ARTICLE

Behavioural prompts to increase early filing of tax returns: a population-level randomised controlled trial of 11.2 million taxpayers in Indonesia

Ruth Persian1*, Gitarani Prastuti2, Adityawarman2, Daniel Bogiatzis-Gibbons2, Muhammad Hakim Kurniawan2, Gatot Subroto2, Muhammad Mustakim2, Laurenz Scheunemann3, Kizzy Gandy4 and Alex Sutherland1

1Behavioural Insights Team, London, UK, 2Directorate General of Taxes, DKI Jakarta, Indonesia, 3Independent Scholar, Germany and 4Independent Scholar, Sydney, Australia

*Corresponding author: Ruth Persian, email: ruth.persian@bi.team

(Received 4 February 2022; revised 12 June 2022; accepted 16 June 2022)

Abstract

In Indonesia, as in other countries, a large proportion of tax returns are filed at the last minute. In a population-wide randomised controlled trial (n = 11,157,069), we evaluated the impact of behavioural email prompts on the proportion of annual tax returns filed at least two weeks before the deadline; and overall filing rate. In two control conditions, taxpayers either received no email or an email used in prior years, emphasising regulatory information. The five treatments informed by behavioural science were (1) a simplified version of the existing email, emphasising early filing; (2) the simplified version with additional guidance on filing taxes; (3) the simplified version with a planning prompt and option to sign up for email reminders; (4) a version combining treatments 1, 2 and 3; and (5) an email appealing to national pride. Compared to the no-email control, all emails led to a statistically significant increase in early and overall filing rates. The planning email (3) was the most effective, increasing early filing from 34.9% to 37% (b 2.07 percentage points (pp), p < 0.001, 95% confidence interval (CI) 1.97–2.17pp), and overall filing from 65.6% to 66.7% (b 1.10pp, p < 0.001, 95% CI 0.99–1.19pp).

Keywords: RCT; tax filing; planning prompts; Indonesia

Introduction

In Indonesia, as in other countries such as the UK and US, a large proportion of tax returns are filed at the last minute or late. For example, in 2017, only 35% of employee taxpayers and 2% of self-employed taxpayers filed a tax return at least two weeks before the deadline.1 For governments, which run on tight budget cycles and rely on timely revenue collection, taxpayer procrastination related to filing or payment

1Data from the Directorate General of Taxes’ database.
© The Author(s), 2022. Published by Cambridge University Press
has significant costs. Furthermore, processing a last-minute rush of annual tax returns can increase administrative costs, for instance, because of having to hire extra staff or upgrade IT infrastructure.

From taxpayers’ point of view, bottlenecks caused by last-minute filing could result in lower tax morale as they get frustrated with overloaded and crashed systems. This is a real risk: In 2017, last-minute demands on the Indonesian tax filing system led to the filing deadline being extended because the system overloaded (Amianti, 2017). Compliance might also decrease if taxpayers fail to file their taxes, either online or in person, when too many people attempt to do so at the same time. In addition, tax returns submitted in haste may be more prone to being incomplete or contain errors (Slemrod et al., 1997). In the medium to long term, citizens’ tax morale and thus compliance might deteriorate further due to bad experience with last-minute filing pressures and rejections due to incomplete fields or errors, as well as the perception that the tax burden is not shared equally (if they are owed a refund). Spreading filing dates over a longer period of time would therefore be in the interest of both governments and taxpayers.

Low-cost communication interventions, such as emails, that leverage insights from the behavioural sciences have been shown to be effective at improving taxpayer behaviour across a range of contexts (Dalton et al., 2021). Even when the individual-level impact is modest, the aggregate effect in a country as large as Indonesia with close to 50 million (Karnadi, 2022) registered taxpayers (a number that is rising) can be substantial. The sample size available and the fact that increasing the sample is costless allow us to test a number of variations of an intervention with a high degree of precision.

In one of the largest, population-wide randomised controlled trials (RCTs) ever conducted by a government (n = 11,157,069), we evaluated the impact of behavioural email prompts on the (i) proportion of annual tax returns filed at least two weeks before the deadline; and (ii) overall filing rate among taxpayers registered for online filing. In the two control conditions, taxpayers either received (i) no email or (ii) the email used in prior years, heavy on regulatory information. The five treatments informed by behavioural science were (1) a simplified version of the existing email, emphasising early filing; (2) the simplified version with additional guidance on filing taxes; (3) the simplified version with a planning prompt and option to sign up for email reminders; (4) a version combining treatments 2 and 3; and (5) an email appealing to national pride.

We were interested in testing which concept from the behavioural science literature, addressing barriers we found during exploratory research into the barriers to filing, would be most effective at encouraging the behaviours of interest; thereby testing whether previously tested concepts could be applied not just to increase tax compliance more narrowly, but also to affect other desirable taxpayer behaviours, such as moving away from last-minute filing.

Compared to the no-email control, all emails lead to a statistically significant increase in early and overall filing rates. The planning email (3) was the most effective, increasing early filing from 34.9% to 37% (point estimate 2.07 percentage points (pp), p < 0.001, 95% confidence interval (CI): 1.97–2.17pps), and overall filing from 65.6% to 66.7% (point estimate 1.10pp, p < 0.001, 95% CI: 0.99–1.19pp). We estimate...
that, if scaled to the entire population of taxpayers registered for online filing, this email would have led to 233,900 tax returns submitted at least two weeks before the deadline and 133,000 extra tax returns submitted overall.

