat of audit increases reported income. In Australia, Wenzel (2006) finds that reminder letters sent to business taxpayers who had not filed their tax declaration on-time increased compliance, with larger effects on compliance evident if the reminder letter included text assuring taxpayers they were not suspected of being dishonest and expressed understanding.

Our research differs from much of the tax compliance literature in studying the payment of a tax liability, rather than an initial declaration of income. Notable exceptions are Del Carpio (2013) and Castro and Scartascini (2015), who study the effect of audit threat letters and appeals to social norms on property tax payments in Peru and Argentina, respectively; Dwenger et al. (2016), who investigate the effect of enforcement, social norm and letter simplification treatments on payments of a church tax in Germany; Pomeranz (2015), who shows that firms in Chile increased VAT payments following receipt of an audit threat letter; and Brockmeyer et al. (2016), who found that an enforcement email treatment tripled the income tax filing rate and doubled the payment rate among previously non-filing firms in Costa Rica. Other studies have investigated the importance of the delivery mechanism. Ortega and Scartascini (2015) find personal visits by tax inspectors to be more effective than letter or email contact in eliciting payment of overdue taxes in Costa Rica, while Boning et al. (2018) find that firms in the US receiving an in-person visit from a tax official remit more than firms receiving a letter.

Our experimental trial is most closely related to Hallsworth et al. (2017), who study the effect of reminder letters on payments of overdue income tax for a large sample of UK taxpayers. They find small but significant effects on payment rates when reminder letters include appeals to social norms. There are three important differences between our study and Hallsworth et al. (2017). First, we study business rather than individual income taxpayers. Second, we investigate the effect on payment rates using variation from the timing at which the reminder letter is sent. Third, our study includes a control group that did not receive a reminder letter for the duration of the trial.[3]

2.2 Theoretical literature on reminders

Payment of tax debts depends on capacity to pay and behavioral factors affecting task completion. The most important behavioral factors are imperfect memory and present biased preferences. Crucially, payment constraints and behavioral factors may interact to affect the optimal timing of a reminder.

Sunstein (2014, p. 587) argues that "For reminders, timing greatly matters; making sure that people can act immediately on the information is critical (especially in light of the occasional tendency to forgetfulness)." For payments, such as tax debts, this means timing reminders when people or businesses have the capacity to pay. On average, businesses in Australia wait 47 days to receive payment of invoices (OECD, 2017) and experience an average late payment time of 15 days (Dun and Bradstreet, 2017); these delays create uncertainty over cashflows and affect firms' ability to pay tax debts. More than half of small business in Australia report having experienced cash flow issues (Roe and Law, 2017).

Holman and Zaidi (2010) and Taubinsky (2014) show that when people are overconfident about their ability to remember an uncompleted task then longer deadlines can reduce the probability of task completion. However, reminders can reverse the effect of longer deadlines (Taubinsky, 2014).

---

[3]While variation in the date at which reminder letters were sent in the first week of the trial enables Hallsworth et al. (2017) to estimate the effect of a generic reminder letter on payments at the end of the first week, this estimate provides only a lower bound because additional payments are likely to be made after the first week.

Payment constraints and imperfect memory may interact to make an early reminder less effective than a late reminder. An early reminder will be useful for those who did not pay a debt because they forgot but will not be useful for those unable to pay. In contrast, a delayed reminder will remain useful for those who forgot but will be additionally useful for those that forgot by the time they had the capacity to pay (Ericson, 2017). However, there may be a point beyond which delay reduces the effectiveness of the reminder. The longer a debt remains unpaid the more likely it is that the debt is forgotten, new demands on a business's cash flow arise, and payment constraints escalate. Consistent with this, Ortega and Scartascini (2015) find evidence that enforcement letters are more effective at prompting full payment for recent than aged overdue tax debts in a field trial in Colombia.

Present bias causes people to overweight the present relative to the future (Laibson, 1997; O'Donoghue and Rabin, 1999). Ericson (2017) shows theoretically how imperfect memory and present bias can interact to affect the efficacy of reminders. The optimal time to send a reminder to a present biased individual can be later than for a time-consistent individual. This is because present biased individuals may procrastinate in response to early reminders and forget about the task again before acting.

Finally, the effect of reminders can depend on whether they were anticipated or unanticipated (Ericson, 2017). Anticipated reminders can reduce the probability of task completion because better memory reduces the cost of procrastination. It appears likely that the reminders in our trial were unanticipated. We discuss this issue in Section 3.

2.3   Empirical literature on reminders

Empirically, reminders sent on or prior to deadlines have been found to be effective. Karlan et al. (2015) find that reminders increase the probability of microloan repayment in the Philippines, but only when the loan officer's name is mentioned. Borrowers in the treatment group received a reminder letter before each weekly repayment deadline over the life of the loan. Reminders sent on the weekly payment deadline, one or two days before are equally effective. However, the weekly reminders are most likely to have been anticipated by borrowers.

Cadena and Schoar (2011) find a simple reminder message to be as effective as a 25 percent reduction in the cost of capital in ensuring on time microloan repayment in Uganda but they do not investigate the issue of timing. Karlan et al. (2016) find that reminder messages increase saving for people with a commitment savings account and hypothesize that reminders mitigate undersaving caused by inattention to large irregular expenses. In contrast to the other papers, Karlan et al. (2016) also test the effect of a reminder sent after a missed deadline but find no effect of sending an additional late reminder to people who have missed their commitment saving goal.

Reminders have been found to be effective in a variety of other domains as well, including goal achievement in the workplace (Cadena et al., 2011) and adherence to medical treatment (Bobrow, 2016). However, there is also evidence that reminders can have unintended effects, with Damgaard and Gravert (2018) finding that reminders increase unsubscriptions from a charity mailing list.

Our study is novel in studying the effect of reminder timing on tax debt payment in a real-world setting. Although there is now a body of theoretical literature, there has been little empirical work investigating the timing of reminders, with the exceptions of Karlan et al. (2016) and Karlan et al. (2015) for commitment saving accounts and microloan repayment. The timing of reminders is a decision facing all tax authorities, and debt collectors more generally.

2.4 Mechanisms and expected behaviors

The literature discussed above identifies payment constraints, imperfect memory and procrastination as relevant mechanisms through which reminders may affect tax compliance. The observed behavior of taxpayers may also be a result of the interaction between these mechanisms. Our experimental setting permits us to identify imperfect memory and payment constraints but not procrastination.

Table 1 summarizes the behavior we expect to observe for each combination of frictions. For businesses that forgot their debt, but do not face a payment constraint, we expect the reminder to cause an increase in payment rates occurring shortly after receipt of the reminder (Case 1). For businesses that forgot their debt and face a payment constraint, we

expect the reminder to cause entry into a partial payment plan, which we describe further in Section 3 (Case 2). Evidence that late reminders are more effective than early reminders would be consistent with payment constraints that decrease with time since the due date and taxpayers quickly forgetting reminders (Case 2). For example, businesses receiving an early reminder may forget to make a payment before the payment constraint eases. Evidence that early reminders are more effective than late reminders would be consistent with payment constraints that increase with time since the due date. If early and late reminders are equally effective, this would be consistent with an absence of payment constraints or payment constraints that do not vary with time since the due date. We expect the reminder to have no effect on taxpayers with perfect memory (Cases 3-4).

[Table 1 about here.]

We compare our findings to these predictions in the results section (Section 5), where we also discuss the implications of our results and the role of potential mechanisms for the optimal timing of reminders.

3 Trial setting

Our experiment studies repayment of Business Activity Statement (BAS) debts. All businesses in Australia registered for the Goods and Services Tax (GST – equivalent to the Value-Added Tax) are required to submit a BAS. The BAS is used to report GST tax collections (the reason for most businesses to lodge a BAS), pay as you go withholding for businesses with employees, pay as you go installments for making regular payments towards the expected end of year income tax liability, wine equalization tax and luxury car tax. The BAS also allows businesses to claim fuel tax credits.

About 93 percent of businesses in our sample lodged a BAS on a quarterly basis, the remaining 7 percent lodged on a monthly basis. About 25 percent of the businesses in our sample lodged their BAS form online. A tax or BAS agent lodged the BAS form on behalf of the business owner in 39 percent of the cases.[4] The BAS can also be lodged through

---

[4]We estimate this fraction by summing over reporting channels used by tax or BAS agents. We obtain a lower bound estimate because agents could have helped prepare BAS forms that were lodged by businesses.

other channels such as mail or telephone (about 4 percent of cases in our sample). In addition to these lodgment channels, the ATO is able to raise the debt automatically. This was done for about 32 percent of the cases in our sample. Raising the debt automatically is akin to a notice of assessment for an estimated tax debt (i.e. pre-payment).

