ONLINE APPENDIX

Disrupting Compliance:

The Impact of a Randomized Tax Holiday in Uruguay

Thad Dunning* Felipe Monestier† Rafael Piñeiro,‡

Fernando Rosenblatt,§ and Guadalupe Tuñón¶

June 7, 2024

*Robson Professor, Department of Political Science, University of California, Berkeley; thad.dunning@berkeley.edu.

†Assistant Professor, Departamento de Ciencia Política, Universidad de la República; felipe.monestier@cienciassociales.edu.uy

‡Professor, Departamento de Ciencias Sociales, Universidad Católica del Uruguay; rafael.pineiro@ucu.edu.uy.

§Senior Lecturer, Department of Politics, The University of Manchester; fernando.rosenblatt@manchester.ac.uk.

¶Assistant Professor, Department of Politics and School of International and Public Affairs, Princeton University; tunon@princeton.edu.

Contents

List of Tables

List of Figures

A Taxpayer reward programs in Latin America

Latin American municipal governments have recently offered many prize programs that provide positive incentives for tax compliance. Cities from Salta, Argentina to Peruíbe, Brazil to Miraflores, Peru now raffle prizes—from televisions to new cars and houses to discounted payments and free travel—to reward and motivate punctual tax payment.

In order to assess the prevalance of such policies, in 2015 we identified all Latin American countries in which the land value tax (*impuesto predial*) is assessed at the municipal level. For countries with more than 500 municipalities, we took a simple random sample and used web searches and interviews with municipal authorities to measure the presence of such prize programs in the sample. We drew a 10 percent random sample of municipalities in Colombia, Brazil, and Mexico, while in Peru and Argentina we drew larger samples (22 and 21 percent of municipalities, respectively). We included a census of municipalities in Uruguay, Bolivia, and Ecuador.

Figure A1 shows our estimates of the percentage of municipalities that offer reward programs for up-to-date taxpayers. These programs are substantially frequent, turning up in 37 percent of sampled municipalities in Ecuador and 24 and 25 percent in Brazil and Ecuador, respectively. The rewards vary substantially in form: they include discounts for cash payments, annual payments in full, or payments in advance. Only a portion of the municipalities award prizes via lotteries: in descending order, 15.8 percent of municipalities in Uruguay, 8.9 percent in Brazil, 8.7 percent in Mexico, 5.9 percent in Peru, and 2.5 percent in Ecuador, and percentages under 1 percent in the other countries. Prizes include automobiles, televisions, refrigerators, free trips, and the like. Only in Uruguay (3 out of 19 municipalities) are rewards offered in the form of tax holidays.

Our survey of the prevalence of these programs arose from our broader interest in understanding the effects of positive incentives for tax compliance. We provide here some information we gathered from qualitative interviews, both on the sources and rationale for these programs and their intended and perceived effects. Ultimately, we relegate this material to this online appendix because many of these other Latin American programs do not interrupt tax compliance, as in Montevideo. Thus, while our survey is relevant to identifying the empirical scope of positive incentive programs, they do not directly reflect the program feature we identified as central to explaining the negative impact on tax compliance of the incentive program in Montevideo.

First, regarding the sources and rationale: the programs have often arisen in the context of widespread non-compliance and amnesties for delinquent taxpayers. As in Montevideo, where the center-left government of the Frente Amplio hoped precisely that its prize policy would counter the perceived negative incentive effects of amnesties following the economic crisis of 2002—which eased the burden on delinquent taxpayers without offering a compensating benefit for punctual compliers—the desire to reward compliant taxpayers was important in many other contexts. Many officials also emphasized difficulties with enforcement of sanctions for non-payment of taxes, heightening the appeal of trying to use positive incentives to boost compliance instead. As a municipal tax official in Argentina noted to us, "proximity means that a neighbor can approach the administration to justify why he doesn't pay…In the last 20 years, we have never auctioned either a commercial or residential property. In general, we end up with an agreement" in which delinquent taxpayers consent to a discounted payment plan.[1]

Second, regarding impact: many of our interviewees certainly believed that the programs boost tax compliance. As an Argentine official noted, "We have a compliance rate of 85% with the Municipal Service Tax today, whereas when we started [prize lotteries] in 2009, it was at 68%."[2] They argue, plausibly, that prizes may not only sustain compliance among good taxpayers but also induce bad taxpayers to bring their accounts up to date. As another Argentine municipal tax official noted, "one of the conditions [to participate in a prize lottery] was not to be delinquent on payments. What did people say? 'Make me a payment plan'" to allow entry to the lottery.[3] A former Brazilian mayor noted that after initiation of the prize program, "many indebted people went to look for payment plans."[4] An Argentine interviewee emphasized the public, credible nature of his muncipality's prize lottery, as well as the importance of social recognition: "We take a photo and put the program on the webpage of the municipality, we publish a

[1]Interview, Daniel Chillo, municipality of Tigre; all translations ours.
[2]Ibid.
[3]Carlos Maisterrena, Ciudad de Paraná, Entre Ríos, Argentina.
[4]Beto Trícoli, mayor of Atibaia, Brazil (2001-2008).

list of the taxpayers among whom we do the lottery.'[5] A Brazilian mayor noted,"the best weapon [against non-compliance] is transparency."[6] Another interviewee noted, "every time we award prizes we make a kind of show.... this helps create a taxpaying culture.... It is not worth it to hit unless you also caress those you need to caress."[7]

The benefits of these programs throughout Latin America are, of course, largely unproven. The impact of the material incentives may be limited, for the reasons we described in the paper; and it is an open question whether the programs indeed boost the expressive benefits of paying taxes, for example, by shaping attitudes towards the equity and transparency of the tax system. Given the cost of such programs in terms of the value of prizes awarded, it is also unclear how their net fiscal impact would compare to more typical negative inducements (threats of audits or punishments for delinquent taxpayers, for example). We emphasize again that—in view of our findings that habit disruption induced a negative effect on tax compliance in Montevideo's lottery—our results may not readily extend to other prize programs (outside of the two other Uruguayan municipalities that have also used holidays). However, our findings suggest limited effects of information about the prize program in our field experiment in Montevideo: information did not shape payment behavior any more than a placebo reminder. This may suggest constraints on the positive impacts of these other programs. Designs such as ours could be used to study the impact of these other programs in Latin America.

[5] Chillo, note 1.
[6] Geraldo Cruz, former mayor of Embu das Artes, São Paulo, Brazil.
[7] Maisterrena, note 3.

Figure A1: **Reward Programs for Good Taxpayers in Latin America**

The figure shows the percentage of municipalities that offer reward programs for up-to-date taxpayers. Based on a 10% random sample of municipalities in Argentina, Bolivia, Brazil, Colombia, Ecuador, Mexico, and Peru, as well as a census of municipalities in Uruguay; in other Latin American countries, we identified no rewards programs using Web searches and ancillary sources.

B Natural experiment: Design and Additional Analyses

Figure A2: **The Tax Holiday Lottery: Selection of Winning Account Numbers**

Institucional | Trámites y tributos | Tránsito y transporte | Servicios y sociedad | Empresas | Ciudad y cultura

Intendencia deMontevideo

Sorteo Bimestral para Buenos Pagadores

Listado de Posibles Cuentas Ganadoras

Sorteo Bimestral Febrero 2009 Terminación: 8662

Tasa General Municipal	Patente de Rodados		Contribución Inmobiliaria		Tarifa de Saneamiento
Cuenta Corriente	Cuenta Corriente	Identificación	Cuenta Corriente	Identificación	Cuenta Corriente
638662	1208662	305612/SAW9167	218662	564/401	2798662
668662	1318662	360356/SAH2587	258662	10282/301	2808662
678662	1328662	382765/SAQ2304	268662	14659/401	2878662
688662	1408662	463531/SAH7615	288662	21837/101	2898662
728662	2208662	904461/SAG7716	298662	26220/803	2908662
768662	3518662	1004335/SBC2701	308662	29734/502	3168662
798662	3618662	1016292/SBD1254	318662	31711/503	
838662	3718662	1028575/SBD7490	328662	34080/601	
948662	3868662	1056616/SBE8813	338662	39851/4	
1008662	3978662	1078082/SBF6388	418662	85343	
1018662			468662	114666	
1038662			478662	121098/3	
1078662			508662	141019/1	
1098662			538662	161294/202	
1118662			548662	168557	
1738662			2028662	151499	
2148662			3778662	1712/202	

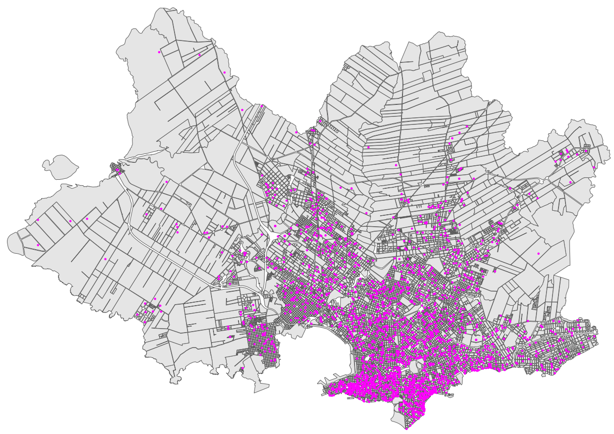
The figure shows a screenshot of winning taxpayer accounts in the February 2009 tax holiday lottery published on the website of Montevideo's municipal government. Across four types of taxes—head ("Tasa General Municipal"), vehicle ("Patente de Rodados"), property ("Contribución Inmobiliaria"), and sewage ("Tarifa de Sanamiento")—the figures shows all taxpayer accounts that end in "8662" and for which taxes have been paid promptly over the previous year. Source: `http://www.montevideo.gub.uy/sorteosBP/pages/consCuentasSorteadas.xhtml`, accessed April 4, 2016.