In addition to adding to the literature on taxpayer behaviour, this study demonstrates that it is possible to run large-scale RCTs in collaboration with governments and that RCTs such as this one can deliver high-quality evidence for policy makers in a timely way.

The Indonesian tax system

Most Indonesian income taxpayers are required to file an annual tax return. This applies even if income tax is automatically paid during the tax year through a pay-as-you-earn scheme. Indeed most taxpayers do not have to pay additional taxes at the point of filing. Those earning below the annual net income threshold\(^2\) for paying income tax of IDR 54 m/year (approx. USD 3700) or those employed can request an exemption from the Director General, which, if approved, would mean that they do not have to file a tax return. However, in reality, very few taxpayers request such an exemption.

The tax year in Indonesia runs from January to December, and personal taxpayers then have until 31 March of the following year to file a tax return. Returns can be filed by submitting them physically at one of approximately 556 local tax offices and service units, sending them by post or filing them online at the Directorate General of Taxes’ (DJP) website or through the online systems of vetted external service providers.

However, compliance among registered taxpayers is relatively low, especially among the self-employed: for the 2017 tax year, only 75% of registered employee taxpayers and 62% of self-employed taxpayers filed a return (Directorate General of Taxes, 2021). Low compliance might partly be explained by the fact that the penalty for not filing is low, at IDR 100,000 (USD 7), and the risk for a taxpayer to actually have to pay this fine is very small.

Related literature

Tax compliance includes four distinct behaviours: (1) registering as a taxpayer; (2) reporting one’s tax liability accurately; (3) paying the full amount of tax reported in a timely manner; (4) filing required returns on time (Forum on Tax Administration, 2004). The traditional economic model describes tax compliance as being determined by a cost-benefit analysis conducted by a rational actor (Allingham and Sandmo, 1972; Srinivasan, 1972); however, it has been criticised for not being able to explain the high levels of compliance observed in the data, given the small risk of being caught (Torgler, 2007). More recently, there has, therefore, been a move towards using behavioural economics and social psychology to explain (non-)compliance. This literature argues that the traditional model has to be extended to include factors such as taxpayer morale (that is, the intrinsic motivation to pay taxes), social considerations, perceived fairness of tax burden and satisfaction with government (Andreoni et al., 1998; Kirchler et al., 2007).

\(^2\)The annual net income is gross income minus exemptions/deductions (e.g. occupational expenses).
Equally, the behavioural science literature offers potential entry points for policy interventions that go beyond costly enforcement and comprehensive audits: Low-cost taxpayer communications which make compliance easy (through simplification and reminders), present compliance as a social norm or highlight the consequences of non-compliance have had success in more than 10 countries (Hallsworth, 2014; Dalton et al., 2021). Brockmeyer et al. (2019) use a message that highlights the availability of third-party information – that is, that makes the perceived risk of non-compliance more salient – to increase tax payments among businesses in Costa Rica. Kettle et al. (2016) show that a message presenting non-compliance as a deliberate choice, rather than negligence, is effective at increasing compliance among taxpayers in Guatemala. Simple reminders can mitigate the effect of inattention and forgetfulness and thereby increase compliance (Guyton et al., 2016; Gillitzer and Sinning, 2020). That said, communication might not just be ineffective; in some contexts, it might even backfire and decrease compliance (Ariel, 2012). While the existing literature can thus provide evidence on what might work, any intervention will therefore need to be evaluated as rigorously as possible before implementation at scale. This is especially the case for interventions developed for and implemented in a context with little previous evidence on what works, such as Indonesia.

What sets our study apart from large parts of the tax compliance literature is that the primary target behaviour does not directly fall under one of the four components of compliance: failure to file early, that is, sometime before the deadline does not make a taxpayer non-compliant, as long as the return has been completed by the deadline. That said, non-compliance is a concern for the Indonesian government and increasing it was a secondary objective of this study.

The behavioural science literature suggests that last-minute tax filing could be associated with the human tendency to procrastinate when a task is difficult or unpleasant (Benzarti, 2015; Martinez et al., 2017; Chirico et al., 2018). Early filing might be ultimately in the interest of the taxpayer by avoiding an overburdened physical and online systems at a later date but comes with the immediate cost of the effort required for filing. Procrastination of filing can have different reasons: Firstly, there is evidence from the United States that the complexity of tax forms – or added friction – delays filing, even for returns which will result in a refund (Slemrod et al., 1997). Secondly, Martinez et al. (2017) demonstrate the link between procrastination of filing and present bias, that is, the tendency to overvalue immediate rewards (avoiding the effort of filing now) relative to long-term consequences (O'Donoghue and Rabin, 1999).

To understand the drivers of last-minute filing in Indonesia specifically, we spoke to taxpayers, some of whom filed their 2016 tax return near the deadline, and frontline tax officials in several local tax offices. This exploratory research supported the hypothesis that procrastination was highly relevant for explaining last-minute filing and exacerbated by friction such as confusion over the rules, forms and deadlines to file an annual tax return in Indonesia. This was especially the case for self-employed taxpayers, who are required to complete a much longer and more complex tax return form than employee taxpayers and need to supply several supporting documents.