Payment is due at the time of lodgment. However, lodgment and payment are separate processes that do not necessarily occur at the same time and that may involve different actors or decision-makers. Although about 94 percent of businesses in our sample have a tax agent and at least 39 percent of the BAS forms were lodged by an agent, this does not imply that agents were also involved in the payment process. Agents may lodge on behalf of a business, but payment is the responsibility of the business. Businesses are encouraged to lodge on time even if they are unable to pay on time (ATO, 2017b). Lodgment and payment are also separate processes for about one-third of the cases in which the tax office raised the debt automatically.

All businesses in the experiment missed the 26 May 2017 BAS payment due date. The ATO has an escalating series of treatments to collect overdue debts. Taxpayers deemed to be of low risk are initially sent a blue reminder letter. If payment is not forthcoming the ATO will typically send an orange letter warning of stronger action followed by a final warning red letter before commencement of stronger action (ATO, 2017c). Stronger action can include issuance of garnishee notices, director penalty notices and insolvency proceedings; some debts are referred to external collection agencies (ATO, 2019a).[5] For the financial year 2016-17, 70 percent of small business tax liabilities were paid on time and 87 percent were paid within 90 days of the due date (ATO, 2017a). Garnishee notices were used in fewer than 1.1 percent of all debt cases in the financial year 2017-18 (ATO, 2019b). Waiting until a warning of stronger action to comply can increase scrutiny of a business's tax affairs and raise the ATO's risk assessment of the business for future interactions.

We study how the effectiveness of the initial blue reminder letter sent to low-risk taxpayers varies depending on when the letter is sent. Unpaid debt cases were randomly allocated to receive a reminder letter either 10, 17 or 25 days following the due date. A

---

[5]Director penalty notices make company directors personally liable for tax debts.

control group did not receive a letter for the duration of the trial. The ATO did not initiate any other contact with businesses for the duration of the trial. The trial lasted until 14 July 2017 (49 days). This is the length of time for which the ATO is willing to use only a simple reminder letter to induce payment. Following the end of the trial unpaid debt cases were treated according to the ATO's standard debt management practices.

The taxpayers in our experiment were classified as low risk according to internal ATO risk modelling. This means that at any point over the three years prior to the experiment they did not have a business tax debt outstanding for more than seven days or of more than $100 in value. The ATO does not send letters to taxpayers if the debt has been outstanding for less than seven days so the businesses in the trial would not have received a reminder letter from the ATO for at least three years. This means we can rule out the possibility that taxpayers in the experiment had become accustomed to delaying payment until receipt of a reminder.

The letter sent to businesses is shown in Appendix A. The letter notes the taxpayer's compliant history and requests payment.[6] Taxpayers are reminded that a general interest charge equal to 8.78 percent per annum (ATO, 2019c), compounded daily, is applied to the debt for each day it remains overdue. The interest rate charged by the ATO is higher than residential-secured small-business lending rates, about the same as unsecured small business lending rates, and lower than credit card and unsecured personal loan rates (Reserve Bank of Australia, 2019). The average interest rate faced by small-to-medium businesses in Australia is about 5.5 percent (OECD, 2017).

The letter reminds taxpayers that they have the option to enter a payment plan if they are unable to pay. This is an agreement to make regular payments over a period generally not longer than 12 months. Typically, a minimum upfront payment of 10 percent is required within seven days of commencement of the payment plan (ATO, 2019d).[7]

---

[6]There were some minor changes made to the wording of the reminder letter a short period of time before the commencement of the trial. The intention was to use language perceived to be more appropriate for a low-risk pool of taxpayers. For example, the title was changed from "You have an overdue tax debt!" to "You have a tax debt". The change in letter wording was not tested experimentally. However, taxpayers in our experiment would not have received a reminder for at least the previous three years and so are unlikely to have adjusted their behavior in response to the revised wording. Furthermore, all taxpayers received the same letter.

[7]Payment plans can be arranged via phone or online for amounts up to $100,000 (ATO, 2019d).

Interest penalties continue to apply during the life of the payment plan. However, entry into a payment plan has the benefit of demonstrating the business is doing its best to meet its obligations and forestalls potential stronger action by the ATO.

It is possible that the letter increased the perceived probability of enforcement of some taxpayers. However, it is more plausible that the letter was primarily perceived as a reminder. First, the ATO prefers to use non-deterrence methods in first interactions with a compliant population of taxpayers. Accordingly, the letter sent to the population in our experiment deliberately uses non-deterrence language without mentioning any consequences associated with non-compliance (see Appendix A). Second, the history of compliant behavior of taxpayers in our sample indicates either a high existing perceived probability of enforcement or a high willingness to comply.

The reminder letter requests that taxpayers make a payment or enter into a payment plan within two weeks of receiving the letter. However, there are no specific consequences associated with this two-week period. The time until the ATO begins escalation of debt collection efforts depended on the length of time since the due date, not the end of this two-week period. Furthermore, interest penalties continued to apply during the two-week period.

Nevertheless, some taxpayers may have perceived the request to comply within two weeks of receiving the reminder as a new deadline. This means we do not have variation in the timing of the reminder relative to a fixed deadline and so cannot test for present bias.[8] However, the letter's request to make payment or enter a payment plan within two weeks does not imply that each of our treatments will have the same effect. If memory is imperfect and payment constraints vary over time, reminders will be more effective if timed when payment constraints are less likely (see Section 2.4).

---

[8]Although the timing of stronger action was fixed, and so varied relative to the reminder for each treatment group, the timing of stronger action was not stated publicly.

4 Trial design, empirical strategy and data

4.1 Trial design

A total of 4,787 unpaid debt cases were quarantined from the usual ATO treatment pathways and randomly allocated to receive a reminder letter either 10, 17 or 25 days following the due date. A control group did not receive a letter for the duration of the trial. About the same number of quarantined cases was randomly assigned to each of the treatment groups and to the control group.[9]

De-identified baseline data were provided by the ATO after the researchers received human ethics approval from the Australian National University.[10] Cases were grouped into strata with similar baseline characteristics to achieve stratified randomization. Within each stratum, each case was assigned at random to one of the treatment groups or to the control group. This procedure ensured that the cases were evenly distributed across groups with regard to baseline characteristics. A complete list of baseline characteristics is provided in Appendix B. Randomization was based on a random variable generator in STATA®, using a random choice of the underlying seed. The user-written command `randtreat` version 1.4 (5 April 2017), available from the `ssc` library, was used for stratified randomization.

The random assignment led to the following allocation of cases: Treatment Group 1: 1,199 cases, Treatment Group 2: 1,200 cases, Treatment Group 3: 1,186 cases and Control Group: 1,202 cases. On 5 June 2017, the ATO sent 1,054 letters to Treatment Group 1 (145 debts had already been paid in the meantime); 926 letters were sent to Treatment Group 2 on 12 June (274 cases had already been paid by that time); Treatment Group 3 received 768 letters that were sent out on 20 June (418 cases had already been paid). De-identified data were made available by the ATO at the end of the data collection period on 14 July 2017.

---

[9]The trial was registered in the RCT Registry of the American Economic Association (https://doi.org/10.1257/rct.3059-1.0).