Table A1: **Natural Experiment: Balance Tests on Pre-Treatment Covariates. Winning vs. Non-Winning Account Numbers.**

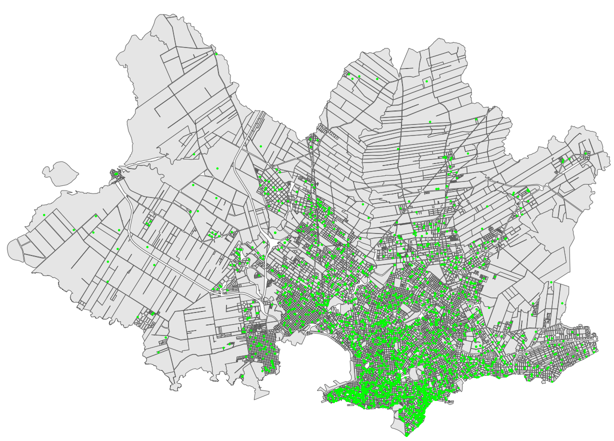
	Control Group Mean	Difference	SE_Diff	N	p-value
Sample: All taxes					
Good taxpayer at t=0	0.346	0.003	0.007	18621	0.651
Good taxpayer at t=-1	0.347	0.001	0.007	18309	0.934
Good taxpayer at t=-2	0.345	0.002	0.007	17991	0.823
Good taxpayer at t=-3	0.346	0.002	0.007	17670	0.769
Good taxpayer at t=-4	0.345	0.005	0.007	17350	0.514
Good taxpayer at t=-5	0.345	0.003	0.007	17021	0.689
Sample: Property, head & sewage					
Automatic debit at t=0	0.106	-0.003	0.005	14027	0.506
Automatic debit at t=-1	0.106	-0.003	0.005	13807	0.503
Automatic debit at t=-2	0.106	-0.004	0.005	13580	0.498
Automatic debit at t=-3	0.106	-0.004	0.005	13350	0.447
2004 Property value	1259241	-199730	164141	13462	0.224
Current property value	2747841	-305151	345883	13998	0.378
Rented Property	0.243	-0.009	0.007	13998	0.229
Sample: Property tax					
Retiree	0.008	-0.001	0.002	5129	0.655
Paid year in full	0.287	-0.010	0.013	5129	0.441

The table compares taxpayers with winning lottery numbers (treatment group) to those in our random sample of non-winning lottery numbers (control group) on pre-treatment covariates, including past tax compliance, whether the taxpayer pays by automatic debit, their parcel's value in 2004 (before the tax holiday program took effect), and status as a retiree at the time of the lottery. Here, $t = 0$ is the payment period in connection with which the taxpayer won or could have won (in the control group) the lottery. To allow higher-powered balance tests, we include both good and bad taxpayers, as measured at the date of each lottery. Note that past compliance should be highly prognostic of future compliance. Taxpayers with winning and non-winning lottery numbers are statistically indistinguishable on these covariates, consistent with random assignment.

Figure A3: **Natural Experiment: Property Plots of Winning Account Numbers.**



(a) Winning Account Numbers



(b) Non-Winning Account Numbers

The maps show the geographic distribution of winning (treatment group - panel a) and non-winning (control group - panel b) taxpayers across Montevideo.

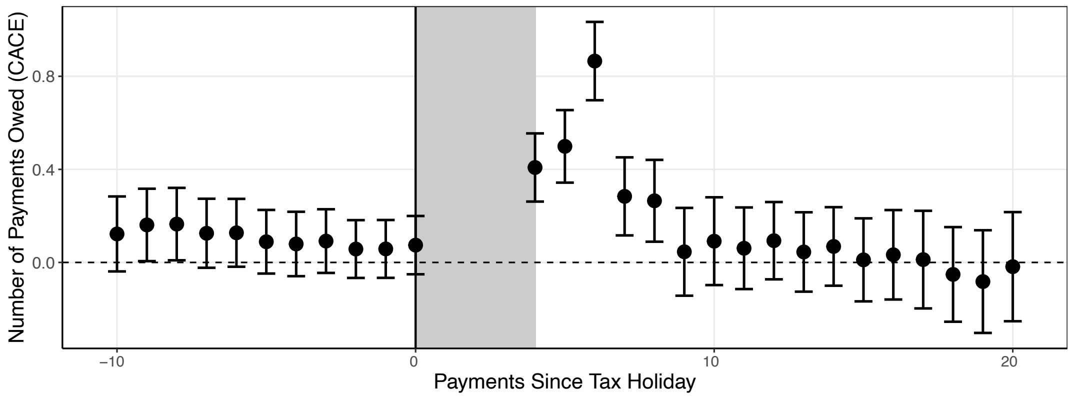
Table A2: **Natural Experiment: Placebo Tests with Post-Treatment Variables - Ineligible Taxpayers. Winning vs. Non-Winning Account Numbers.**

	Mean, Non-Winning Account Numbers	Difference of Means	SE	N	*p*-value
Good Taxpayer t = 1	0.085	0.006	0.005	11639	0.288
Good Taxpayer t = 2	0.095	0.007	0.006	11489	0.187
Good Taxpayer t = 3	0.114	0.002	0.006	11339	0.697
Good Taxpayer t = 4	0.116	0.002	0.006	11192	0.688
Good Taxpayer t = 5	0.115	0.004	0.006	11052	0.493
Good Taxpayer t = 1—5	0.105	0.005	0.005	11639	0.352

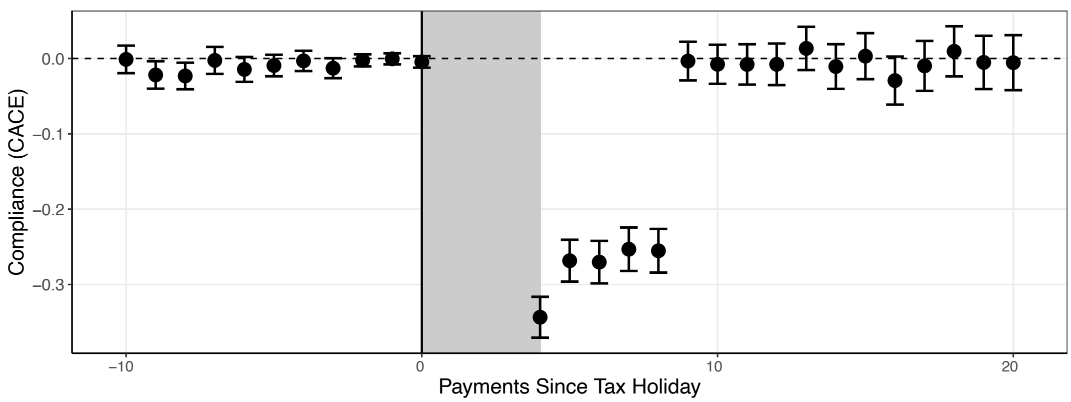
The table shows treatment effects of a winning lottery number for taxpayers who were ineligible to win a tax holiday as of the date of the relevant lottery—and for whom no treatment effects should therefore exist. For the outcome variable, we use the proportion who are good taxpayers at different post-treatment payment periods. Here, "t = 1" refers to the first payment period at which we can compare treatment and control groups in a symmetric way (e.g., just to the right of the grey strip in Figure 2 in the paper).