This study contributes to the tax literature by expanding the application of communication interventions based on insights from the behavioural science literature.
from increasing compliance in a more narrowly defined sense to the challenge of increasing early tax returns to spread the burden on systems. While a number of experiments have used behavioural science-based interventions to increase on-time and accurate tax declarations and payment, to the best of our knowledge no previously published study has tested these types of communication interventions to specifically encourage early tax returns to smooth tax administration.

**Intervention messages**

For this study, we developed and tested five different email treatments to address the reasons behind taxpayer procrastination, against two control interventions. For some treatments, we combined different behavioural insights concepts to maximise the likely effect size. While this means that we cannot disentangle the separate impact of different behavioural science concepts, we made this decision as the primary interest of the study was to find and implement the most effective approach, rather than test specific mechanisms.

Several of our tested concepts (simplification, social norms, reminders, messenger effect, national pride, and salience of negative consequences) have been previously tested to directly increase short-term tax compliance (see above). We also tested additional concepts which are new in the tax administration literature but have been tested in other policy areas: planning, enhanced active choice and implementation intentions. Table 1 summarises the seven treatment arms, which are presented in more detail in this section. The full emails can be found in Supplementary Appendix A.

All emails addressed the taxpayer by their name. This simple way of personalising a message has been shown to increase the impact of government communications (Haynes et al., 2013) and was easy to implement as DJP already held taxpayers’ names in the dataset we used for randomising and implementing the intervention. To strengthen the credibility and persuasiveness of the emails through a messenger effect (Dolan et al., 2012), they were also signed by the Director General of DJP.

**Control 2 – Status quo email**

This email was very similar to an email previously used to remind taxpayers of their filing obligations. It opened with the regulations on taxpayers’ filing obligations, outlined the options that taxpayers had for filing and where they could obtain additional information. Compared to no email at all, it served as a reminder of taxpayers’ obligation to file and aimed to overcome forgetfulness and inattention.

**Treatment 1 – Simplification**

The status quo email opened with a lengthy description of the legal framework for tax filing – information that is ultimately of little relevance for the individual taxpayer. Evidence shows that, on average, people spend not more than 10–13 s reading certain types of branded emails (Ceci, 2021). Even if emails from a government source may be given more attention by recipients, a long message would reduce the chances people engage with the full content and see or derive the call to action. Previous research
shows that simplifying the message and including a clear call-to-action can indeed increase tax compliance (see, for example, Kettle et al., 2016). Our email simplified the content and focused on the main information – that it is time to file the tax return and what the options for doing so are. The word count of the main body of the email was thereby reduced from 198 in the status quo version to 93.

It also added a weak dynamic social norm by telling the recipient that more Indonesians than ever are filing a return.

**Treatment 2 – Link to guidance**

Our interviews with taxpayers suggested that many taxpayers were put off by the complexity of completing a return, especially where additional documentation was required or more complex forms had to be completed. We also heard that, after queuing for a long time to submit their form to a tax official, taxpayers would sometimes get to the front of the queue only to learn they were missing certain required information or attachments. The same happened for online filing, where the system would not save the information and taxpayers would have to restart the filing process once they had assembled all the documents. Having to return at a later date might reduce tax compliance in the short term if there is not enough time to assemble all the required documentation or lead to more returns being filed just before the deadline; and in the long term if tax morale is negatively affected.

The existing taxpayer guidance published by DJP was lengthy and used complex language. For example, the guide for form 1770 (the form used by self-employed
taxpayers) was close to 70 pages long. In this email, we therefore applied one of the core principles of behavioural science-based interventions: making the target behaviour as easy as possible (Harford, 2019). We linked to a checklist outlining the documents required for the different types of tax forms and a two-page FAQ document on form 1770, to remove the friction of having to look for the required information after having started the filing process and potentially having to restart the filing process at a later date.

In addition to stating the same weak dynamic social norm as in Treatment 1, the email addressed procrastination due to present bias by making the potential cost of procrastination (queues and a slow online system closer to the due date) salient. Finally, we aimed to reduce the perceived effort cost of filing by stating that it takes only 15 min to file a return if all required documents have been assembled in advance.

**Treatment 3 – Planning prompt & reminders**

This email built on research that asking people to plan in advance when and how they will complete an action makes them more likely to follow through on their intentions (Gollwitzer 1999; Milkman et al., 2011, 2012) and that reminders help increase action by addressing forgetfulness.

The email also nudged the recipient to make an enhanced active choice (Keller et al., 2011) by asking them to choose either ‘YES; I will choose a suitable date to report the SPT (the tax return, note by authors) before 16 March 2018 to make it easier’ or ‘NO; I will not choose a date to report the SPT before 16 March, even if this makes it more difficult for me’. Enhanced active choice messages encourage the desired behaviour (in this case, choosing ‘yes’) by triggering the recipient’s desire to avoid future regret from not choosing the option suggested by the sender (Keller et al., 2011). The ‘yes’ option took the taxpayer to a website where they could choose a date for filing their return. Those picking a date were sent reminder emails two days before, and on, the selected date, advising them to prepare and visualise the steps they needed to take to file their tax return.

As before, the email included the weak social norm from Treatment 1 and made the potential cost of procrastination salient.

**Treatment 4 – Combined guidance & planning prompt**

This email was a combination of the guidance and planning emails (Treatments 2 and 3), in order to test whether integrating versions we believed to increase filing would strengthen or weaken the effect. To reduce the length of the email, we removed the list of problems taxpayers could avoid by filing early.