[10]Human ethics protocol number 2017/454, title: *Testing the Effect of the Timing of Reminder Letters on the Payment Behavior of Business Taxpayers.*

4.2   Empirical strategy

4.2.1   Survival analysis

We estimate Kaplan-Meier survivor functions to study differences in tax debt payments between the treatment and control groups over time. By definition, the survivor function $S(t)$ tells us the probability that a tax debt remains outstanding past time t. We observe a set of "failure times", $\tau_1, \tau_2, \ldots, \tau_J$, where J is the number of distinct uncensored failure times in our data. Kaplan and Meier (1958) propose a nonparametric estimate of the survivor function,

$$\widehat{S}(t) = \prod_{j|\tau_j \leq t} \left(\frac{n_j - d_j}{n_j} \right),$$

where n_j is the number of individuals "at risk" (of paying their taxes) at time τ_j and d_j is the number of "failures" (tax payments) at time τ_j. Following Kalbfleisch and Prentice (2002), we calculate confidence intervals using the asymptotic variance of $\ln(-\ln(\widehat{S}(t)))$,

$$\sigma^2(t) = \frac{\sum_{j|\tau_j \leq t} \frac{d_j}{n_j(n_j - d_j)}}{(\sum_{j|\tau_j \leq t} \ln(\frac{n_j - d_j}{n_j}))^2},$$

to ensure that the confidence intervals are bounded between 0 and 1. The confidence intervals are given by $\left[[\widehat{S}(t)]^{\exp(z_{\alpha/2}\widehat{\sigma}^2)}, [\widehat{S}(t)]^{\exp(-z_{\alpha/2}\widehat{\sigma}^2)} \right]$, where $z_{\alpha/2}$ is the $(1 - \alpha/2)$th quantile of the standard normal distribution. To simplify the interpretation of our results, our discussion will focus on failure functions, which can be estimated by $\widehat{F}(t) = 1 - \widehat{S}(t)$. Failure functions allow us to consider the probability of paying taxes ("failing") at a given point in time, which appears more natural in the context of our analysis than studying the probability of not paying taxes ("surviving").

4.2.2   Treatment effects estimation

Our analysis of treatment effects is based on estimating separate regression models to compare the outcome measures of members of each of the treatment groups to those of the control group. We use the following model to estimate the effect of a reminder letter

on an outcome measure of interest:

$$Y_i = \beta_0 + \beta_1 D_i + X_i \beta_2 + \varepsilon_i,$$

where Y_i refers to one of the outcome measures of taxpayer i, D_i is the treatment indicator for the comparison of one of the treatment groups to the control group, X_i is a set of observed taxpayer characteristics that are used as control variables to balance out potential differences between treatment and control group, and ε_i is the model error term. A complete list of control variables is provided in Appendix B. Our parameter of interest is β_1, the average treatment effect on the treated.

Our analysis focuses on three outcome measures: a 0/1-variable indicating whether or not payments were made by the end of the trial, a 0/1-variable indicating whether or not a partial payment was made at any point during the trial and a continuous variable including the amount (in AUS$) paid by the end of the trial. We use separate regression models excluding and including control variables to estimate unconditional and conditional treatment effects on the treated, respectively. The regression results are presented in Section 5. We estimate a linear probability model to obtain the effect of each of the treatments on the probability of payment by the end of the trial. We compare these estimates to the marginal effects of a binary Probit model to account for the nonlinear nature of the outcome variable. A linear regression model is used to estimate the effect of each of the treatments on the amount paid by the end of the trial. Finally, we perform a range of robustness checks to examine the validity of our results.

4.3 Data

Table 2 includes summary statistics of the control variables used in our analysis. We also present the p-values that refer to the comparison of sample means between each treatment group and the control group. We observe that the average total business income varies from about AUS$730,000 to AUS$970,000 across groups. However, due to the large variation in business income within each of the groups, these differences are not statistically significant (the p-values of the comparison of sample means range from 0.42

14

to 0.96). In our analysis, we will use income quartiles as control variables to be consistent with our stratified randomization.

Table 2 also reports the distribution of businesses across three broad debt levels used by the ATO to categorize tax debts: Debt Level 1: AUS$0 – AUS$2,499, Debt Level 2: AUS$2,500 – AUS$7,499, Debt Level 3: AUS$7,500+. Most businesses (almost 60%) have a relatively low initial tax debt of less than AUS$2,500. Only about 15% of business owe AUS$7,500 or more at the beginning of the trial. Around 7-9% of businesses experience an increase in their outstanding tax debt during the trial period. We perform robustness checks to investigate the impact of including these businesses in our analysis.

[Table 2 about here.]

More than 90% of the businesses in our sample are "micro enterprises" with an annual turnover of less than AUS$2 million and almost half (around 45%) of the businesses are individual/sole traders. The ATO classifies businesses into "risk clusters" based on an internal analytics model. Our analysis sample consists of two risk clusters comprised of businesses that the ATO considers as low-risk clients.

As noted earlier, the vast majority of businesses in our sample (more than 90%) have a tax agent. We do not observe the number of employees for a large fraction of businesses in our data but most businesses appear to have a relatively small number of employees. There is considerable heterogeneity with regard to industries and the geographic distribution of businesses across states. A relatively large fraction of businesses belong to the construction sector (about 13%), financial and insurance services (about 14%), rental, hiring and real estate services (about 11%) and professional, scientific and technical services (about 11%). More than 70% of businesses are located in New South Wales, Victoria or Queensland.

In the following analysis, we will use the variables reported in Table 2 as control variables to balance out potential differences between observed characteristics when estimating treatment effects. We will also study heterogenous treatment effects for the most important control variables. Section 5.3 includes a summary of the main findings of our analysis of heterogenous treatment effects.

5 Results

5.1 Payment rates and share of debt paid

Figure 1 shows the Kaplan-Meier failure function for the control group and each of the treatment groups over the duration of the trial. Each line depicts the share of cases paid (either partially or fully) by group and by day since the beginning of the trial. The solid line in Figure 1 reveals that the share of cases with a payment in the control group remains low over the entire study period. No payments had been made for almost 50% of the cases in the control group up to 52 days after the due date. In contrast, more than 75% of the cases in the three treatment groups received payment within 52 days (no payments were made for less than 25% of the cases). The difference between the treatment groups and the control group constitutes the causal effect of the reminder. Differences between the three treatment groups and the control group are statistically significant.

[Figure 1 about here.]

Among the treatment groups, we observe that the early reminders resulted in earlier payment of outstanding debts than the late reminders. Critically, the three treatment groups converge to the same payment proportion by the end of the trial period. This indicates that early reminders accelerate tax debt payment without reducing the cumulative probability of payment. To compare the *rate* at which payments are made across the three treatment groups, we estimate the payment profiles for each group using the reminder send date (instead of the due date) as "day 0". The estimates, which are presented in Figure 2, reveal that the payment profiles do not differ significantly between the three treatment groups, indicating that the rate at which payments were made did not depend on whether the reminder was sent out earlier or later.

For each treatment, we observe a sharp increase in payment immediately after the reminder and a smooth profile thereafter (Figures 1 and 2). Some research has interpreted task completion close to a deadline as evidence of present-biased preferences (e.g. Frakes and Wasserman, 2017). However, Heidhues and Strack (2019) show theoretically that, without making parametric assumptions, any distribution of completion times is

consistent with both time-consistent and present-biased preferences. The payment profile is like that observed by Brockmeyer et al. (2016), who found a sharp increase in filing and payment immediately after delivery of an enforcement email to previously non-filing firms in Costa Rica.

[Figure 2 about here.]

So far, we have focused on the share of cases with payments. Figure 3 provides information about the share of debt that was paid over time. We observe that businesses in the control group repaid about 75% of their debt by the end of the trial period. In contrast, businesses in the three treatment groups repaid close to 90% of their debt. The share of debt paid in each group is larger than the share of payments (see Figure 1) because the payment probability is larger for high value debts than for low value debts.

[Figure 3 about here.]

5.2 Treatment effects

Estimates of the unconditional treatment effects obtained from a linear probability model (Table 3, Panel A) indicate that the reminder letters increased the likelihood of tax debt payment by about 24 percentage points by the end of the trial period, regardless of the point in time at which the reminders were sent out. We observe considerable heterogeneity in the treatment effects when initial debt levels are considered. While the reminders increased the payment rates of businesses with an initial tax debt of less than AUS$7,500 by about 28 percentage points, businesses with an initial debt of AUS$7,500 or more did not significantly change their behavior as a result of receiving a reminder (Table 3, Panel A). The treatment effect estimates presented in Panel A of Table 3 are remarkably stable and do not change much when we control for business characteristics (Panel B). We also find that the estimates obtained from a binary Probit model (Panel C) are about the same as those of the linear probability model.

[Table 3 about here.]

Table 4 reports OLS estimates of the treatment effects on the amount paid by the end of the trial. The numbers in Panel A of Table 4 indicate that the overall treatment effects are not statistically significant. However, we observe significant treatment effects when we estimate separate regressions by initial debt level. While businesses with an initial debt of below AUS$7,500 increase their tax debt payments by about AUS$500 if they receive a reminder, we find no significant effect of receiving a reminder on tax debt payments of businesses with a tax debt of AUS$7,500 or more. Our findings do not change qualitatively if we add control variables to our model (Panel B).