Figure A4: **Natural Experiment: The Negative Impact of Holidays on Compliance. Effects on the Number of Payments Owed, Compliance, and Total Debt.**

(a) Number of Payments Owed



(b) Compliance

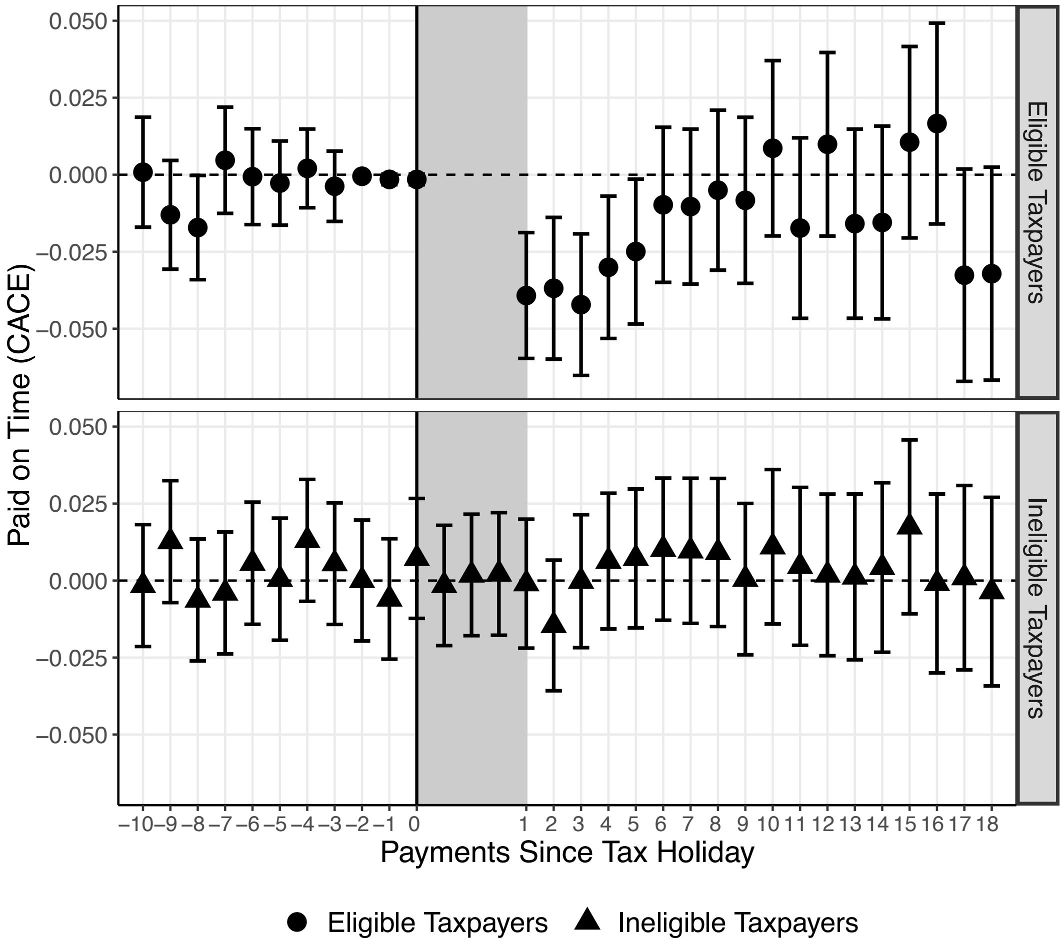


(c) Total Debt as of October 2014

Estimate	**SE**	**p-value**	**N**
183.64	296.68	0.536	6355

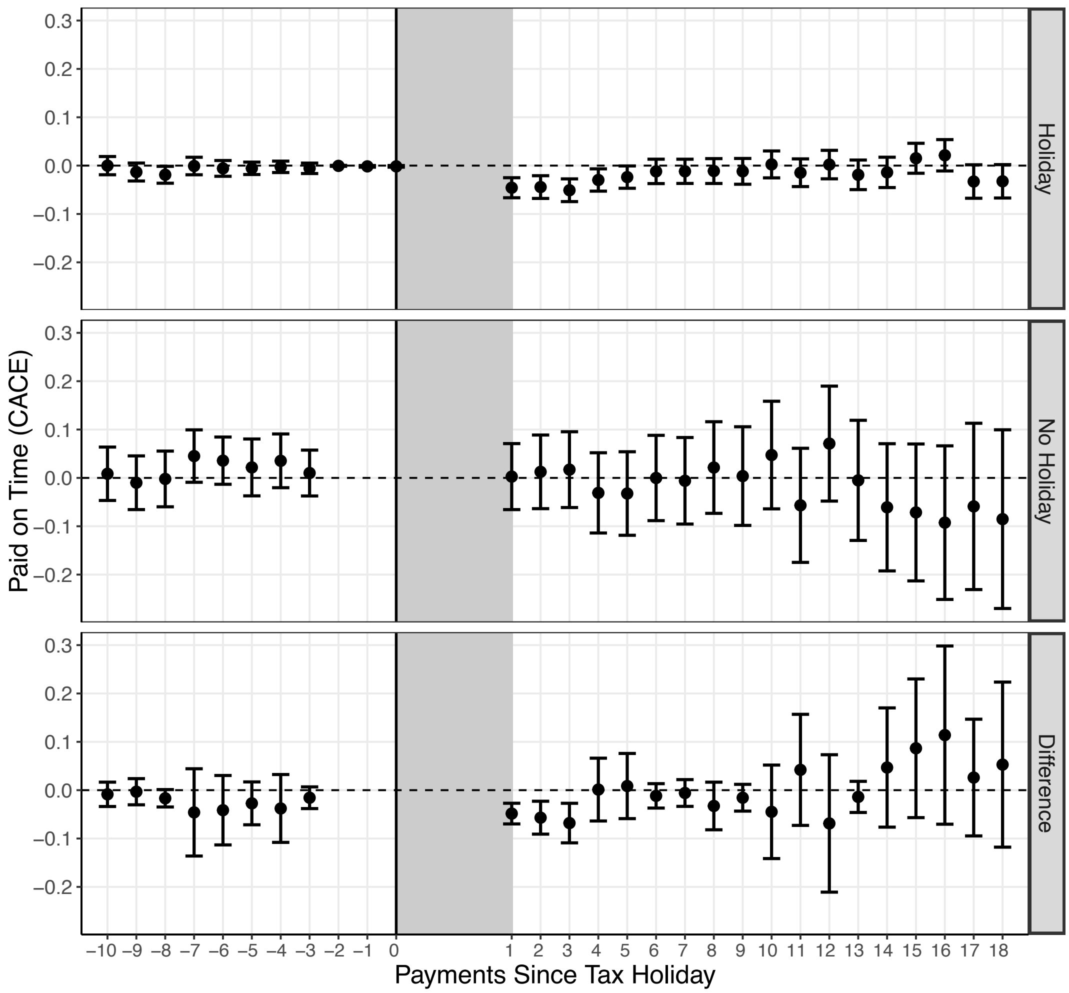
The figure depicts balance tests and estimated treatment effects for three additional outcomes in the natural experiment. For panels (a) and (b), the horizontal axis measures tax payment periods before or after the tax holiday. The grey vertical strip indicates the period of the tax holiday (the treatment). Comparisons between winners and non-winners to the left of zero—the date at which each taxpayer won or could have won a particular tax holiday lottery—test for balance on pre-treatment tax compliance. Post-treatment comparisons among eligible taxpayers to the right of the vertical grey strip estimate the treatment effects of the tax holiday. The vertical axis shows the estimated complier average causal effect ($\widehat{CACE}$) for (a) the number of payments owed and (b) the proportion of taxpayers who are fully up to date with their taxes. Compliers are taxpayers with winning account numbers who would claim the exoneration. Vertical lines show 95% confidence intervals. In panel (c), we additionally report results for a difference in means test using total debt data as of October 2014.

Figure A5: **Main Results - Full Post Treatment Period**



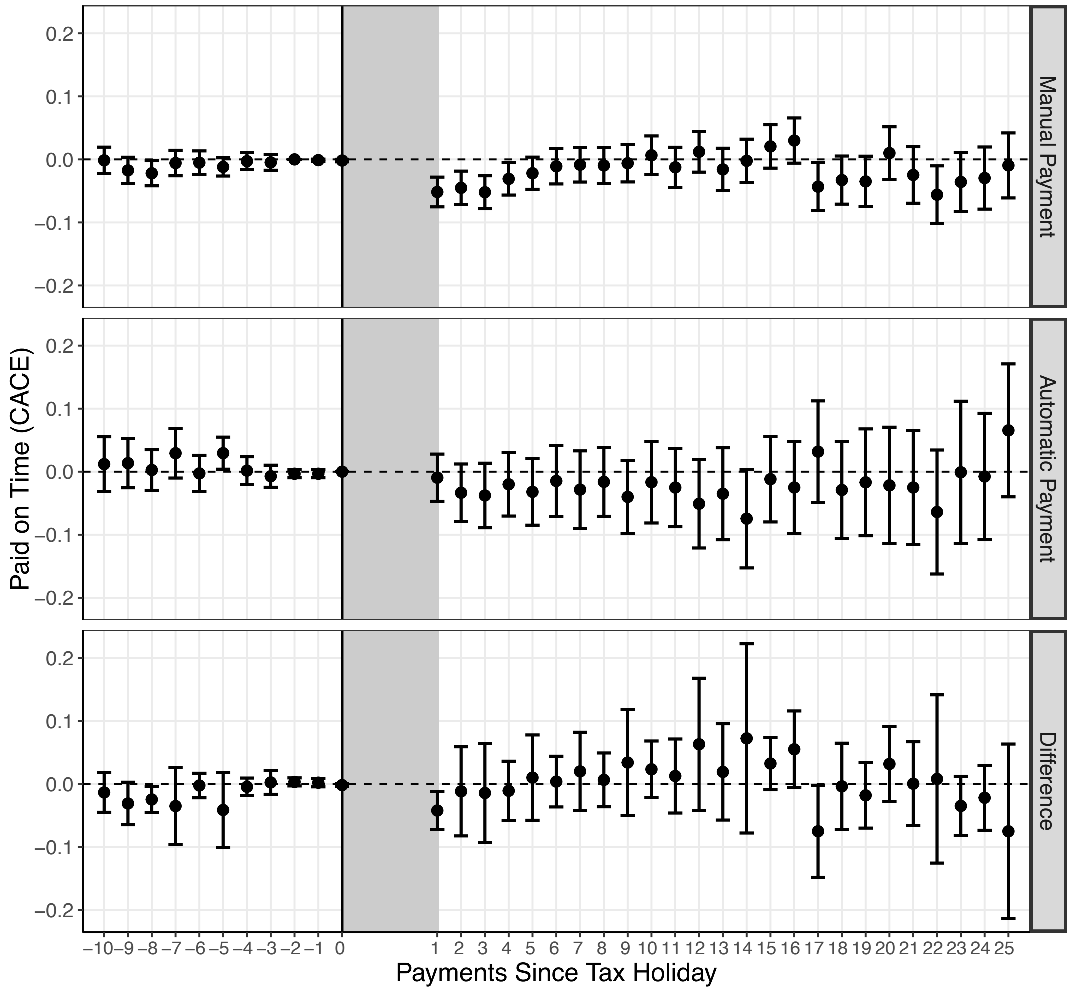
The figure depicts balance tests, placebo outcome tests, and treatment effects in our natural experiment. The horizontal axis measures tax payment periods before or after the period of the tax holiday (grey vertical strip). The vertical axis shows the estimated complier average causal effect (CACE) for the proportion of taxpayers who paid on time at each payment period. Vertical lines show 95% confidence intervals.