**Treatment 5 – National pride**

The evidence on the effectiveness of invoking national pride is mixed: there are examples where such messaging has been used successfully to increase tax compliance in the lab (Gangl et al., 2016; Macintyre et al., 2021); by contrast, in an experiment in Guatemala, the appeal to the taxpayers’ national identity was less effective at
increasing tax compliance than a social norms message and one framing non-compliance as a deliberate choice (Kettle et al., 2016). On the other hand, transparency about how tax revenue is spent can positively influence tax compliance (Djawadi and Fahr, 2013).

This email addressed taxpayers as heroes for building the nation and contained an image illustrating how government revenue would be spent under the 2018 national budget. Unlike the other four behavioural science-based emails, the structure of this message was more consistent with conventional Indonesian communication, which is often less direct than English. For example, it opened with an explanation of why the email was sent.

**Trial design**

**Sample**
The study included the universe of Indonesian individual income taxpayers registered for online filing by December 2017, excluding expatriates paying taxes in Indonesia (n = 11,157,069, 30% of all registered taxpayers). While online filers may not be representative of all Indonesian personal income taxpayers, they make up a substantial proportion of the taxpayer base and can be targeted with email interventions. Furthermore, their importance as a share of all Indonesian taxpayers is likely to increase with the increased digitalisation of tax administration. For example, between 2015 and 2017, online filing in Indonesia increased from 9.7 to 10.5 million taxpayers (Mediatama, 2017, 2018). The email addresses of taxpayers registered for online filing had furthermore been verified by DJP, making this group an easy target group for an email intervention.

**Randomisation**

This trial was randomised, and outcomes measured, at the level of the individual taxpayers, using their tax identification number (Nomor Pokok Wajib Pajak, NPWP) as the unique identifier. The participants were assigned equally via stratified randomisation to one of the seven treatment arms. We stratified randomisation by two characteristics we expected to be associated with the target behaviour of tax filing: (i) self-employed/employee status; and (ii) the regional tax office where the taxpayer is registered. We also confirmed the following characteristics were balanced across trial arms: age in decade bands; whether the taxpayer had registered within the 3 years before trial launch; and whether the taxpayer had filed a 2016 tax return (see Supplementary Appendix B). Taxpayers were randomised on site by one of the co-authors, using Stata code developed jointly by the study team.

**Implementation**

The emails were sent approximately five weeks before the filing deadline of 31 March 2018 to give taxpayers sufficient time to file before 16 March, while also ensuring that most employees would already have received a tax withholding slip from their employer which had to be attached to the tax return.

To account for the fact that DJP’s server could only send a few million emails/day, roll-out of the intervention was split over two days (19 and 26 February 2018). On the
first day, several implementation challenges were encountered, including a delay in delivery due to DJP’s server being temporarily blocked by a major email provider and a relatively small number of emails (compared to the overall sample size) not being sent out as intended (for more detail see Supplementary Appendix C). All issues were resolved before roll-out on Day 2 and did not have an impact on results (see below). Due to the high volume of emails, it took several hours for all emails to leave DJP’s server on each of the two implementation days.

We carefully monitored social media and internet publications for the duration of the trial. We identified one instance of a DJP staff publishing an article about the trial, which was removed immediately.

Overall, we believe that few taxpayers will have spoken to others about the communication and that any spillovers would thus have been minimal: DJP routinely communicates with taxpayers through email (for example, the control email was based on the email that DJP had sent out in previous years), so receiving an email will not have been anything noteworthy. Furthermore, anecdotal evidence from speaking to Indonesian taxpayers who had been part of the sample indicated that they do not routinely speak to others about DJP communication.

We are not able to estimate/control for potential bias caused by spillovers, if it exists, but believe that it would overall lead to attenuation of the treatment effect as taxpayer behaviour across arms would have been more similar than without spillovers.

Outcome measures
The two primary outcome measures were: (1) whether 2017 returns were filed early (by 16 March 2018, two weeks before the deadline); and (2) whether 2017 returns were filed at all among the universe of eligible Indonesian personal income taxpayers registered for online filing. The date chosen as the early filing cut-off was chosen as our exploratory research suggested that virtually all employee taxpayers had received their tax withholding slip by then, but was still early enough to meaningfully contribute to smoothing out the distribution of filing dates.

We also analysed the impact of the emails on two additional, secondary outcomes: (1) the net amount of taxes paid (in IDR), and (2) the proportion of people who pay any amount of tax (relevant mostly for self-employed taxpayers, as their tax is not automatically withheld by an employer). Since the intervention targeted filing and most taxpayers in the sample do not have to pay additional taxes/receive a refund (as they are taxed through a pay-as-you-earn scheme), we did not expect any significant positive impact on these outcomes, but wanted to ensure that there was no negative impact on these outcomes either. Table 2 summarises the outcome measures as specified before trial launch.3

Finally, we explored the correlation between signing up to the commitment website (Treatments 3 and 4 only) and early filing.

3 All analysis was pre-specified in a non-published research protocol. We had also planned to plot ‘survival curves’ against time for each treatment arm. We decided instead to plot filing dates by treatment group. However, the differences were too small to be visible and have, therefore, been excluded from this paper. Instead, we present exploratory analysis looking at filing during the last week and day before the deadline.
Our analysis is based on intention-to-treat (ITT) estimates (i.e. the impact of sending emails), rather than treatment-on-the-treated (TOT) estimates (i.e. the effect of opening, reading and acting on them). Emails are not received by all taxpayers and those that received emails might not open them, so it is reasonable to expect a difference between the two estimates. However, the ITT, rather than the TOT, reflects the impact of practical interest as emails would always be sent to all taxpayers. Emails were sent with requests for ‘read receipts’ so that we could track whether emails had been opened. By mid-April 2018 we recorded that between 71% and 84% of emails had been opened across different treatment conditions (see Supplementary Appendix D). However, many email programmes require recipients to consent to read receipts being sent, so these numbers are lower bounds.