[Table 4 about here.]

Taken together, the results in Tables 3 and 4 indicate that the reminders could increase both payment rates and the amount of payments but that the effects are driven entirely by businesses with relatively low (below AUS$7,500) initial debt levels. This finding suggests that reminders are ineffective if initial debt levels are too high. To determine the "tipping point" beyond which the reminders are no longer effective, we examine the effect of receiving a reminder (i.e. the effect of being in one of the three treatment groups) on debt payments by the initial debt level of the taxpayer. Figure 4 shows the difference between businesses that do and do not receive a reminder within each decile of the initial debt level. We find that the reminders were highly effective for initial debt levels of about AUS$8,000 or less. Beyond that threshold, the effect of the reminder on debt payments is not significant, suggesting that reminders are ineffective if the initial amount of debt is larger than AUS$8,000. However, it is worth noting that high value debt cases not receiving a reminder received payment at a similar rate to low value debt cases that did receive a reminder.

[Figure 4 about here.]

The results presented here do not differentiate between full and partial payments. About 13 percent of all payments made over the trial period were partial payments. To understand the contribution of partial payment to the overall patterns, we can exclude partial payments from our analysis sample. We find that the treatment effects reported in

Table 3 increase if partial payments are removed from the sample. For the unconditional linear probability model, the treatment effects rise by about 7 percentage points for each group. The estimated treatment effects and standard errors are 0.318 (0.019) for Treatment 1, 0.308 (0.019) for Treatment 2 and 0.301 (0.019) for Treatment 3. The additional effects resulting from the exclusion of partial payment on the amount of debt paid range from AUS$38 (Treatment 1) to AUS$93 (Treatment 3) for debts less than AUS$7,500, but we continue to see no significant effect of the reminder on high value debt cases (initial debt levels more than AUS$7,500).

We also investigate the effect of the reminder on partial payments by excluding full payments from the analysis sample. Unconditional linear probability model estimates indicate that the reminder significantly increased the probability of a partial payment being made by the end of the trial period. The observed effects and standard errors of Treatments 1-3 are 0.171 (0.036), 0.149 (0.035) and 0.117 (0.034), respectively. These estimates indicate that the probability of a partial payment being made by the end of the trial period was larger for early than late reminders. This is the only piece of evidence we find that the timing of the reminders affects the payment probability. However, we cannot reject equality across the three treatment groups at a 5 percent significance level. The amount of debt paid was higher in each treatment group for initial debt levels below AUS$7,500 but the amounts are imprecisely estimated owing to the smaller sample size after excluding full payments.

5.3   Robustness checks

We perform a number of robustness checks to test the validity of the results presented in Section 5.2. First, the conditional models in Tables 3 and 4 include a large set of control variables that may be correlated and may therefore offset each other. To address this concern, we re-estimate the models presented in Tables 3 and 4 but instead of including all control variables, we only include single (groups of) control variables (consistent with the grouping of variables in Table 2). We find that the inclusion of single (groups of) control variables does not affect our findings qualitatively.

Second, Tables 3 and 4 report heterogenous treatment effects by initial debt level but

we ignore heterogeneity with regard to other factors. Excluding debt cases automatically raised by the ATO (32% of cases) or businesses with a monthly filing deadline (7% of cases) has a negligible effect on the results. Excluding BAS cases lodged through channels used exclusively by tax agents modestly increases the estimated probability of payment by the end of the trial, by around 5 percentage points. We observe a similar pattern of heterogeneity in treatment effects by initial debt level and by total business income, but we do not observe heterogeneity by business debt-to-income ratio.

Third, we also estimate heterogenous treatment effects for all other (groups of) control variables presented in Table 2. We only find heterogeneity when we estimate separate treatment effects for cases that did or did not experience an increase in their outstanding debt over the trial period. We observe that the reminder letters did not lead to a significant increase in payment probabilities if businesses experienced an increase in outstanding debt. The treatment effects observed for businesses that did not experience an increase in outstanding debt are about the same as those presented in Table 3. We do not find heterogeneity when we estimate the corresponding treatment effects on the amount paid by the end of the trial. We study the implications of excluding cases that experienced an increase in outstanding debt over the trial period from our sample and find that their exclusion does not affect the results presented in Tables 3 and 4 qualitatively. Our treatment effects are even a bit (although not significantly) larger when we exclude these cases from our sample.

Fourth, we consider multiple hypothesis testing to account for the possibility of false positives. Following List et al. (2016), we carry out hypothesis testing for multiple outcomes (the payment probability and the amount paid by the end of the trial), treatments (variation in the timing of reminders) and subgroups (initial debt levels below or above AUS$7,500). We also compare the treatment groups to each other. Table C.1 in Appendix C reports the group differences in means together with alternative p-values associated with these differences.[11]

The numbers in Table C.1 reveal that the multiple hypothesis test results are consistent with our main findings (presented in Tables 3 and 4). All treatments cause significant

---

[11]See List et al. (2016) for a detailed description of the approach and the interpretation of alternative p-values.

differences in outcomes between treatment and control businesses with an initial debt of less than AUS\$7,500 (Subgroup 1). In contrast, differences in outcomes between treatment and control businesses with an initial debt of AUS\$7,500+ (Subgroup 2) are not significant. We also do not observe significant differences between treatment groups. Overall, these results indicate that multiple hypothesis testing does not affect our conclusions.

Finally, the ATO provided us with additional data on payment rates beyond the end of the trial period. The share of debts paid in each treatment group remained almost identical beyond the end of the trial period. Strikingly, the share of debts paid in the control group after about 150 days remained below that of the treatment groups. We do not present any results based on data collected after the end of the trial period because we have no information on further actions taken by the ATO after the end of the trial.

5.4 Mechanisms

Our empirical results shed light on the presence of imperfect memory and payment constraints as mechanisms affecting small-business tax debt payment. The spike in payments observed shortly after the receipt of the reminder is consistent with at least some taxpayers experiencing imperfect memory (Figure 1 and Table 1, Case 1). This is because it is optimal to pay immediately if the debt is in memory and the business has a cost of funds lower than the ATO's interest rate penalty. We also find evidence of payment constraints for some businesses, with about 13 percent of the payments observed over the trial period being partial payments. The reminder letter caused an increase in partial payments, indicating that some business taxpayers facing payment constraints also experienced imperfect memory (Table 1, Case 2). The ineffectiveness of the reminder for high-debt cases is consistent with payment constraints but not with imperfect memory (Case 3). We do not find evidence that variation in the timing of the reminder affects the probability of full payments. This is inconsistent with payment constraints that decay quickly and taxpayers forgetting early reminders (Case 2). It is also inconsistent with payment constraints that increase with time since the due date (Case 2). Taken together, our findings are consistent with imperfect memory and payment constraints being frictions affecting tax debt payment. However, we do not find evidence that payment constraints and imperfect

memory interact to make early reminders more or less effective than late reminders.

5.5 Implications for the optimal timing of reminders

Our results indicate that the payment probability is the same for reminders sent one, two and three weeks after the due date. However, there are still trade-offs between early and late reminder letters. We discuss first the revenue implications for the tax authority and then the welfare implications.

For each day a tax debt is outstanding the ATO levies interest penalties at the rate of 8.78 percent per annum, compounded daily. Thus, delaying the date at which a debt is paid raises more revenue from penalties. Delaying the reminder also reduces cost, because some debts will be paid in the intervening period and fewer letters need to be sent. However, compared with interest penalties, the cost of sending the reminder letters is low, at about AUS$1.25 per letter, including postage. Revenue raised is around AUS$4,200 higher when the letters are sent in week three rather than week one (Table 5). However, this is modest compared with the approximately AUS$25.6m of total debt outstanding in the analysis sample at the beginning of the trial.

[Table 5 about here.]

If taxpayers experience only imperfect memory then delaying the reminder could be socially desirable. There is no behavioral response available to taxpayers to reduce their debt so the additional penalty revenue raised from a delayed reminder is akin to a lump-sum tax. However, if taxpayers became aware the tax authority was delaying reminders to increase penalty income, tax morale could decrease, leading to less compliance in the future. For businesses experiencing payment constraints and imperfect memory, an early reminder may lead to a payment plan being setup more quickly, provided the business can meet the minimum payment requirement. This benefits the business through lower interest penalties and demonstrating to the tax authority a willingness to comply. We discussed the results with the ATO and they preferred an early reminder. This prioritizes timely compliance and engagement with businesses, which is viewed to be important for maintaining a high rate of voluntary compliance.