Figure A6: **Treatment Effects By Type of Tax: Holiday vs. No Holiday - Full Post Treatment Period**



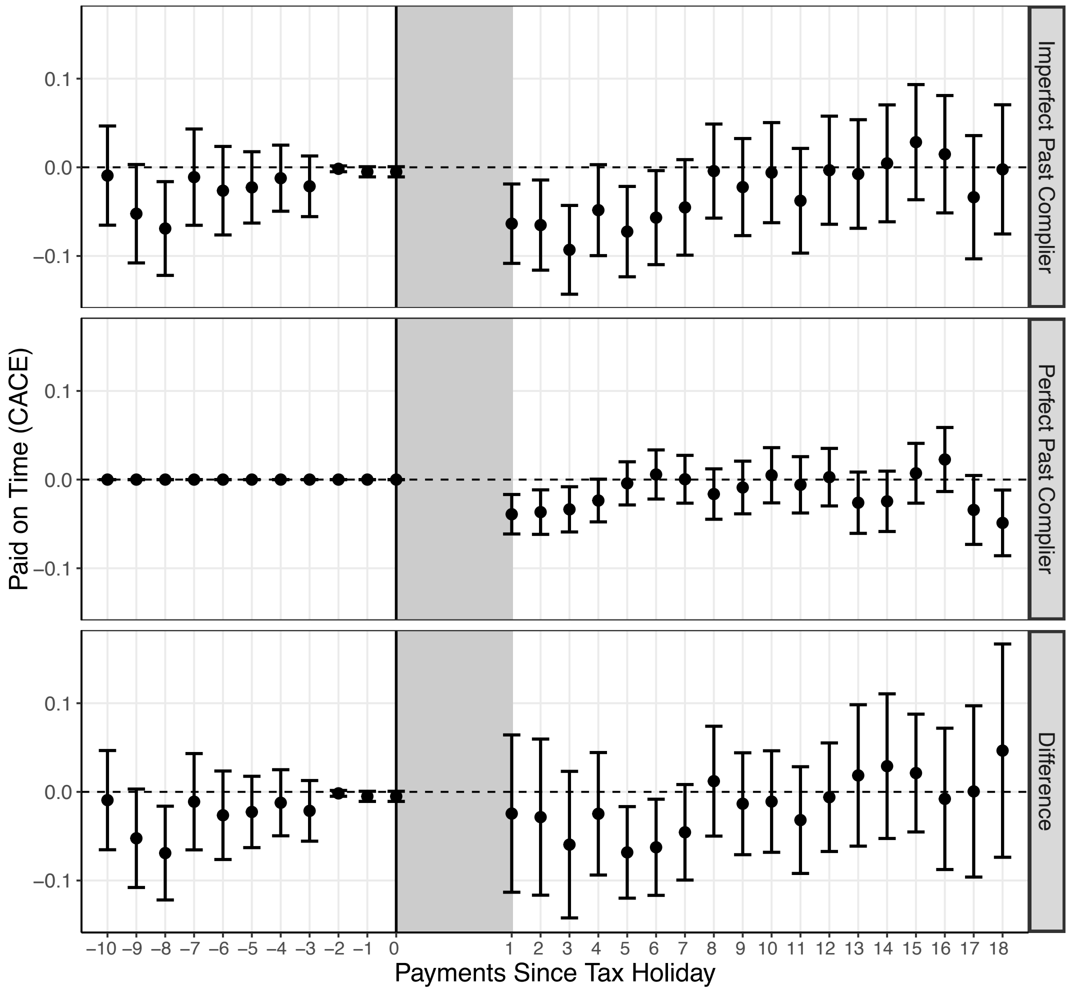
The negative effect only occurs when the payment habit is actually interrupted, in the "Holiday" taxes. See also Online Appendix Figure A5 for pre-specified analysis of heterogeneous effects by type of tax.

Figure A7: **Placebo test: Treatment Effects for Automatic vs. Manual Payers - Full Post Treatment Period**



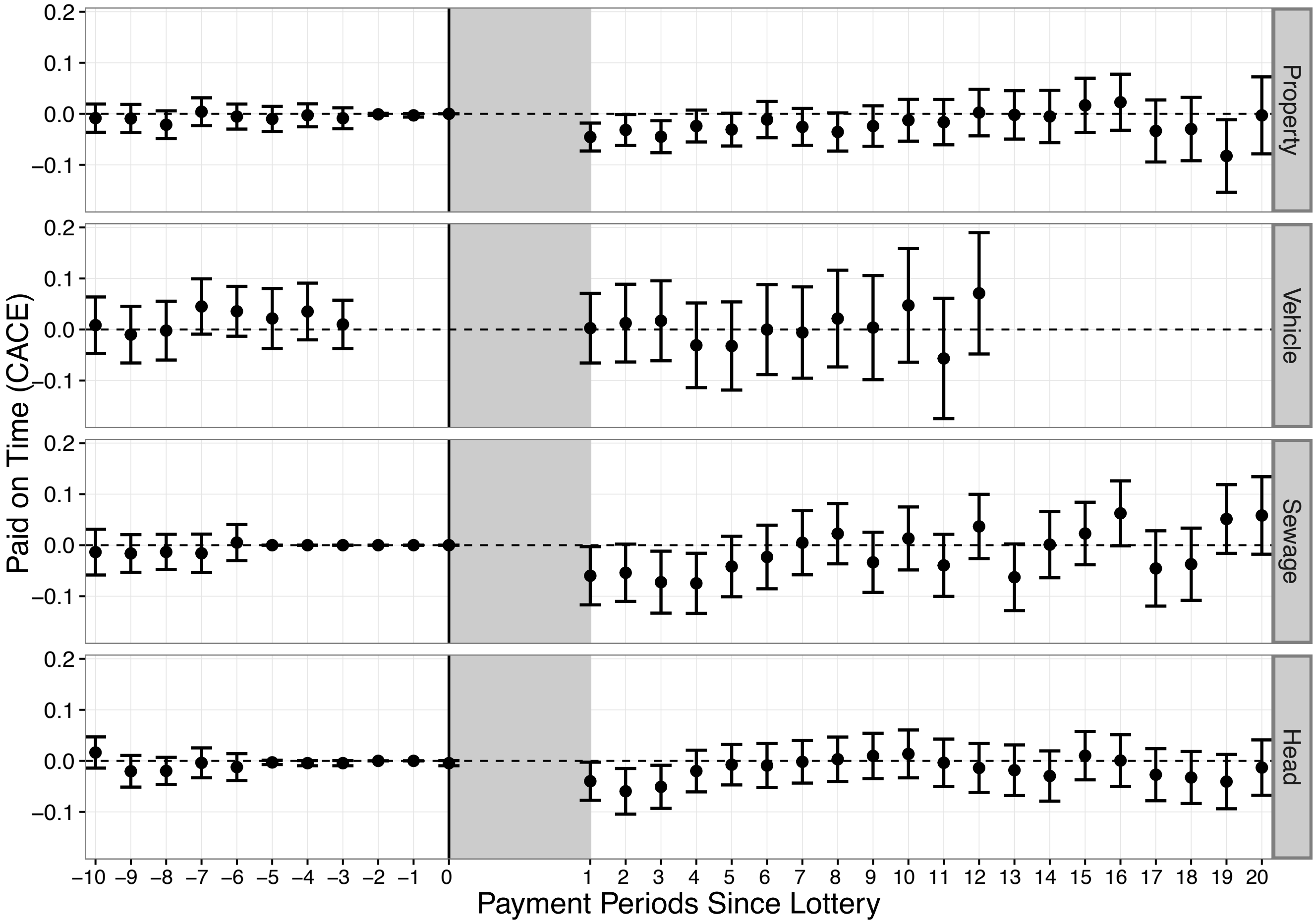
The negative effect of the holiday only appears for manual taxpayers, not for those in automatic withdrawal programs. The figure depicts effects for the property, head, and sewage taxes. (Registered test of Mechanism 1B.2 in the PAP).