Our primary analysis consisted of covariate-adjusted linear probability models with robust standard errors to account for possible heteroskedasticity, comparing each treatment with the control group. The outcomes were binary variables for if the individual filed (1) early (yes/no) or (2) at all (yes/no). The covariates adjusted for were whether the email was sent on Day 2 (rather than Day 1, binary), n-1 dummy variables indicating the regional tax office the recipient was registered at, whether the taxpayer had registered within the last 3 years or not (binary), whether the taxpayer filed a tax return for the 2016 tax year or not, and taxpayer age in decade bands.

We were interested in ranking the seven treatment arms (controls and treatments) against each other. This makes a total of 21 comparisons per outcome, for a total of 42 comparisons across the two primary outcomes. We, therefore, employed the Benjamini–Hochberg step-up procedure, which reduces the Family-Wise Error Rate by setting the false discovery rate (the probability of making at least one Type I error) at 0.05 (Benjamini and Hochberg, 1995).

Secondary analysis was conducted using the same specification on the total amount of tax collected (in IDR) and the payment rate, although we did not adjust

### Table 2. Outcome measures

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Cut-off date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary 1: Filing a tax return by 16 March 2018 (i.e. 2 weeks before the deadline).</td>
<td>16 March 2018</td>
</tr>
<tr>
<td>Primary 2: Filing a tax return at all.</td>
<td>4 November 2018</td>
</tr>
<tr>
<td>Secondary 1: Amount of tax collected from annual tax returns (in IDR)</td>
<td>4 November 2018</td>
</tr>
<tr>
<td>Secondary 2: Paying any tax of any amount. This is most relevant to self-employed taxpayers whose tax is not automatically withheld by an employer.</td>
<td>4 November 2018</td>
</tr>
<tr>
<td>Secondary 3 (non-experimental): correlation between signing up to commitment website and early filing (Treatments 3 and 4).</td>
<td>4 November 2018</td>
</tr>
</tbody>
</table>

**Statistical analysis**

Our analysis is based on intention-to-treat (ITT) estimates (i.e. the impact of sending emails), rather than treatment-on-the-treated (TOT) estimates (i.e. the effect of opening, reading and acting on them). Emails are not received by all taxpayers and those that received emails might not open them, so it is reasonable to expect a difference between the two estimates. However, the ITT, rather than the TOT, reflects the impact of practical interest as emails would always be sent to all taxpayers. Emails were sent with requests for ‘read receipts’ so that we could track whether emails had been opened. By mid-April 2018 we recorded that between 71% and 84% of emails had been opened across different treatment conditions (see Supplementary Appendix D). However, many email programmes require recipients to consent to read receipts being sent, so these numbers are lower bounds.

Our primary analysis consisted of covariate-adjusted linear probability models with robust standard errors to account for possible heteroskedasticity, comparing each treatment with the control group. The outcomes were binary variables for if the individual filed (1) early (yes/no) or (2) at all (yes/no). The covariates adjusted for were whether the email was sent on Day 2 (rather than Day 1, binary), n-1 dummy variables indicating the regional tax office the recipient was registered at, whether the taxpayer had registered within the last 3 years or not (binary), whether the taxpayer filed a tax return for the 2016 tax year or not, and taxpayer age in decade bands.

We were interested in ranking the seven treatment arms (controls and treatments) against each other. This makes a total of 21 comparisons per outcome, for a total of 42 comparisons across the two primary outcomes. We, therefore, employed the Benjamini–Hochberg step-up procedure, which reduces the Family-Wise Error Rate by setting the false discovery rate (the probability of making at least one Type I error) at 0.05 (Benjamini and Hochberg, 1995).

Secondary analysis was conducted using the same specification on the total amount of tax collected (in IDR) and the payment rate, although we did not adjust
this for multiple comparisons as it was not being used to draw direct policy conclusions.

Finally, given possible differences in response for the self-employed we tested for treatment effect heterogeneity by self-employment status and, given the cultural diversity across Indonesia, we checked whether the treatment effects varied by regional tax office by re-running the regressions in each of the 33 tax offices (with a specific focus on potential backfires).

Details of the regressions run are presented in Supplementary Appendix E, and the Stata code is available from the authors upon request. The analysis was run by members of the study team on DJP computers. Due to Indonesian privacy and tax law, it is not possible to provide the raw data.

Data

Data to create outcome and covariate variables were obtained from DJP’s IT systems (for more detail, see Supplementary Appendix E). The baseline, randomisation and tax return data were merged by one of the co-authors using Knime. We used Stata to merge this with the information on sign-ups to the commitment website (Treatments 3 and 4 only), using the taxpayer ID number as a unique identifier for individuals. We based our analysis on the final data extracted in early November 2018. This means that returns filed after this date were not captured. However, only a very small number of returns are filed after this date and this is, therefore, not of any practical concern for the interpretation of our results. We handled duplicate taxpayer ID numbers in the outcome dataset (i.e. cases where individuals had submitted an amendment of their original tax return) by keeping the date of the first submission as filing date. To determine the total taxes paid/claimed at the point of submission, we calculated the net amount across submissions.