6   Conclusions

This paper provides evidence on the relevance of the timing of nudges by studying how varying the timing of reminder letters to taxpayers affects their tax payment behavior. All debt collectors face a choice over the timing of when to send a reminder. Despite this, there is little rigorous empirical evidence on the effect of the timing of reminders on payment behavior. We contribute to the literature by generating evidence on the effect of the timing of reminders in a large-scale field experiment with business taxpayers in Australia. Small businesses were randomly chosen to receive a reminder letter either one, two or three weeks following their missed tax debt due date; a control group received no contact from the tax authority for the duration of the seven-week field experiment.

We find that the probability of payment at the end of the seven-week field experiment was 25 percentage points greater for each group that received a reminder letter relative to the control group. However, there was no difference in the probability of payment between treatment groups at the end of the seven-week trial period, regardless of whether the reminder letter was sent one, two or three weeks after the due date. Measured relative to the date at which the reminder letter was sent, the probability of payment evolved almost identically across the three groups receiving a reminder letter.

Our results imply that sending reminder letters early accelerates the collection of tax debts and has no effect on the ultimate probability of payment. While delaying the timing of the reminder letters increases revenue collected through interest penalties, the amount of additional revenue collected relative to debt outstanding is modest. Furthermore, although penalties levied on debt already incurred raises lump-sum revenue, deliberately not alerting taxpayers about their overdue debts is likely to be counter-productive in the long-term.

Our results provide insight into some of the mechanisms affecting tax debt payment. We find evidence businesses did not pay their debts on time because of payment constraints and imperfect memory. Some businesses showed behavior consistent with both payment constraints and imperfect memory. The only meaningful heterogeneity in payment behavior is related to the size of debts. Reminder letters did not increase the

23

probability of payment for high value debts (AUS$7,500+). This could indicate that payment constraints are a more important barrier to payment than imperfect memory for taxpayers with high value debts.

Tables and Figures

TABLE 1: MECHANISMS AND EXPECTED BEHAVIORS

Case	Imperfect memory	Payment constraint	Expected behavior
1	✓	✗	The reminder causes an increase in re-payments. Early and late reminders are equally effective.
2	✓	✓	Late (early) reminders more effective than early (late) reminders if payment constraints decrease (increase) with time since the due date. Business enters a payment plan if it can meet the minimum payment amount.
3	✗	✓	Business enters a payment plan if it can meet the minimum payment amount. The reminder has no effect on the likelihood or timing of entry into a payment plan.
4	✗	✗	The reminder has no effect on repayment behavior.

TABLE 2: BASELINE CHARACTERISTICS

	Control		Treatment 1			Treatment 2			Treatment 3		
	Mean	N	Mean	N	p-value	Mean	N	p-value	Mean	N	p-value
Total Business Income	733,688	1,186	968,881	1,184	0.423	898,547	1,186	0.606	746,371	1,172	0.960
Initial Debt Level											
$0 - $2,499	0.577	1,202	0.580	1,199	0.877	0.583	1,200	0.799	0.586	1,186	0.669
$2,500 - $7,499	0.268	1,202	0.269	1,199	0.970	0.265	1,200	0.873	0.265	1,186	0.863
$7,500+	0.155	1,202	0.151	1,199	0.797	0.152	1,200	0.879	0.149	1,186	0.708
Increase in Outstanding Debt	0.077	1,202	0.067	1,199	0.313	0.080	1,200	0.811	0.086	1,186	0.441
Business Market Segment											
Micro enterprise	0.911	1,186	0.919	1,184	0.470	0.906	1,186	0.670	0.910	1,172	0.928
Small/Medium enterprise	0.041	1,186	0.039	1,184	0.760	0.046	1,186	0.615	0.047	1,172	0.507
Not for profit enterprise	0.048	1,186	0.042	1,184	0.494	0.049	1,186	0.924	0.044	1,172	0.598
Client Type											
Australian Private Company	0.230	1,202	0.229	1,199	0.949	0.227	1,200	0.863	0.228	1,186	0.909
Individual/Sole Trader	0.452	1,202	0.452	1,199	0.988	0.452	1,200	0.997	0.453	1,186	0.959
Other	0.319	1,202	0.319	1,199	0.967	0.322	1,200	0.874	0.320	1,186	0.961
Risk Cluster											
Risk Cluster 2	0.667	1,202	0.666	1,199	0.965	0.671	1,200	0.851	0.670	1,186	0.872
Risk Cluster 7	0.333	1,202	0.334	1,199	0.965	0.329	1,200	0.851	0.330	1,186	0.872
Lodgement Channel											
ATO Online	0.053	1,175	0.067	1,161	0.142	0.055	1,178	0.796	0.051	1,158	0.843
Auto Finalised	0.328	1,175	0.309	1,161	0.339	0.312	1,178	0.427	0.328	1,158	0.980
Agent Portal	0.017	1,175	0.036	1,161	0.004	0.025	1,178	0.155	0.022	1,158	0.346
Business Portal	0.130	1,175	0.134	1,161	0.767	0.111	1,178	0.157	0.133	1,158	0.843
Corporate Data Capture	0.060	1,175	0.065	1,161	0.557	0.070	1,178	0.284	0.070	1,158	0.309
Electronic Lodgement	0.220	1,175	0.218	1,161	0.883	0.215	1,178	0.740	0.216	1,158	0.791
Tax Agent Portal	0.134	1,175	0.132	1,161	0.896	0.159	1,178	0.084	0.117	1,158	0.214
Web Services	0.049	1,175	0.029	1,161	0.013	0.041	1,178	0.314	0.050	1,158	0.936
Other	0.009	1,175	0.009	1,161	0.848	0.011	1,178	0.686	0.013	1,158	0.409

TABLE 2 (CONTINUED)

	Control		Treatment 1			Treatment 2			Treatment 3		
	Mean	N	Mean	N	p-value	Mean	N	p-value	Mean	N	p-value
Tax Agent	0.937	1,202	0.932	1,199	0.610	0.932	1,200	0.672	0.946	1,186	0.335
Number of Employees											
1	0.097	1,202	0.093	1,199	0.796	0.112	1,200	0.224	0.094	1,186	0.863
2-5	0.106	1,202	0.118	1,199	0.388	0.116	1,200	0.466	0.113	1,186	0.612
6-19	0.052	1,202	0.053	1,199	0.989	0.052	1,200	0.934	0.051	1,186	0.914
20+	0.012	1,202	0.018	1,199	0.243	0.019	1,200	0.189	0.018	1,186	0.295
Missing	0.732	1,202	0.718	1,199	0.442	0.702	1,200	0.098	0.723	1,186	0.634
Industry											
Agriculture, Forestry, Fishing	0.071	1,186	0.068	1,184	0.755	0.063	1,186	0.460	0.072	1,172	0.936
Manufacturing	0.028	1,186	0.025	1,184	0.707	0.039	1,186	0.137	0.029	1,172	0.863
Construction	0.138	1,186	0.141	1,184	0.846	0.131	1,186	0.588	0.128	1,172	0.462
Wholesale Trade	0.022	1,186	0.024	1,184	0.778	0.024	1,186	0.682	0.027	1,172	0.399
Retail Trade	0.048	1,186	0.036	1,184	0.155	0.038	1,186	0.225	0.038	1,172	0.207
Accommodation and Food Services	0.033	1,186	0.034	1,184	0.903	0.025	1,186	0.272	0.024	1,172	0.189
Transport, Postal, Warehousing	0.049	1,186	0.046	1,184	0.705	0.053	1,186	0.641	0.061	1,172	0.183
Financial and Insurance Services	0.142	1,186	0.149	1,184	0.630	0.152	1,186	0.524	0.144	1,172	0.906
Rental, Hiring, Real Estate Services	0.107	1,186	0.120	1,184	0.324	0.116	1,186	0.514	0.109	1,172	0.868
Professional, Scientific, Technical Services	0.109	1,186	0.118	1,184	0.467	0.108	1,186	0.947	0.115	1,172	0.621
Administrative and Support Services	0.030	1,186	0.045	1,184	0.049	0.026	1,186	0.618	0.041	1,172	0.132
Health Care and Social Assistance	0.053	1,186	0.046	1,184	0.456	0.052	1,186	0.927	0.049	1,172	0.690
Miscellaneous	0.170	1,186	0.148	1,184	0.134	0.173	1,186	0.870	0.162	1,172	0.593
State											
New South Wales	0.345	1,186	0.326	1,184	0.331	0.361	1,186	0.414	0.363	1,172	0.345
Queensland	0.218	1,186	0.231	1,184	0.447	0.209	1,186	0.616	0.219	1,172	0.918
South Australia	0.054	1,186	0.057	1,184	0.780	0.056	1,186	0.857	0.047	1,172	0.435
Tasmania, ACT, NT†	0.030	1,186	0.037	1,184	0.359	0.030	1,186	0.904	0.030	1,172	0.944
Victoria	0.262	1,186	0.249	1,184	0.466	0.245	1,186	0.345	0.256	1,172	0.729
Western Australia	0.091	1,186	0.101	1,184	0.435	0.099	1,186	0.484	0.084	1,172	0.572