Figure A8: **The Stock of Habit: Perfect vs. Imperfect Past Compliers - Full Post Treatment Period**



The negative effect of winning is less pronounced for taxpayers with a greater reserve of payment habit. "Perfect Past Compliers" paid punctually in all 15 payment periods prior to winning the lottery, while "Imperfect Past Compliers" failed to do so in at least one period. Differences in estimated average causal effects are statistically significant at the 0.05 level for the first seven post-treatment payment periods. (Registered test of PAP Hypothesis 2C, though the PAP not discuss that test with respect to the habit mechanism).

Figure A9: **Natural Experiment: Heterogeneous Treatment Effects by Type of Tax.**



The figure A9 depicts balance tests and treatment effects by tax. The horizontal axis measures tax payment periods before or after the tax holiday for panel. The grey vertical strip indicates the period of the tax holiday (the treatment). Comparisons between winners and non-winners to the left of zero—the date at which each taxpayer won or could have won a particular tax holiday lottery—test for balance on pre-treatment tax compliance. Post-treatment comparisons among eligible taxpayers estimate the treatment effects of the tax holiday. The vertical axis shows the estimated complier average causal effect ($\widehat{CACE}$) for the proportion of taxpayers paying on time at each period. Compliers are taxpayers with winning account numbers who would claim the exoneration. Vertical lines show 95% confidence intervals.

B.1 Additional alternative explanations

As discussed in the text, results from our field experiment cast doubt on several important alternative explanations for the negative impact of winning the tax holiday, especially those focused on the information conveyed by the lottery. Here we assess several other possible explanations for negative effects that we considered in our pre-analysis plan and that could serve as alternative explanations but for which space constraints did not allow discussion in the paper.

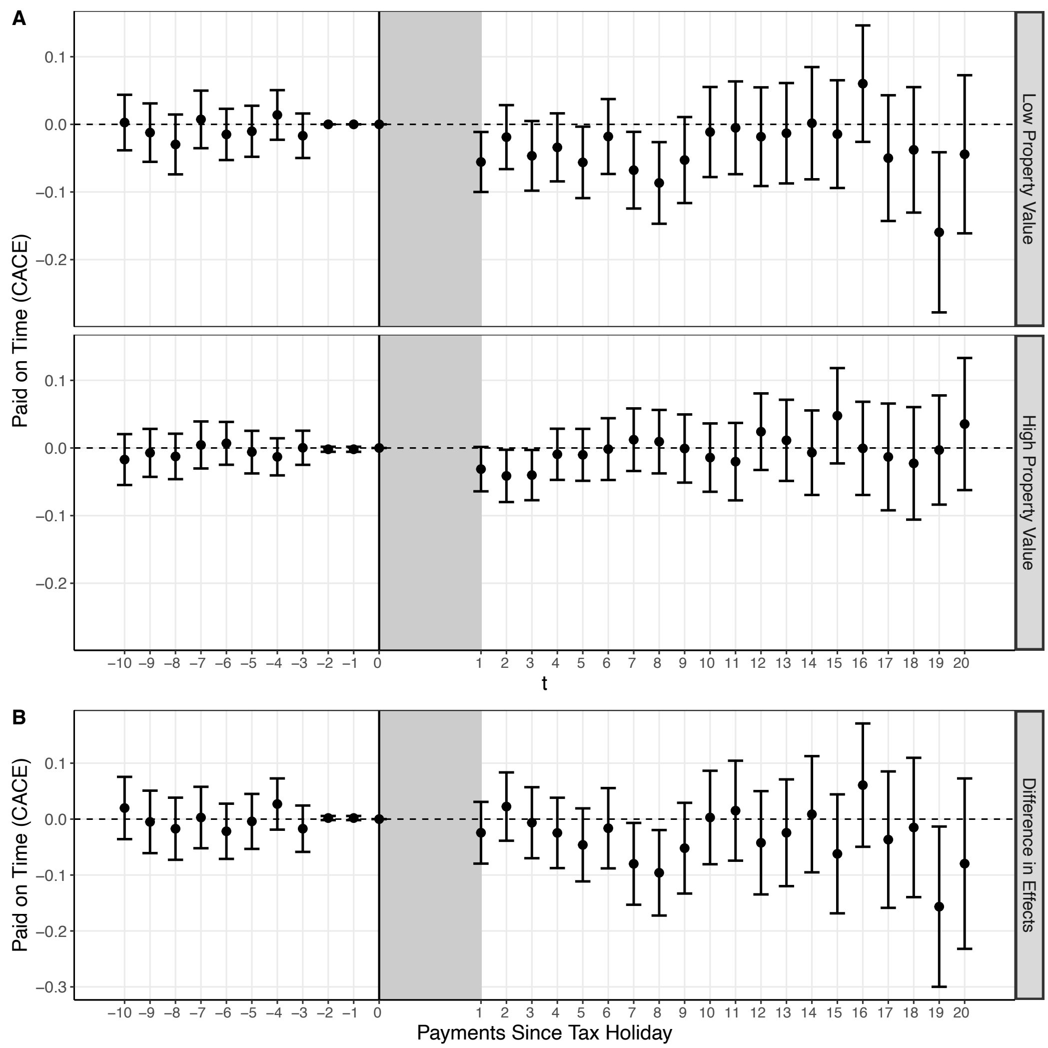
First, per Mechanism 1B.1 (original PAP, p. 37)—"income effects" might in principle explain a negative effect of winning the lottery on future compliance. Perhaps the additional income from a year free of paying taxes buttresses winners against the costs of punishment in case of non-payment for future taxes, effectively lowering the parameter c in our decision model. In qualitative interviews, we asked winners of property tax holiday what they did with the extra income gained from the holiday. Many did mention using this to cover other costs, though taxpayers afforded differential importance to the monetary prize. One spoke fondly of using the extra money to eat out in restaurants (Interviewee CTA 512794), while another said "it didn't change my life at all, it was enough to buy a good pair of shoes" (Interviewee CTA 334095).

To assess this hypothesis quantitatively, we conduct heterogeneous effect analysis to assess whether the effects of winning the tax holiday vary according to the cost of payment, operationalized in terms of the property value. We do not find evidence for such disparate impacts for richer and poorer taxpayers (Online Appendix Figure A10). Admittedly, the interpretation of these tests is tricky: the idea is that differences in the economic importance of the rebate to different taxpayers might suggest heterogeneous effects of the holiday for richer and poorer taxpayers. Yet, the wealthy have more valuable properties and therefore receive larger tax breaks during the holiday—but they also tend to have greater income. Still, the fact that patterns of effects are similar for rich and poor taxpayers suggests that income effects may be unlikely to explain a negative effect of winning the lottery.

Second and perhaps more powerfully, the null effects for the vehicle (no holiday) tax are also inconsistent with the alternative interpretation in Mechanism 1B.1. After all, winners of the vehicle tax lottery receive an important temporary income shock that they can likewise divert to consumption goods. If substitution effects explain our findings, we should thus observe a negative effect for winners of this lottery. Instead, consistent with our interpretation, we find a null effect for the vehicle tax—the one tax in which the payment habit is not interrupted (Figure A9).

We can also use our analyses in the paper to evaluate other alternative hypotheses we did not consider in our PAP. Perhaps experiencing the exoneration breaks a "taboo" against non-compliance or encourages taxpayers to think of themselves as non-compliers. This explanation is not unrelated to habit—in that failure to pay has a causal impact on future payment propensities—but seems closer to what we could call "identity." Another possibility is that by stimulating extrinsic incentives to comply, the material rewards offered by holiday might crowd out intrinsic incentives, such as a sense of civic duty (**?**). In both cases, however, we would expect the impact of receiving the holiday to be more or less permanent. Once a taxpayer has broken the taboo or learned of the extrinsic incentives, she should tend to comply at consistently lower rates in the future. Instead—consistent with our theory of habit in Section 2 of the paper but not with these alternative explanations—we find negative but decaying effects of winning the lottery. Moreover, if the breaking of taboos or the crowding out of extrinsic incentives account for our findings, we might also expect the negative impact to be greatest for those who have previously always fulfilled their civic duty. Instead, the effect is weaker for those with a greater stock of habit.

