Nudging Businesses to Pay Their Taxes: Does Timing Matter?
Analysis Plan

Christian Gillitzer
The University of Sydney

Mathias Sinning[1]
Australian National University,
RWI and IZA

29 May, 2018

Abstract

This document is the *ex-ante* analysis plan for a trial conducted as part of the Payment Thinking (PT) project being run by the Debt Research and International Relations Team of the Australian Taxation Office (ATO), in conjunction with Australian National University and with the support of the Australian Research Council (ARC). This document is part of the trial registration that has been submitted to the RCT Registry of the American Economic Association (AEA). The document includes an overview, regression specifications and hypotheses tested in the analysis.

JEL Classification: H25, H26, H83
Keywords: tax compliance, reminder letter, timing

---

[1] All correspondence to Mathias Sinning, Tax and Transfer Policy Institute, Crawford School of Public Policy, JG Crawford Building #132, Lennox Crossing, The Australian National University, Canberra ACT 2601, T: +61 2 6125 1284, E: mathias.sinning@anu.edu.au.

Description of trial

The Australian Taxation Office (ATO) sends a reminder letter to business taxpayers who have missed their tax debt due date. This trial evaluated the effect of varying the send date of the reminder letter. The target population is restricted to taxpayers with a history of compliant payment behaviour over at least the previous three years. The trial was conducted based on the 26th May 2017 Business Activity Statement (BAS) lodgment date. 4,787 unpaid debt cases were quarantined from the usual ATO treatment pathways. Debt cases were randomly allocated to receive a reminder letter in either the first, second, or third week following their missed tax debt due date; a control group did not receive a letter for the duration of the trial. An equal number of quarantined debt cases were randomly assigned to each of the treatment groups and to the control group. The outcomes of interest are: (1) the number of days since the tax debt due date until payment is received, including information on partial payments; and (2) details of any payment arrangements clients may have entered into during the trial.

Details of the trial

Timing of events

- Tax debt due date: 26 May, 2017

- Baseline data received and cases randomly assigned to groups: 29 May to 2 June, 2017

- Letters sent to first group: 7 June, 2017

- Letters sent to second group: 14 June, 2017

- Letters sent to third group: 22 June, 2017

- Trial End Date: 14 July 2017

Interventions

The trial was conducted based on the 26th May 2017 Business Activity Statement (BAS) lodgment date. Cases with overdue tax debts were randomly allocated to receive a reminder letter in either the first, second, or third week following their missed tax debt due date (Appendix A contains a template for the letter); a control group did not receive a letter for the duration of the trial. All cases were quarantined from the ATO's usual treatment pathways until 14 July.

Randomization

Debt cases were randomly allocated to receive a reminder letter in either the first, second, or third week following their missed tax debt due date; a control group did not receive a

letter for the duration of the trial. The first step of the randomisation procedure involved grouping the tax debt cases into strata with similar baseline characteristics. Within each strata, each case was assigned at random to one of the treatment groups or to the control group. This stratification procedure ensures that the mix of tax debt debt case baseline characteristics is similar across the treatment and control groups. See Appendix B for a list of stratification variables. Randomisation was based on a random variable generator in `STATA`, using a random choice of the underlying seed, and the user-written `STATA` command `randtreat` version 1.4 (5 April 2017), available from the `ssc` library.

Econometric analysis

Our analysis is based on four sub-samples: (W1) cases chosen at the start of the trial to be sent a reminder letter in week 1 after their missed tax-debt due date, (W2) cases chosen at the start of the trial to be sent a reminder letter in week 2 after the missed tax-debt due date, (W3) cases chosen at the start of the trial to be sent a reminder letter in week 3 after the missed tax-debt due date, and (C) a control group not receiving a reminder letter. We are primarily interested in comparing outcomes for cases in groups W1, W2, and W3 to the cases in group C. We will also make pairwise comparison of outcomes across the treatment groups: W1 vs. W2, W1 vs. W3, and W2 vs. W3.

The general strategy for studying the effect of a treatment on an outcome measure of interest will be to estimate the following model:

$$Y_{i,t} = \alpha_t + \beta_{j,t} T_{i,j} + \varepsilon_{i,t} \tag{1}$$

where,

- $Y_{i,t}$ is a given outcome for taxpayer i at t-days since the tax debt due date;

- $T_{i,j}$ is an indicator for whether taxpayer i was allocated at the start of the trial to treatment group j, where $j \in \{Week1, Week2, Week3, Control\}$;

- $\beta_{i,j}$ is an (unconditional) average treatment effect for treatment j at t-days since the tax debt due date;

- $\varepsilon_{i,t}$ is a model error term.

The estimator used depends on the form of the outcome variable:

- Continuous dependent variables: ordinary least-squares (OLS) regression;

- Binary dependent variables: OLS and Probit models;

- Failure time data: Kaplan-Meier non-parametric estimator and Cox proportional hazard model.