Note: p-values refer to the comparison of means between treatment and control groups.
† ACT: Australian Capital Territory, NT: Northern Territory.
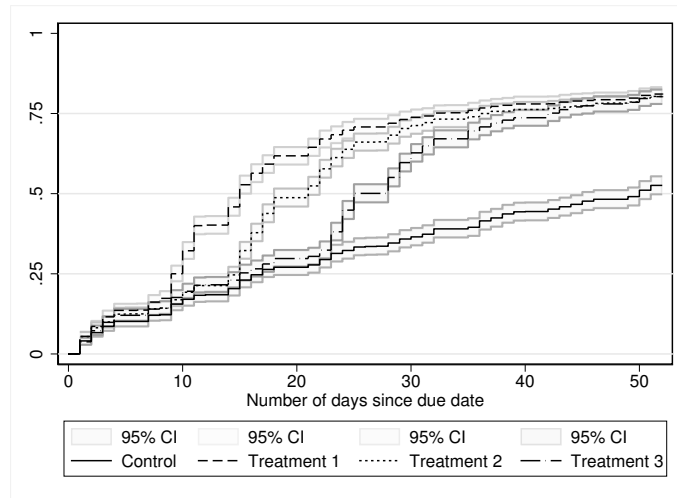
FIGURE 1: Payment Rates: Kaplan-Meier Failure Estimates

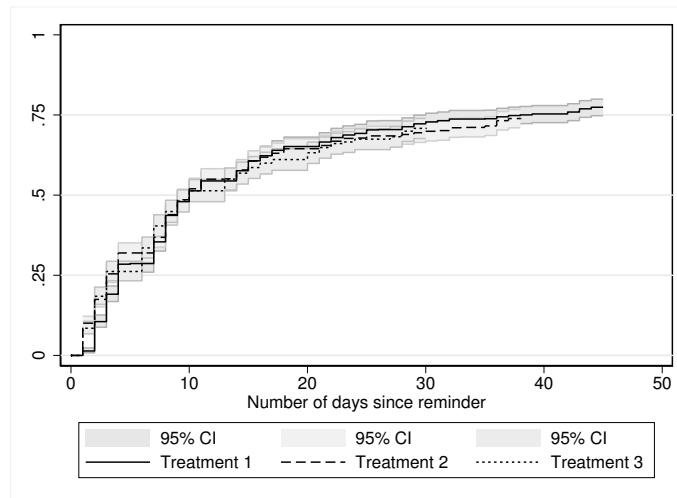FIGURE 2: Comparison of Payment Profiles
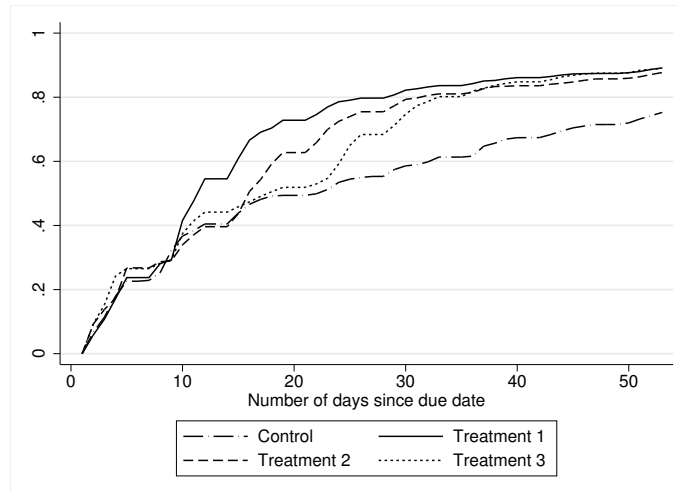

FIGURE 3: Share of Debt Paid

TABLE 3: TREATMENT EFFECTS ON PAYMENT MADE BY END OF TRIAL

	Treatment 1	Treatment 2	Treatment 3
PANEL A: UNCONDITIONAL LINEAR PROBABILITY MODEL			
Payment Made by End of Trial	0.248**	0.238**	0.234**
	(0.017)	(0.017)	(0.017)
	[2,401]	[2,402]	[2,388]
By Initial Debt Level			
$0 - $7,499	0.289**	0.279**	0.282**
	(0.019)	(0.019)	(0.019)
	[2,034]	[2,033]	[2,025]
$7,500+	0.025	0.015	-0.028
	(0.029)	(0.030)	(0.033)
	[367]	[369]	[363]
PANEL B: CONDITIONAL LINEAR PROBABILITY MODEL			
Payment Made by End of Trial	0.248**	0.235**	0.229**
	(0.017)	(0.017)	(0.017)
	[2,305]	[2,323]	[2,303]
By Initial Debt Level			
$0 - $7,499	0.290**	0.278**	0.278**
	(0.019)	(0.019)	(0.019)
	[1,947]	[1,959]	[1,949]
$7,500+	0.009	0.000	-0.028
	(0.030)	(0.030)	(0.033)
	[358]	[364]	[354]
PANEL C: CONDITIONAL PROBIT MODEL (MARGINAL EFFECTS)			
Payment Made by End of Trial	0.250**	0.235**	0.232**
	(0.017)	(0.017)	(0.017)
	[2,305]	[2,323]	[2,303]
By Initial Debt Level			
$0 - $7,499	0.296**	0.283**	0.284**
	(0.019)	(0.019)	(0.019)
	[1,947]	[1,959]	[1,949]
$7,500+	0.021	-0.002	-0.035
	(0.028)	(0.019)	(0.027)
	[260]	[325]	[298]

Note: Robust standard errors in parentheses. Number of observations in brackets.
* $p < 0.05$, ** $p < 0.01$.

TABLE 4: TREATMENT EFFECTS ON AMOUNT PAID BY END OF TRIAL

	Treatment 1	Treatment 2	Treatment 3
PANEL A: UNCONDITIONAL LINEAR REGRESSION MODEL			
Amount Paid by End of Trial	590.94	252.64	634.77
	(762.23)	(530.19)	(587.46)
	[2,401]	[2,402]	[2,388]
By Initial Debt Level			
$0 - $7,499	540.69**	456.25**	509.95**
	(78.71)	(90.75)	(91.73)
	[2,034]	[2,033]	[2,025]
$7,500+	1303.17	-918.90	1793.00
	(4656.92)	(2957.92)	(3368.23)
	[367]	[369]	[363]
PANEL B: CONDITIONAL LINEAR REGRESSION MODEL			
Amount Paid by End of Trial	797.05	135.41	614.75
	(802.71)	(449.77)	(515.95)
	[2,305]	[2,323]	[2,303]
By Initial Debt Level			
$0 - $7,499	549.24**	465.36**	542.27**
	(62.27)	(78.11)	(74.44)
	[1,947]	[1,959]	[1,949]
$7,500+	-621.35	-2095.43	615.78
	(4536.22)	(2719.29)	(3368.91)
	[358]	[364]	[354]