Figure A10: **Natural Experiment: Heterogeneous Treatment Effects by Property Value.**



The figure depicts balance tests and treatment effects for the property tax by property value. We consider all properties valued above the median property value as "high" whereas those below the median are classified as "low property value." In both panels A and B, the horizontal axis measures tax payment periods before or after the tax holiday for panel. The grey vertical strip indicates the period of the tax holiday (the treatment). Panel A shows comparisons between winners and non-winners to the left of zero—the date at which each taxpayer won or could have won a particular tax holiday lottery—test for balance on pre-treatment tax compliance. Post-treatment comparisons among eligible taxpayers estimate the treatment effects of the tax holiday. The vertical axis shows the complier average causal effect (CACE) for the proportion of taxpayers paying on time at each period. Compliers are taxpayers with winning account numbers who would claim the exoneration. Panel B shows the difference in effects across the two property value levels. Vertical lines show 95% confidence intervals.

B.2 Adjustment for multiple comparisons

On pp. 43-44 of the original pre-analysis plan, we specified Bonferroni and Benjamini and Hochsberg (FDR) adjustment of p-values for multiple comparisons as follows:

"How large is m under our study design? This differs for the field, natural, and survey experiments. For the natural experiment, we have one randomization into treatment (winner) and control (eligible non-winner) groups. (As discussed, this is really blocked randomization, where the blocking is by type of lottery; however, the blocks are all in expectation the same size for each type of tax). Meanwhile, we will have nominal p-values associated with each of the following comparisons:

- K-S test (three outcomes: compliance, missed payments, and total debt)
- Diff-in-diff (three outcomes)
- Persistence of effects, heterogeneous effects (three outcomes)
- Difference of means (six outcomes: trust in municipality, trust in civil servants, evaluation of mayor, fairness of taxes, fairness of the tax specific to the corresponding lottery, and opinion of lottery)
- Heterogeneous effects, by cost of payment (three outcomes)
- Heterogeneous effects, by time since winning (three outcomes)
- Heterogeneous effects, by beliefs about non-independence of winnings (three outcomes)

The original total number of comparisons is 24. We also have the p-value for the comparison of effects in the natural and field experiment. This makes a total of m = 25 p-values for the natural experiment."

However, we were not able to collect the relevant data and conduct all of these tests: for reasons explained in the second amendment to our PAP, the survey data with which we planned to measure the six outcomes (trust in municipality, trust in civil servants, evaluation of mayor, fairness of taxes, fairness of the tax specific to the corresponding lottery, and opinion of lottery) and beliefs about non-independence of winning were not gathered in a way that made feasible the estimate of treatment effects in the natural experiment. In our revision to Table 7.3 specifying tests, we also got rid of the K-S test (p. 24 of second amendment). While this was not amended in the analysis, a review of these tests revealed that the heterogeneous effects looking at persistence were essentially equivalent to the heterogeneous effects by time since winning. This therefore leaves the following set of registered multiple comparisons (leaving $m = 10$):

- Diff-in-diff (three outcomes)
- Persistence of effects, heterogeneous effects (three outcomes)
- Heterogeneous effects, by cost of payment (three outcomes)
- Effects in the nat. exp vs. field experiment (one outcome)

The second amendment to our PAP also notes that "our primary outcomes for the field and natural experiments are the administrative measures of tax compliance, missed payments, and total debt... We will adjust statistical tests for multiple comparisons with respect to [the] primary outcomes measured through administrative data." For the measure of total debt, the Montevideo government provided us with information on the total number of payments owed by the taxpayer in each period. For all comparisons we measured the outcomes as pre-registered for the main analysis, which is the mean of the first year post-treatment versus the first year pre-treatment.[8] For ease of interpretation and because the pre-treatment levels of compliance often do not substantially differ among eligible taxpayers, we report differences in levels in the paper (e.g. Figure 4 and Table 4.3), but a complete set of results using the difference-in-differences analysis is reported in the reproduction of our replication code (section F) as well as in the PAP analysis (Section G.4). For the persistence tests, we focus on the comparison of the effects between the first and second year after treatment.[9] Finally, as stated in the first PAP amendment, we operationalized the cost of payment using the property value, which reduces the sample of that test to the property tax. For the comparison between the natural and field experiment, we use missed payment in the first period after the intervention as our main outcome.

[8]"Effects of the tax holiday (difference in differences analysis). Effects of the tax holiday. Comparing winners to non-winners, difference in difference analysis (comparison A=mean of the year before winning vs. mean of the year after the tax holiday; comparison B= mean of three years before winning vs. mean of three years after tax holiday). Tests using compliance as an outcome are conditional on finding effects for either missed payments or number of payments owed for the relevant period. This is because compliance is a stricter test, and if we find effects for neither missed payments of number of payments owed, there will be no effects by construction for compliance."

[9]To calculate the second year outcome, we again take the mean of each outcome for the second year and subtract the mean of the outcome for the year before the treatment.

Table A3 below reports nominal p-values as well as those adjusted using the Bonferroni method and the Benjamin and Hochsberg correction to control the false discovery rate. All three of the p-values for the one-year difference-in-differences are nominally significant and survive both corrections. The decay of the effects over time is nominally reflected in the first versus second year comparisons for the number of payments owed as well as overall compliance, but not for missed payments. These two effects also survive both multiple comparison adjustments. The difference in the heterogeneous treatment effects by property value was not nominally significant for either the mean of missed payment of the number of payments owed. However, the results on compliance were nominally significant—with a negative effect sightly larger for taxpayers with a higher property value. This result survives both the Bonferroni and FDR corrections. Finally, the difference in effects across the natural and field experiments is nominally significant and survives the FDR adjustment but the Bonferroni-adjusted p-value is .09.

Table A3: Natural Experiment Pre-Registered Tests: Adjusted p-Values.

Test	**Outcome**	**Nominal p-value**	**Bonferroni p-value**	**FDR p-value**
Effects of Tax Holiday	1yr DiD Missed Payment	0.00004	0.00036	0.00007
	1yr DiD Nr Payments Owed	0.00000	0.00000	0.00000
	1yr DiD Compliance	0.00000	0.00000	0.00000
Persistence (year 1 vs 2)	1yr DiD Missed Payment	0.29096	1.00000	0.32329
	1yr DiD Nr Payments Owed	0.00001	0.00014	0.00003
	1yr DiD Compliance	0.00000	0.00000	0.00000
Property Value HTEs	1yr DiD Missed Payment	0.35993	1.00000	0.35993
	1yr DiD Nr Payments Owed	0.11746	1.00000	0.14683
	1yr DiD Compliance	0.00525	0.05246	0.00874
Natural vs Field Experiment	Missed Payment	0.00912	0.09119	0.01303

C Field Experiment: Design and Additional Analyses

Figure A11: **Field Experiment: Flyer with text of informational intervention (Spanish) - Reminder**

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

Por consultas:

FONO TRIBUTOS
1950 3000

Figure A12: **Field Experiment: Text of informational intervention (Spanish) - Reminder + lottery + individual benefit**

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

La Intendencia de Montevideo quiere premiar a los buenos pagadores. **Si usted paga en fecha participará automáticamente de un sorteo por la exoneración de un año de Contribución Inmobiliaria.**

Los sorteos se realizan todos los meses pares del año junto con la Lotería Nacional. Los beneficiados serán debidamente informados y se publicarán los resultados en el sitio web de la Intendencia.

¡Usted puede ser el próximo!

Por consultas:

FONO TRIBUTOS

1950 3000

Figure A13: **Field Experiment: Flyer with text of informational intervention (Spanish)- Reminder + lottery + individual benefit + probability of winning**

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

La Intendencia de Montevideo quiere premiar a los buenos pagadores. **Si usted paga en fecha participará automáticamente de un sorteo por la exoneración de un año de Contribución Inmobiliaria.**

En cada sorteo, 1 de cada 5.000 hogares recibe este beneficio.

Los sorteos se realizan todos los meses pares del año junto con la Lotería Nacional. Los beneficiados serán debidamente informados y se publicarán los resultados en el sitio web de la Intendencia.

¡Usted puede ser el próximo!

Por consultas:

FONO TRIBUTOS
1950 3000

Figure A14: **Field Experiment: Flyer with text of informational intervention (Spanish) - Reminder + individual punishment**

Intendencia
deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

Quienes no paguen en fecha podrían estar sujetos a multas y recargos. La Intendencia de Montevideo podría tomar acciones administrativas y legales para hacer cumplir la normativa en los casos que correspondan.

¡Pague en fecha, evite multas y recargos!

Por consultas:

FONO TRIBUTOS
1950 3000

Figure A15: **Field Experiment: Flyer with text of informational intervention (Spanish) - Reminder + lottery + social benefit**

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

La Intendencia de Montevideo quiere premiar a los buenos pagadores. **Si usted paga en fecha participará automáticamente de un sorteo por la exoneración de un año de Contribución Inmobiliaria.**

Los sorteos se realizan todos los meses pares del año junto con la Lotería Nacional. Los beneficiados serán debidamente informados y se publicarán los resultados en el sitio web de la Intendencia.

La Intendencia de Montevideo realiza este sorteo para reconocer a los buenos pagadores por su contribución a la construcción de una ciudad más justa y mejor para todos/as.