We will also estimate specifications of Equation 1 including control variables:

$$Y_{i,t} = \alpha_t + \beta_{j,t}T_{i,j} + \gamma_{\mathbf{t}}\mathbf{X_i} + \varepsilon_{i,t} \tag{2}$$

where,

- $\mathbf{X_i}$ is the vector of control variables for taxpayer i. We are primarily interested in the set of variables used for stratification; see Appendix B for a list of these variables.

- $\beta_{i,t}$ is a conditional average treatment effect.

Debt outstanding will be observed daily and we will use the following outcome measures in our analysis:

- Continuous:

  - Amount(s) of payments;
  - Amount of debt still owed at conclusion of trial;
  - Payment proportion: proportion of debt paid vs. owed;
  - Amount of payment arrangement entered into.

- Binary:

  - Indicator for payment in full;
  - Indicator for payment arrangement entered into;

- Failure time data:

  - Indicator for payment in full;
  - Indicator for receipt of any payment;
  - Indicator for payment arrangement entered into.

Hypotheses

We are mainly interested in testing the statistical significance of the estimated coefficients $\hat{\beta}_{j,t}$. Because we do not know the sign of the estimated coefficients, we will use the following null hypotheses (H_0) and corresponding alternative hypotheses (H_1) to perform one-sided and two-sided t-tests at conventional levels of significance (1%, 5% and 10%):

1. $H_0 : \beta_{j,t} = 0$, $H_1 : \beta_{j,t} \neq 0$;

2. $H_0 : \beta_{j,t} = 0$, $H_1 : \beta_{j,t} > 0$;

3. $H_0 : \beta_{j,t} = 0$, $H_1 : \beta_{j,t} < 0$.

A Treatment group letter

Figure A1: Reminder 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>.

B List of control variables

The following control variables were supplied by the ATO. (S) denotes a variable used for stratification.

- Total business income (S);

- Initial debt level (S);

- Increase in outstanding debt;

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

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

- Risk cluster (based on internal ATO analytics model) (S);

- 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) (S);

- Number of employees Industry (5 digit ANZSIC code) (S);

- State/Territory;

- Industry type (ANZSIC code) (S).