Note: Robust standard errors in parentheses. Number of observations in brackets.
* $p < 0.05$, ** $p < 0.01$.
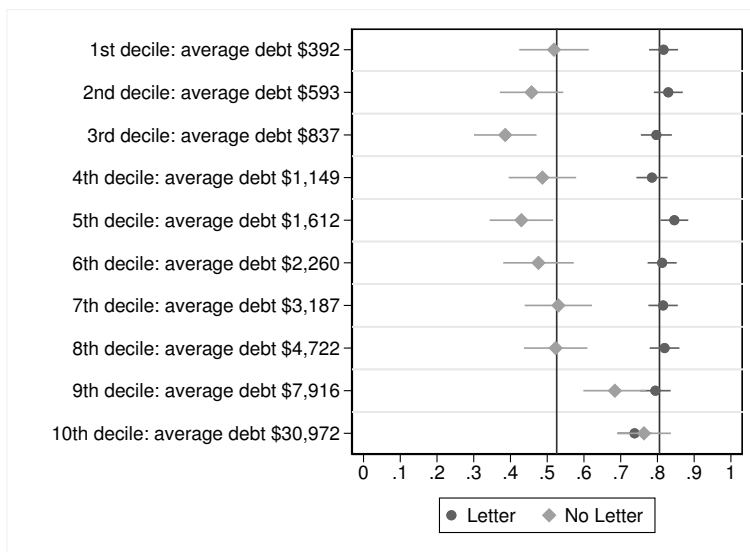
FIGURE 4: Treatment Effect by Initial Debt Level

TABLE 5: REVENUE IMPLICATIONS

Trial group	Total interest penalties by day 52	Number of letters sent	Cost of letters (at $1.25 per letter)	Interest penalties less cost of letters	Share cases paid by day 52
No letter	$23,742	0	$0	$23,742	0.53
Week 1	$14,532	1,054	$1,318	$13,214	0.81
Week 2	$16,561	926	$1,158	$15,403	0.80
Week 3	$18,414	768	$960	$17,454	0.80

References

Allcott, H., "Social Norms and Energy Conservation," *Journal of Public Economics*, 2011, *95* (5), 1982–2095.

Ariely, D., *Predictably Irrational*, New York: Harper Collins, 2008.

ATO, "Commisioner of Taxation Annual Report 2016-17," Australian Taxation Office (ATO), Australian Government, Commonwealth of Australia, Canberra 2017a.

_ , "If You Can't Lodge and Pay on Time," `https://www.ato.gov.au/Business/Reports-and-returns/What-if-you-can-t-lodge-and-pay-on-time-/` 2017b. (Accessed: 1 March 2019).

_ , "Tax Practitioner Stewardship Group Minutes 1 December 2017, Section 5: Debt," `https://www.ato.gov.au/General/Consultation/In-detail/Stewardship-groups-minutes/Tax-Practitioner-Stewardship-Group/TPSG-minutes-1-December-2017/?page=7#5__Debt` 2017c. (Accessed: 1 March 2019).

_ , "If You Don't Pay," `https://www.ato.gov.au/general/paying-the-ato/if-you-don-t-pay/` 2019a. (Accessed: 1 March 2019).

_ , "Small Business: The Big Picture," `https://www.ato.gov.au/about-ato/managing-the-tax-and-super-system/in-detail/insight--big-pictures/small-business--the-big-picture/` 2019b. (Accessed: 1 March 2019).

_ , "General Interest Charge (GIC) Rates," `https://www.ato.gov.au/rates/general-interest-charge-%28gic%29-rates/` 2019c. (Accessed: 1 March 2019).

_ , "Payment Plan Estimator," `https://www.ato.gov.au/Calculators-and-tools/` 2019d. (Accessed: 1 March 2019).

BIT, "EAST - Four Simple Ways to Apply Behavioural Insights," Report, Behavioural Insights Team (BIT), UK 2014.

Boning, William C., John Guyton, Ronald H. Hodge II, Joel Slemrod, and Ugo Troiano, "Heard it Through the Grapevine: Direct and Network Effects of a Tax Enforcement Field Experiment," NBER Working Paper Series 24305, National Bureau of Economic Research (NBER), Cambridge, MA 2018.

Brockmeyer, Anne, Marco Hernandez, Stewart Kettle, and Spencer Smith, "Casting the Tax Net Wider: Experimental Evidence from Costa Rica," World Bank Policy Research Working Paper No. 7850 2016.

Cadena, X. and A. Schoar, "Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments," NBER Working Paper Series 17020, National Bureau of Economic Research (NBER), Cambridge, MA 2011.

Cadena, Ximena, Antoinette Schoar, Alexandra Cristea, and Heber M. Delgado-Medrano, "Fighting Procrastination in the Workplace: An Experiment," NBER Working Paper Series 16944, National Bureau of Economic Research (NBER), Cambridge, MA 2011.

Carpio, Lucia Del, "Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru," 2013. Unpublished Manuscript, Princeton University.

Castro, Lucio and Carlos Scartascini, "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment," *Journal of Economic Behaviour and Organization*, 2015, *116*, 65–82.

Cialdini, R. B., *Influence: Science and Practice*, Boston: Pearson, 2008.

Damgaard, Mette Trier and Christina Gravert, "The Hidden Cost of Nudging: Experimental Evidence from Reminders in Fundraising," *Journal of Public Economics*, 2018, *157*, 15–26.

Dun and Bradstreet, "Late Payments in Australia: Second Quarter Analysis 2017," `http://dnb.com.au/_media/documents/AU%20Late%20Payments%202Q17%20Final.pdf` 2017. (Accessed: 1 March 2019).

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke, "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany," *American Economic Journal: Economic Policy*, 2016, *8* (3), 203–232.

Ericson, K. M., "On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation," *Journal of the European Economic Association*, 2017, *15* (3), 692–719.

et al Bobrow, Kristen, "Mobile Phone Text Messages to Support Treatment Adherence in Adults With High Blood Pressure (StAR): A Single-Blind, Randomized Trial," *Circulation*, 2016, *133* (6), 592–600.

Frakes, M. D. and M. F. Wasserman, "Procrastination in the Workplace: Evidence from the U.S. Patent Office," NBER Working Paper Series 22987, National Bureau of Economic Research (NBER), Cambridge, MA 2017.

Gächter, S., H. Orzen, E. Renner, and C. Starmer, "Are Experimental Economists Prone to Framing Effects? A Natural Field Experiment," *Journal of Economic Behavior and Organization*, 2009, *70* (3), 443–446.

Hallsworth, Michael, "The Use of Field Experiments to Increase Tax Compliance," *Oxford Review of Economic Policy*, 2014, *30* (4), 658–679.

_ , John A. List, Robert D. Metcalfe, and Ivo Vlaev, "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance," *Journal of Public Economics*, April 2017, *148*, 14–31.

Harju, Jarkko, Tuomas Kosonen, and Olli Ropponen, "Do Honest Hairdressers Get a Haircut?," Working Paper, Government Institute for Economic Research (VATT) 2014.

Hasseldine, John, Peggy Hite, Simon James, and Marika Toumi, "Persuasive Communications: Tax Compliance Enforcement Strategies for Sole Proprietors," *Contemporary Accounting Research*, 2007, *24* (1), 171–194.

Heidhues, P. and P. Strack, "Identifying Present-Bias from the Timing of Choices," Working Paper, Available at: https://arxiv.org/abs/1905.03959 (Accessed: 10 September 2019) 2019.

Holman, Jeff and Farhan Zaidi, "The Economics of Prospective Memory," Working Paper, Available at SSRN: http://dx.doi.org/10.2139/ssrn.1662183 (Accessed: 1 March 2019) 2010.

IRS, "Tax Gap Estimates for Tax Years 2008-2010," Available at: https://www.irs.gov/newsroom/the-tax-gap (Accessed: 1 March 2019), US Department of the Treasury, Internal Revenue Service (IRS) 2016.

Kahneman, D. and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kalbfleisch, J. D. and R. L. Prentice, *The Statistical Analysis of Failure Time Data, 2nd edition*, New York: Wiley, 2002.

Kaplan, E. L. and P. Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 1958, *53* (282), 457–481.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman, "Getting to the Top of Mind: How Reminders Increase Saving," *Management Science*, 2016, *62* (12), 3393–3411.

_ , **Melanie Morten, and Jonathan Zinman**, "A Personal Touch in Text Messaging Can Improve Microloan Payments," *Behavioral Science and Policy*, 2015, *2*, 25–32.

Kleven, Henrik Jacobsen, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez, "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

Levin, I. P., S. L. Schneider, and G. J. Gaeth, "All Frames are Not Created Equal: A Typology and Critical Analysis of Framing Effects," *Organizational Behavior and Human Decision Processes*, 1998, *76* (2), 149–188.

List, John A., Azeem M. Shaikh, and Yang Xu, "Multiple Hypothesis Testing in Experimental Economics," NBER Working Paper Series 21875, National Bureau of Economic Research (NBER), Cambridge, MA 2016.