Results

Primary results

When comparing raw means, relative to the no-email control, all emails led to an increase in early and overall filing rates (Table 3). The increases were more pronounced for filing early than for filing overall: For the former, differences ranged from 0.8 to 2.1pp, or 2.2–6% in relative terms. Overall filing rates increased by 0.5 to 1.1pp, equivalent to a 0.7–1.7% relative increase compared to the no-email control.

Table 4 provides a regression results from the primary analysis, Figures 1 and 2 the visualisation. We find that, compared to the no-email control and for both outcomes, all treatments produce a statistically significant increase in the respective outcome variable, suggesting that the pure reminder effect of an email impacts filing behaviour. The point estimate is highest for the planning prompt & reminders treatment (Treatment 3): it increased the likelihood of filing early by 2.07pp (Column 1, p < 0.001, 95% CI: 1.97–2.17pp), and that of filing overall by 1.10pp (Column 2, p < 0.001, 95% CI: 0.99–1.19pp). It is lowest for the national pride treatment (Treatment 5) in both regressions, with an increase of 0.77 and 0.53pp compared to the no-email control, respectively.
In a pairwise comparison, Treatment 3 performed not only statistically significantly better in encouraging early filing than the no-email control, but also when compared to all other emails (see Appendix E for detail). When considering overall filing, it performed better than all emails except for the Treatment 2 (link to guidance), with which the difference was not statistically significant.

These seemingly small effects could substantially decrease the burden on tax offices and DJP’s IT system: if the planning prompt (Treatment 3) had been sent to the whole sample of Indonesian taxpayers registered for online filing, it would have translated to an additional 133,000 tax returns filed overall, and an additional 233,900 returns submitted two weeks or more before the deadline. The planning prompt website incurred an initial set-up cost (in terms of human resources used). However, we believe that this cost is justified – especially for scale-up, given that the cost for updating the site is small – given the significantly better performance of Treatment 3 for the main outcome of interest.

Secondary analysis

Our intervention had no impact on the amount of (net) taxes paid at point of filing (for detail, see Table E.3, Supplementary Appendix E). We believe that this is at least in part driven by the fact that few taxpayers had to make an additional payment at the point of filing, as employed taxpayers generally have their income taxes withheld and paid by their employer.

We also looked at the correlation between signing up to the commitment website and the outcomes, for those individuals in Treatments 3 and 4 (see Table E.4, Supplementary Appendix E). While these results cannot be interpreted causally, we found that signing up to the planning website was strongly associated with a higher likelihood of filing early (a 14.16 percentage point higher likelihood for Treatment 3, and 12.23 percentage point higher likelihood for Treatment 4). We interpret this, cautiously, as the planning website having played an important role in the overall impact of Treatments 3 and 4, while recognising that an alternative

<table>
<thead>
<tr>
<th>Email treatment</th>
<th>N</th>
<th>Filed early (%)</th>
<th>Filed overall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control 1 – No email</td>
<td>1,593,840</td>
<td>34.9</td>
<td>65.6</td>
</tr>
<tr>
<td>Control 2 – Status quo email</td>
<td>1,593,867</td>
<td>35.7</td>
<td>66.1</td>
</tr>
<tr>
<td>Treatment 1 – Simplification</td>
<td>1,593,868</td>
<td>36.1</td>
<td>66.4</td>
</tr>
<tr>
<td>Treatment 2 – Link to guidance</td>
<td>1,593,863</td>
<td>36.2</td>
<td>66.6</td>
</tr>
<tr>
<td>Treatment 3 – Planning prompt &amp; reminders</td>
<td>1,593,869</td>
<td>37.0</td>
<td>66.7</td>
</tr>
<tr>
<td>Treatment 4 – Combined link to guidance &amp; planning prompt</td>
<td>1,593,867</td>
<td>36.6</td>
<td>66.4</td>
</tr>
<tr>
<td>Treatment 5 – National pride</td>
<td>1,593,895</td>
<td>35.7</td>
<td>66.2</td>
</tr>
<tr>
<td>Total (Column 1)/Average (Column 2 &amp; 3)</td>
<td>11,157,069</td>
<td>36.0</td>
<td>66.3</td>
</tr>
</tbody>
</table>
interpretation is that those taxpayers who were more likely to sign up to the website (due to factors such as motivation) might also have been more likely to file early without the website.

### Subgroup analysis

We first compared treatment effects on the primary outcome measures for self-employed to the effect for employed taxpayers. While Treatment 3 is less effective
Figure 1. Proportion of taxpayers who file tax return early.

Figure 2. Proportion of taxpayers who file tax return.
among self-employed taxpayers, this decline in effectiveness is not large enough to wipe out the impact completely (for details see Table E.5 Supplementary Appendix E).

An important concern for governments of countries as culturally diverse as Indonesia is that interventions that use a ‘one-size-fits-all’ approach might be effective on average, but lead to backfires in certain parts of the country. When looking at outcomes by regional tax office, for the early filing outcome only, we find that reassuringly, there appeared to be no backfire effects by region: there are only three regions with some negative point estimates, and none are statistically significant at conventional levels. Furthermore, the impact of Treatment 3 compared to the pure control was statistically significant (at the 5% level) and positive for all regions.