C   STATA randomization code

```stata
1   *  ----------------------------------------------------------------
2   * Perform stratified random assignment for Blue Letter trial
3   *  ----------------------------------------------------------------
4
5   clear
6   cls
7   set more off
8   capture restore
9   set matsize 4000
10
11  *  ----------------------------------------------------------------
12  * Load and save dataset
13  *  ----------------------------------------------------------------
14
15  import excel using "${SOURCE_DATA}SoftBluePilot_DataForRandomisation_NoClientIDs_20170601.xlsx", sheet("Sheet1") cellrange(
    "a1:x6875") firstrow clear
16  save "${DERIVED_DATA}BLT_baseline", replace
17  use "${DERIVED_DATA}BLT_baseline", clear
18
19  *  ----------------------------------------------------------------
20  * Recode covariates: group where necessary to avoid small strata
21  *  ----------------------------------------------------------------
22
23  * Recode debt level variable
24  replace DebtLvl="Debt Level >= 3" if (DebtLvl=="Debt Level 3") |  (DebtLvl=="Debt Level 4") |  ///
25    (DebtLvl=="Debt Level 5") |  (DebtLvl=="Debt Level 6")
26
27  * Recode number of employees into bins
28  g NumofEmployees_cat=1 if !missing(Num_of_Employees) & Num_of_Employees==1
29  replace NumofEmployees_cat=2 if Num_of_Employees>=2 & Num_of_Employees<=5
30  replace NumofEmployees_cat=3 if Num_of_Employees>=6 & Num_of_Employees<=20
31  replace NumofEmployees_cat=4 if Num_of_Employees>=21 & !missing(Num_of_Employees)
32  label define NumofEmployees_cat_lab 1 "[1] 1" 2 "[2] 2-5" 3 "[3] 6-20" 4 "[4] >20"
33  label value NumofEmployees_cat NumofEmployees_cat_lab
34
35  * Client type
36  replace ClntTyp="Other Trust" if (ClntTyp=="Corporate Unit Trust") |  ///
37    (ClntTyp=="Fixed Trust") | (ClntTyp=="Fixed Unit Trust") | (ClntTyp=="Hybrid Trust")
38  replace ClntTyp="Other" if (ClntTyp=="APRA RegSupFnd SmallAPRA") |  ///
39    (ClntTyp=="Australian Public Company") | (ClntTyp=="Cwealth Govt Company") | ///
40    (ClntTyp=="Deceased Estate") | (ClntTyp=="Discretionary Trust Serv Mgt") | ///
41    (ClntTyp=="Other Unincorporated Entity") | (ClntTyp=="Discretionary Trust Invest.") | ///
42    (ClntTyp=="Other Incorporated Entity")
43
44  * Generate 2-digit ANZSIC codes
45  g ANZSIC_2digit=floor(ANZSIC/1000)
46
47  * Generate grouped ANZSIC codes
48  g ANZSIC_1digit=floor(ANZSIC/10000)
49  g ANZSIC_coarse=1 if ANZSIC_1digit<=2
50  replace ANZSIC_coarse=2 if ANZSIC_1digit==3
```

```stata
51   replace ANZSIC_coarse=3 if ANZSIC_1digit>=4 & ANZSIC_1digit<=5
52   replace ANZSIC_coarse=4 if ANZSIC_1digit==6
53   replace ANZSIC_coarse=5 if ANZSIC_1digit>=7 & ANZSIC_1digit<=9
54   label define ANZSIC_coarse_lab 1 "[1] 0-2" 2 "[2] 3" 3 "[3] 4-5" 4 "[4] 6" 5 "[5] 7-9"
55   label value ANZSIC_coarse ANZSIC_coarse_lab
56
57   * Construct quartiles of Effective TBI
58   xtile EffectiveTBI_quartile=Effective_TBI, n(4)
59
60   * Business market segment
61   g BusinessMarketSegment_cat="MIC" if Business_Market_Segment=="MIC"
62   replace BusinessMarketSegment_cat="SME" if Business_Market_Segment=="SME"
63   replace BusinessMarketSegment_cat="Other" if (Business_Market_Segment=="GOV") | ///
64   (Business_Market_Segment=="INB") | (Business_Market_Segment=="LGE") | ///
65   (Business_Market_Segment=="NFP")
66
67   * Indictaor for tax agent (based on whether tax agent number supplied)
68   g TaxAgent=0
69   replace TaxAgent=1 if !missing(Agent_no)
70
71   * ----------------------------------------------------------------
72   * Recode missing values: permits stratification
73   * ----------------------------------------------------------------
74
75   * Recode missing values: N/A for string variables and -99 for numeric variables
76   local RecodeList DebtLvl Cluster_num NumofEmployees_cat ClntTyp ANZSIC_coarse EffectiveTBI_quartile BusinessMarketSegment_cat
     TaxAgent
77   foreach var in `RecodeList' {
78       capture confirm numeric variable `var'
79
80       if _rc==0 {
81
82           quietly tab `var' if `var'==-99
83           if r(N)==0 {
84               disp "Variable `var': missing values recoded to -99"
85               replace `var' = -99 if missing(`var')
86           }
87           else disp "Variable `var': no replace, -99 already present"
88       }
89
90       if _rc==7 {
91
92           quietly tab `var' if `var'=="N/A"
93           if r(N)==0 {
94               disp "Variable `var': missing values recoded to N/A"
95               replace `var' = "N/A" if missing(`var')
96           }
97           else disp "Variable `var': no replace, N/A already present"
98       }
99   }
100
```

```stata
101  * ----------------------------------------------------------
102  * Encode string variables
103  * ----------------------------------------------------------
104
105  foreach var of varlist _all {
106      capture confirm numeric variable `var'
107
108      if _rc==0 {
109          disp "`var' already numeric"
110      }
111
112      if _rc==7 {
113
114          disp "`var' encoded from string to numeric"
115          encode `var', gen(`var'_num)
116          drop `var'
117          rename `var'_num `var'
118      }
119
120  }
121  * ----------------------------------------------------------
122  * Randomisation
123  * ----------------------------------------------------------
124
125  * Stratification variables
126  local strata DebtLvl Cluster_num NumofEmployees_cat ClntTyp ANZSIC_coarse EffectiveTBI_quartile BusinessMarketSegment_cat TaxAgent
127
128  * Generate strata
129  egen strata = group(`strata')
130
131  * Look at size of strata
132  sort strata
133  by strata: gen N = _N
134  tab N
135
136  * Use randtreat v1.4 (5 April 2017)
137  randtreat, replace strata(`strata') generate(treatment_optionA) setseed(987654321) unequal(1/4 1/4 1/4 1/4) misfits(wstrata)
138  randtreat, replace strata(`strata') generate(treatment_optionB) setseed(987654321) unequal(4/16 3/16 4/16 5/16) misfits(wstrata)
139  g treatment=treatment_optionA
140
141  * Check balance
142  tab treatment
143  *tab strata treatment
144
145  * Change treatment group names
146  gen group = "."
147  replace group = "No letter" if treatment==0
148  replace group = "Week 1" if treatment==1
149  replace group = "Week 2" if treatment==2
150  replace group = "Week 3" if treatment==3
151
```

```stata
152   * Test stratification
153   forvalues i=1/3 {
154       T`i'=treatment==`i'
155       g treatment`i'=0 if treatment==0
156       replace treatment`i'=1 if treatment==`i'
157   }
158   foreach var in `strata' {
159       reg `var' T1
160       reg `var' T2
161       reg `var' T3
162       reg `var' treatment1
163       reg `var' treatment2
164       reg `var' treatment3
165   }
166   tab treatment_optionA
167   tab treatment_optionB
168
169   * Export treatment assignment to csv file
170   sort rcrd_num group
171   outsheet rcrd_num group using "${DERIVED_DATA}BLT_treatmentGroups.csv", comma replace noquote
172   exit
173
```