Mascagni, Giulia, "From the lab to the field: A review of tax experiments," *Journal of Economic Surveys*, 2018, *32* (2), 273–301.

Mullainathan, Sendhil and Eldar Shafir, *Scarcity: The New Science of Having Less and How it Defines our Lives*, New York: Times Books, 2014.

O'Donoghue, T. and M. Rabin, "Doing it Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

OECD, "Australia," in "Financing SMEs and Entrepreneurs 2017: An OECD Scoreboard" Organization for Economic Cooperation and Development, Paris. Available at: `https://doi.org/10.1787/fin_sme_ent-2017-11-en`. (Accessed: 1 March 2019) 2017.

Ortega, Daniel E. and Carlos Scartascini, "A Field Experiment on Delivery Methods for Increasing Tax Compliance," Working Paper, Inter American Development Bank Working paper no. 627 2015.

_ **and Pablo Sanguinetti**, "Deterrence and Reciprocity Effects on Tax Compliance: Experimental Evidence from Venezuela," Working Paper, CAF Working paper 2013/08 2013.

Pomeranz, Dina, "No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 2015, *105* (8), 2539–2569.

Reserve Bank of Australia, "Statistical Table F5," `https://www.rba.gov.au/statistics/tables/` 2019. (Accessed: 1 March 2019).

Roe, Katie and Katina Law, "Small Business Cash Flow Education Research 2017," Technical Report, IMS Management Consultants for the Australian Taxation Office. Available at: `https://www.ato.gov.au/uploadedFiles/Content/CR/downloads/Small%20Business%20Cash%20Flow%20Education%20Research%202017.pdf`. (Accessed 1 March 2019) 2017.

Scott, P. J. and C. Lizieri, "Consumer House Price Judgments: New Evidence on Anchoring and Arbitrary Coherence," *Journal of Property Research*, 2012, *29* (1), 49–68.

Slemrod, Joel, "Tax Compliance and Enforcement," NBER Working Paper Series 24799, National Bureau of Economic Research (NBER), Cambridge, MA July 2018.

_ **, Marsha Blumenthal, and Charles Christian**, "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota," *Journal of Public Economics*, 2001, *79* (3), 455–483.

Sunstein, Cass R., "Nudging: A Very Short Guide," *Journal of Consumer Policy*, 2014, *37* (4), 583–588.

Taubinsky, Dmitry, "From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice," Working Paper, Working Paper, Available at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:12269825 2014.

Thaler, R. and C. Sunstein, *Nudge: Improving Decisions about Health, Wealth and Happiness*, New Haven: Yale University Press, 2008.

Tversky, A. and D. Kahneman, "Judgement under Uncertainty: Heuristics and Biases," *Science (New Series)*, 1974, *185* (4157), 1124–1131.

Wansink, B., R. J. Kent, and S. J. Hoch, "An Anchoring and Adjustment Model of Purchase Quantity Decisions," *Journal of Marketing Research*, 1998, *35* (1), 71–81.

Wenzel, Michael, "A Letter from the Tax Office: Compliance Effects of Informational and Interpersonal Justice," *Social Justice Research*, 2006, *19* (3), 345–364.

Appendix A – Treatment Letter

Australian Government
Australian Taxation Office

||||||||||||||||||||||||||||||||||||||||||||||||||
<TITLE> <FIRST NAME> <MIDDLE NAME> <SURNAME>
<ORGANISATION>
<ADDRESS LINE 1>
<ADDRESS LINE 2>
<LOCALITY <STATE> <POSTCODE>
<COUNTRY>

Our reference:	<Corres ID>
Phone:	< >
<TFN<ABN>:	<TFN><ABN>
	<Issue date>

You have a tax debt

Total due now
<$Total amount>

> Have you missed a tax bill?
> You can choose from various payment methods

Dear <FIRST NAME><Sir/Madam>,

We haven't received your payment for your <tax type> bill yet. We can see you normally lodge and pay on time, so in case you can't find the details, here they are again.

Tax type	Amount owing
<Account type>	<$Total amount>

If you have paid in the last 7 days, thank you. No further action is required.

What you need to do
You need to pay your overdue debt by [14 days from issue date]<due date>. Each day your debt remains unpaid it increases. We currently charge interest at <GIC rate> a year, compounding daily, until the debt is paid in full.

If you can't pay
If you can't pay the total amount now contact us on < > between <8.00am> and <6.00pm>, <Monday to Friday> to see how we can help you.

Most people pay their tax on time and, in doing so, help pay for the essential services we all need and use. Thank you for your payment.

Yours <sincerely><faithfully>
<Deputy Commissioner's Name>
Deputy Commissioner of Taxation

PAY NOW

Your payment reference number (PRN) is: <PRN>

BPAY®

Biller code: < >

CREDIT CARD
Pay online with your credit card at **< >**
or phone **< >**.
A card payment fee applies.

For other payment options, visit **<ato.gov.au/howtopay>**

NEED HELP?

Visit us at **<ato.gov.au/contactus>**
Or
Contact us on **< >**
between <8:00am> and <6:00pm>, <Monday to Friday>.

Appendix B – List of Variables

Outcome variables

- Payment in full (Yes/No)

- Amount of tax debt payment (in AUS$)

Control variables

- Total business income

- Initial debt level

- Increase in outstanding debt

- Business market segment (Micro, small/medium, not for profit)

- Client type (Australian private company, individual/sole trader, other)

- Risk cluster (based on internal ATO analytics model)

- Lodgment channel (ATO online, auto finalised, agent portal, business portal, corporate data capture, electronic lodgment, tax agent portal, web services, other)

- Tax agent (yes/no)

- Number of employees

- Industry (5 digit ANZSIC code)

- State/Territory

Appendix C – Multiple Hypothesis Testing

TABLE C.1: HYPOTHESIS TESTING WITH MULTIPLE OUTCOMES, TREATMENTS AND SUBGROUPS

				p-values			
Group 1	Group 2	Subgroup	Mean Difference	Remark 3.1	Theorem 3.1	Bonferoni	Holm
Payment made by end of trial							
Control	Treatment 1	1	0.2886	0.0003	0.0003	0.0080	0.0066
Control	Treatment 2	1	0.2786	0.0003	0.0003	0.0080	0.0063
Control	Treatment 3	1	0.2815	0.0003	0.0003	0.0080	0.0076
Treatment 1	Treatment 2	1	0.0099	0.5063	0.9983	1	1
Treatment 1	Treatment 3	1	0.0070	0.6493	0.9990	1	1
Treatment 2	Treatment 3	1	0.0028	0.8476	0.9976	1	1
Control	Treatment 1	2	0.0249	0.3916	0.9956	1	1
Control	Treatment 2	2	0.0148	0.6346	0.9990	1	1
Control	Treatment 3	2	0.0275	0.4070	0.9956	1	1
Treatment 1	Treatment 2	2	0.0101	0.7066	0.9996	1	1
Treatment 1	Treatment 3	2	0.0524	0.0820	0.6706	1	1
Treatment 2	Treatment 3	2	0.0423	0.1823	0.9080	1	1
Amount paid by end of trial							
Control	Treatment 1	1	548.25	0.0003	0.0003	0.0080	0.0080
Control	Treatment 2	1	513.29	0.0003	0.0003	0.0080	0.0070
Control	Treatment 3	1	553.40	0.0003	0.0003	0.0080	0.0073
Treatment 1	Treatment 2	1	34.96	0.6446	0.9996	1	1
Treatment 1	Treatment 3	1	5.15	0.9480	0.9480	1	0.9480
Treatment 2	Treatment 3	1	40.11	0.6110	0.9990	1	1
Control	Treatment 1	2	1303.17	0.7903	0.9976	1	1
Control	Treatment 2	2	918.90	0.7480	0.9993	1	1
Control	Treatment 3	2	1793.00	0.6030	1	1	1
Treatment 1	Treatment 2	2	2222.07	0.6716	0.9976	1	1
Treatment 1	Treatment 3	2	489.83	0.9306	0.9976	1	1
Treatment 2	Treatment 3	2	2711.90	0.4283	0.9953	1	1

Note: See List et al. (2016) for a detailed description of the approach and the interpretation of mean differences and alternative p-values. Subgroup 1: Initial debt level < AUS\$7,500; Subgroup 2: Initial debt level AUS\$7,500+.