Robustness checks

In order to assess the effect of the implementation issues described above during Day 1 of the roll-out, we conducted additional robustness checks running the primary analysis separately by day. Reassuringly, Treatments 3 and 4 remained the most effective across both days (Table E.7, Supplementary Appendix E).

Around 9% of taxpayers in the sample filed before the intervention and their filing behaviour would thus not have been affected by our treatment. Excluding individuals filing before the intervention increased the size of the coefficients, but did not change the results overall (see Table E.6).

Exploratory analysis

Given the encouraging and consistent results of the pre-specified analysis, we decided to conduct additional, exploratory analysis that had not been pre-specified. We found a 6.4% and 4.8% decrease in the number of taxpayers filing on the last day and during the last week before the deadline, respectively, between the pure control group and Treatment 3 (Table E.8, Supplementary Appendix E). Had the Treatment 3 email been sent to all taxpayers, we estimate that this would have brought forward 143,400 and 29,000 returns from the final week and day, respectively. This result is encouraging as it suggests that the email shifted the behaviour not only of people who had previously filed between one and two weeks before the deadline but also the ‘serious’ procrastinators. This could further unburden tax offices overwhelmed by filings close to the deadline.

We also looked at the impact of the treatments on the likelihood of filing on time (i.e. by 31 March) rather than early or overall filing. We see very similar results, in terms of the relative effectiveness of the different treatments (Table E.9, Supplementary Appendix E).4

Discussion

To our knowledge, this is the first published study experimentally evaluating the impact of low-cost communication on taxpayer behaviours not immediately linked

---

4There was no statistically significant difference in the impact of the different treatments on early filing between taxpayers who had filed a tax return for the 2016 tax year and those that hadn’t.
to compliance defined more narrowly (registration, honest declaration, on-time filing and payment). We show that emails can be used to encourage behaviours of interest to tax authorities and governments more generally, such as early filing, in addition to improving short-term compliance.

All versions of the email reminders, including the status quo email without a behavioural re-design, increased the share of taxpayers filing their taxes early as well as overall, compared to not sending an email. The increases were significant, both statistically and practically and confirm the findings from other research that simple reminders can positively impact taxpayer behaviour (Guyton et al., 2016; Gillitzer and Sinning, 2020).

The simplified email that included planning prompts and asked people to sign up for further automatic email reminders (Treatment 3) performed better than all other treatment and control emails for the early filing outcome, and for all treatments except the simplified email with guidance (Treatment 2) for the overall filing outcome. The planning email increased overall and early filing rates by 1.10 and 2.07pp, respectively.

There is a strong correlation between signing up to the treatment website and filing. As people self-selected into choosing a date and signing up for a reminder email, we cannot conclude that this relationship is causal. Future research could therefore look at disentangling whether the overall effect was driven by the opportunity to choose a date in the calendar and receive reminders or simply by asking taxpayers to plan when they will file their tax return.

To our knowledge, this is the first time planning prompts have been used to change taxpayer behaviour, adding another potential tool to tax authorities’ arsenal. They could be particularly useful when combined with a message to increase initial motivation for behaviour change, such as a social norm or deterrence frame. Planning prompts can then help individuals follow through on their intention, rather than giving up when they encounter frictions in the process.

The email appealing to taxpayers’ national pride by graphically showing what public goods the taxes are spent on, performed worse than all treatments except for the two control conditions. This is in line with the results from Guatemala by Kettle et al. (2016) where an appeal to national pride was less effective than other message frames.

The treatment impacts were more pronounced for filing tax returns early than for filing them overall. This suggests that email reminders are more effective in making taxpayers move forward their filing date rather than inducing those to file that wouldn’t otherwise. This might be the case because the former represents a gradual change likely to be eventuated more easily than the more pronounced behaviour change to start filing tax returns. That said, it might also have encouraged filing among some taxpayers who otherwise would have given up when encountering system pressures or having to restart the filing process several times. Furthermore, it is possible that earlier filing has led to a better taxpayer experience and thus leads to higher compliance in the long run, an outcome that this study did not target. Future research in the Indonesian context should, therefore, aim to unpack the complex causal pathways of compliance in the short and long run.

Due to the good balance on observables and the very large sample size, the effect estimates are very precise. The secondary analysis produced some other, tentative insights. The email reminders had a smaller effect on early filings among self-employed
taxpayers, but a comparable effect on their overall filing rate. This might be due to tax returns for self-employed taxpayers being more complex and/or often being prepared by tax accountants who are less likely to adjust their behaviour in response to a light-touch nudge like our email intervention.\(^5\) Self-employed taxpayers are also more likely to have to pay taxes at the point of filing, which means it is rational for them to file at the last possible minute (to take advantage of interest gained on any cash withheld).

The emails have only a small and mostly insignificant effect on the overall amount of taxes collected and taxpayers’ overall probability to pay taxes. This is unsurprising as most taxpayers do not have to pay taxes at the point of filing (only 1.2% of taxpayers in the no-email control group make any payments during the filing process). We find variation in the effectiveness of the different treatments between regions. Importantly, however, there is no evidence of a backfire effect in any of the tax offices included – no-email reminder diminishes overall or early filing compared to not sending an email. This is a crucial insight for DJP, as well as tax offices in other countries, considering sending out similar reminder emails in future years: Not only are emails likely to be effective at changing taxpayer behaviour in the intended direction, there is little risk of negative effect based on the evidence from this trial. If it is easier to send the same email to all regional tax offices, we would generally not expect any harm in doing so. That said, more caution should be used with messages that might disenfranchise certain segments of the population, for example, because they might highlight a perceived unfairness. In these cases, testing impact by subgroup (defined geographically or otherwise) as done in this study, can address concerns around backfires.