Por consultas:

FONO TRIBUTOS
1950 3000

Figure A16: **Field Experiment: Flyer with text of informational intervention (Spanish) - Reminder + social punishment**

Intendencia de Montevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

Quienes no paguen en fecha podrían estar sujetos a multas y recargos. La Intendencia de Montevideo podría tomar acciones administrativas y legales para hacer cumplir la normativa en los casos que corresponda.

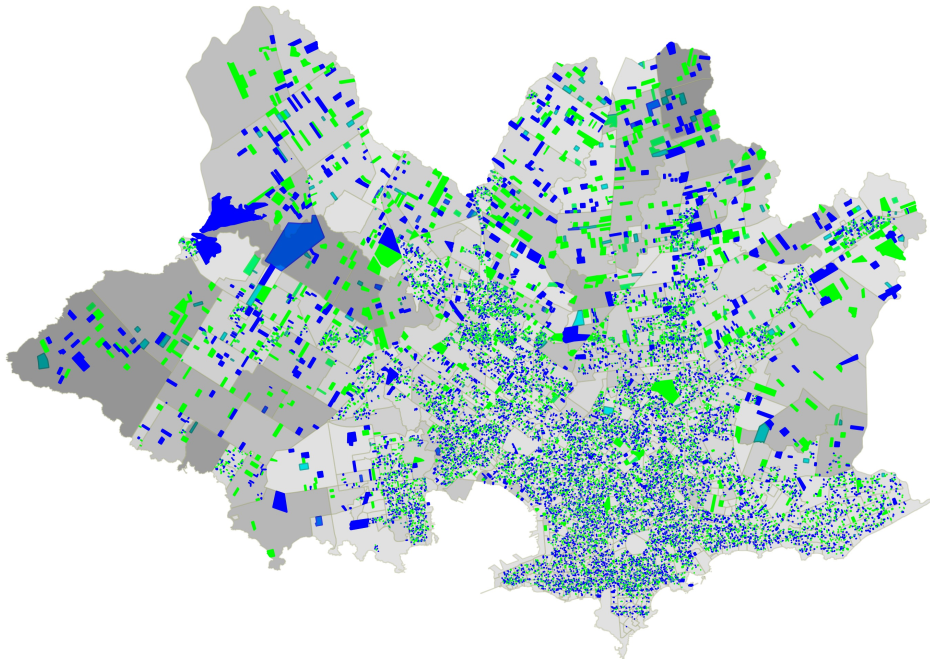
Las multas y recargos son una sanción para quienes no pagan sus impuestos y no contribuyen a la construcción de una ciudad más justa y mejor para todos/as.

Por consultas:

FONO TRIBUTOS
1950 3000

Figure A17: **Field Experiment: Informational intervention - Reverse side of flyers with municipal logo**

Figure A18: **Field Experiment: Geographic Distribution of Eligible and Ineligible Taxpayers.**



The figure depicts properties registered to eligible (in green, N=14,784) and ineligible (in blue, N=13,862) taxpayers in the field experiment. Eligible taxpayers are those that have been up to date during all of the previous year. Taxpayers in the field experiment sample are a random sample from the population of eligible and ineligible taxpayers. The grey background shows the level of development of the district. Darker grey denotes less developed areas. The map shows that eligible and ineligible taxpayers are similarly distributed across more and less developed areas.

Table A4: **Field Experiment: Pre-Treatment Balance.**

Treatment Comparison	Outcome	Estimate	SE	N	p-value
Treatment versus Pure Control	Paid on Time MAR 2010	0.001	0.007	16531	0.840
Treatment versus Placebo	Paid on Time MAR 2010	0.002	0.010	10022	0.838
Placebo versus Pure Control	Paid on Time MAR 2010	-0.001	0.009	13241	0.887
Treatment versus Pure Control	Paid on Time JUL 2010	0.004	0.007	16582	0.534
Treatment versus Placebo	Paid on Time JUL 2010	0.002	0.010	10060	0.847
Placebo versus Pure Control	Paid on Time JUL 2010	0.005	0.009	13282	0.564
Treatment versus Pure Control	Paid on Time NOV 2010	0.020	0.007	16616	0.007
Treatment versus Placebo	Paid on Time NOV 2010	-0.002	0.010	10091	0.857
Placebo versus Pure Control	Paid on Time NOV 2010	0.022	0.009	13297	0.012
Treatment versus Pure Control	Paid on Time MAR 2011	0.008	0.007	16896	0.236
Treatment versus Placebo	Paid on Time MAR 2011	-0.002	0.010	10245	0.809
Placebo versus Pure Control	Paid on Time MAR 2011	0.011	0.009	13507	0.215
Treatment versus Pure Control	Paid on Time JUL 2011	0.008	0.007	16937	0.284
Treatment versus Placebo	Paid on Time JUL 2011	-0.002	0.010	10270	0.824
Placebo versus Pure Control	Paid on Time JUL 2011	0.008	0.009	13541	0.336
Treatment versus Pure Control	Paid on Time NOV 2011	0.001	0.007	16984	0.896
Treatment versus Placebo	Paid on Time NOV 2011	-0.002	0.010	10295	0.851
Placebo versus Pure Control	Paid on Time NOV 2011	0.001	0.008	13583	0.890
Treatment versus Pure Control	Paid on Time MAR 2012	0.012	0.007	17234	0.098
Treatment versus Placebo	Paid on Time MAR 2012	-0.006	0.010	10447	0.550
Placebo versus Pure Control	Paid on Time MAR 2012	0.012	0.008	13777	0.153
Treatment versus Pure Control	Paid on Time JUL 2012	-0.001	0.007	17367	0.841
Treatment versus Placebo	Paid on Time JUL 2012	0.001	0.010	10494	0.934
Placebo versus Pure Control	Paid on Time JUL 2012	-0.006	0.009	13883	0.506
Treatment versus Pure Control	Paid on Time NOV 2012	-0.002	0.007	17378	0.733
Treatment versus Placebo	Paid on Time NOV 2012	-0.007	0.010	10503	0.467
Placebo versus Pure Control	Paid on Time NOV 2012	0.004	0.008	13895	0.605
Treatment versus Pure Control	Paid on Time MAR 2013	0.003	0.007	17567	0.653
Treatment versus Placebo	Paid on Time MAR 2013	-0.007	0.010	10598	0.450
Placebo versus Pure Control	Paid on Time MAR 2013	0.010	0.008	14065	0.235
Treatment versus Pure Control	Paid on Time JUL 2013	-0.007	0.007	17660	0.297
Treatment versus Placebo	Paid on Time JUL 2013	-0.008	0.010	10654	0.393
Placebo versus Pure Control	Paid on Time JUL 2013	-0.001	0.008	14134	0.914
Treatment versus Pure Control	Paid on Time NOV 2013	0.004	0.007	17686	0.532
Treatment versus Placebo	Paid on Time NOV 2013	0.001	0.010	10669	0.882
Placebo versus Pure Control	Paid on Time NOV 2013	0.003	0.008	14155	0.711
Treatment versus Pure Control	Paid on Time MAR 2014	-0.002	0.007	17808	0.806
Treatment versus Placebo	Paid on Time MAR 2014	-0.003	0.010	10778	0.749
Placebo versus Pure Control	Paid on Time MAR 2014	0.000	0.008	14244	0.968

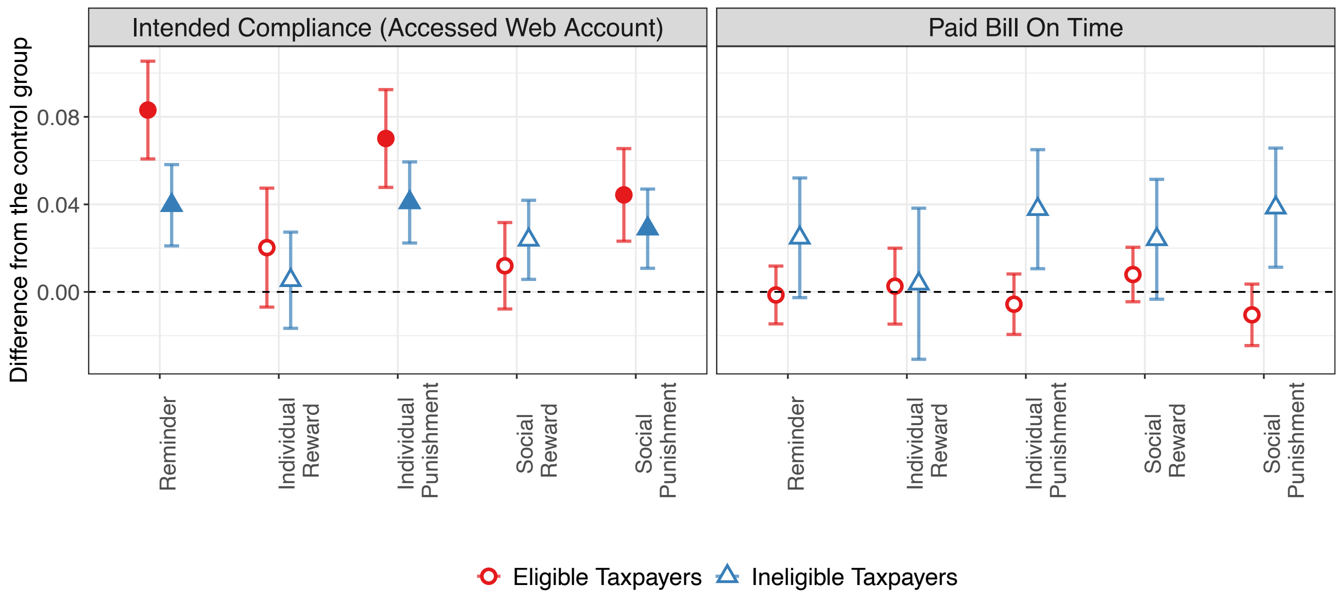
The table compares taxpayers assigned to the treatment and the pure control groups on pre-treatment payment history. To allow higher-powered balance tests, we include both eligible and ineligible taxpayers. Taxpayers in the treatment and the control group are statistically indistinguishable on these covariates, consistent with random assignment.