**Limitations**

The population of interest for this study was all taxpayers registered for online filing in Indonesia as of late 2017. Since we covered online taxpayers across Indonesia, we are confident that the results hold across the country. However, we cannot say with certainty whether these results would also hold for those taxpayers not registered for online filing, as these might differ on important unobservable characteristics. That said, nothing in the most effective message is targeted specifically at online filers and the reminder emails provide planning prompts both for online and manual filing. While it might be more difficult to reach other taxpayers through emails and while the external validity of our findings is subject to limitations, we do not expect backfires if the same email was sent to all Indonesian taxpayers for which DJP holds an email address.

We cannot rule out the possibility that a few implementation challenges on the first day of implementation might have affected the results. These include one email provider blocking the emails due to suspected spam (this issue was resolved in collaboration with the provider), a small number of taxpayers (less than 0.2%) in the status quo email (Control 2) group receiving the Treatment 1 email and approximately 16% of the individuals in the Control 2 group not receiving the emails at all due to issues with the uploading process. While we could not identify which email addresses were affected, given the overall sample size it is unlikely that this

---

\(^5\)When taxpayers use tax accountants, the email addresses registered with DJP might be those of the accountants rather than the taxpayer themselves. It is hence possible that some of the emails were received by tax accountants.
biased estimates substantially. Crucially, robustness checks for the primary outcomes found negligible differences in treatment effects across days.

**Conclusion**

In countries across the globe, a large share of tax returns is submitted at the last minute, overloading pressured systems, and resulting in higher costs for tax administrations as well as lower tax compliance rates. This, in turn, diminishes governments’ revenue and hence the ability to pay for public goods. This study, one of the largest RCTs ever conducted by a government (>11 million participants), finds that designing an email intervention based on simple behavioural science principles significantly changes behaviour. A range of varying email reminders increased the share of Indonesians filing their tax returns two weeks prior to the deadline, compared to sending no email. If the most effective email, which helped participants to plan ahead and offered to opt-in for further email reminders, had been sent to the whole sample of Indonesian taxpayers registered for online filing, this would have resulted in an additional 133,000 tax returns filed overall and 233,900 returns submitted two weeks early.

The effect size, of 2.07pp for early filing and 1.10pp for overall filing, is in line with the average effect size of nudge interventions that Della Vigna and Linos (2022) find in the meta analysis of 126 RCTs run in the US (covering a variety of policy areas) and somewhat lower than the 3.1–7.7 percentage point effect found by Antinyan’s and Asatryan’s (2020) analysis of tax compliance nudges. Overall, it is at the lower end of the range of effect sizes found in some of the studies cited in this article (for detail, see Table F.9, Supplementary Appendix F). The significant differences in population, context and target behaviour across studies may partly explain this: our interventions aimed to encourage a ‘voluntary’ behaviour (early filing) in a context with low enforcement capacity, targeting a population which was not non-compliant at the time the emails were sent out.

Based on our findings, we recommended that DJP send the most effective email reminder to all personal taxpayers registered for online filing in future years. Moreover, given the negligible risk of negative effects, we propose expanding access to the planning website for manual filers. We believe that the findings of this study indicate that similar measures would prove beneficial to other countries as well, but strongly recommend designing different versions tailored to the pertaining context and testing which one works best.

Our main interest was to find the most effective email communication, rather than testing specific behavioural science concepts. Future research should aim to disentangle whether the planning prompt, the reminder emails or a combination of the two drove the overall impact of the most effective email. Similarly, it would be important to measure the impact of repeated treatment, both within and across years.

On a broader level, this study adds to a growing body of evidence on how simple, easily applicable scientific insights into human behaviour can substantially increase the effectiveness of policies, such as government communications. It demonstrates that with large enough samples – such as those often available to tax authorities – it is possible to rigorously test many different versions of government communication to establish which works best and to explore heterogeneity in treatment effects to establish the need for more targeted interventions to avoid backfires.
Finally, it illustrates the case for collaborations between tax administrations interested in basing their interventions in evidence and researchers interested in advancing our understanding of how behavioural sciences can be applied for more effective policy making, but who might not be able to access relevant data due to data protection and privacy laws.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/bpp.2022.25.

Acknowledgements. We would like to thank the staff at DJP, whose determination and support made this project possible: Puspita Wulandari, Hantriono Joko Susilo, Yon Arsal, Iis Insawati, Romadhaniah and Mohammad Wangsit Supriyadi for their support and encouragement throughout and Ikhwan Alisyahbani for spending countless days and nights to ensure that the emails got sent out as intended. We would also like to thank current and former BIT colleagues for their assistance: Rifki Akbari, James Watson and Paul Calcraft. BIT gratefully acknowledges funding from the Global Innovation Fund (GIF) for the programme of work with DJP. GIF was not involved in the design, execution or analysis of this study.

Competing interest. Gitarani Prastuti, Adityawarman, Muhammad Hakim Kurniawan, Gatot Subroto, and Muhammad Mustakim are employees of DJP. Ruth Persian and Alex Sutherland are employees at BIT.

References