Table A5: **Field Experiment: Treatment Conditions and Sample Sizes (Full Experimental Design)**

Treatment condition	**Sample of eligibles (Good taxpayers)**	**Sample of ineligibles (Bad taxpayers)**
0. Pure Control	N=7,243	N=3,412
1. Reminder of Taxes Due (Placebo Control)	N=1,532	N=2,080
2A. Reminder + Lottery/Individual Reward	N=767	N=1,050
2B. Reminder + Lottery/Individual Reward + Probability of Winning	N=751	N=1,043
3. Reminder + Individual Punishment	N=1,465	N=2,109
4. Reminder + Lottery + Social Benefit*	N=1,519	N=2,057
5. Reminder + Social Punishment*	N=1,507	N=2,111
TOTAL N	N=14,784	N=13,862

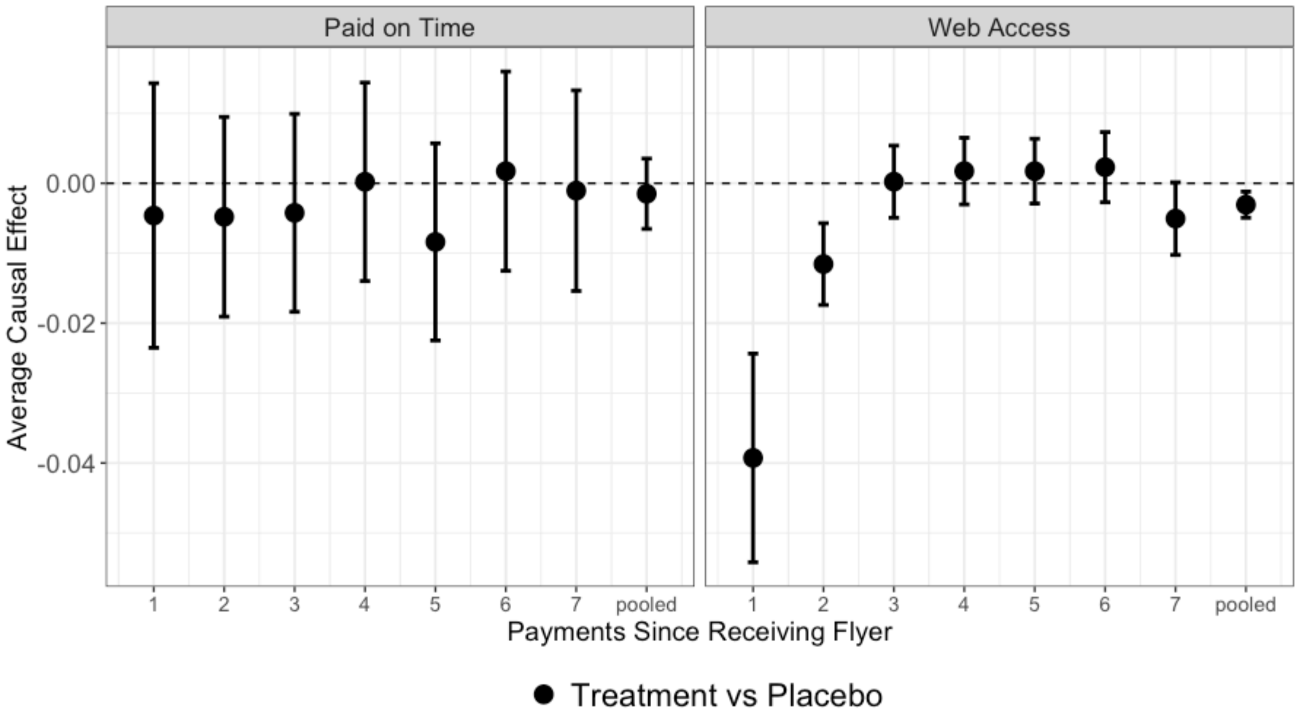
The figure depicts the full experimental design. The pure control and placebo control groups are described in the text. In the article, we pool conditions 2A, 2B, and 4 as the "Reminder + Information About Lottery" condition, since all three conditions inform taxpayers about the existence of the tax holiday. Conditions 2A and 2B prime individual benefits of winning the lottery, while condition 4 primes the social benefit; see our pre-analysis plan for a fuller description and Figures A12, A13, and A15 for the texts of the respective flyers. In the paper, we do not analyze conditions 3 (Figure A14) or 5 (Figure A16), which prime the individual and social aspects of punishment for non-payment of taxes. Total N=28,646.

Figure A19: **Field Experiment: Complete Results.**



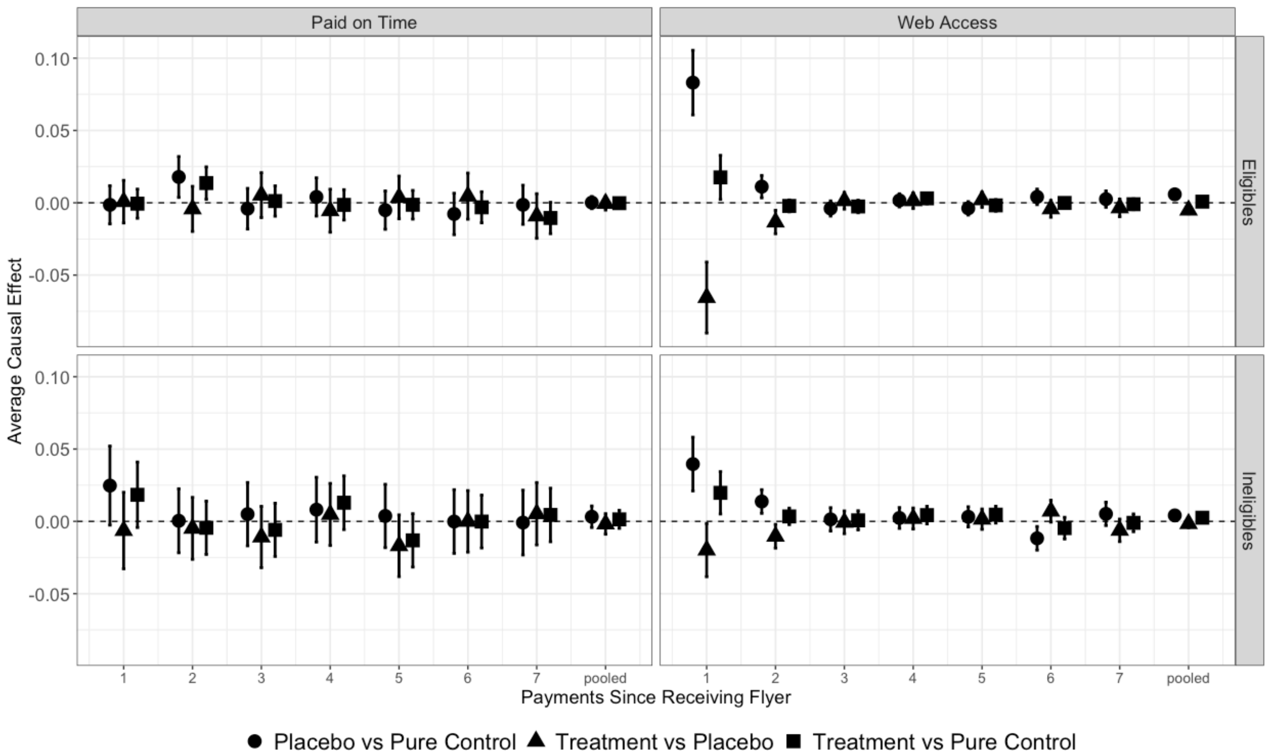
The figure shows the treatment effects for the five treatment levels described in table A5 (pooling treatment 2A and 2B as "Individual Reward") relative to a group that received no flyers (pure control). It depicts the effects for "good taxpayers" (those that have been up to date for the past year and are thus eligible to win the lottery; marked in red circles) and "bad taxpayers" (ineligible to win the lottery; marked in light blue triangles. In the left panel, Web Access measures whether the taxpayer paid the bill punctually in each payment period; on the right, Paid On Time measures whether the account holder accessed his or her online taxpayer account during the relevant period. Here, both outcomes are measured only for the payment period that immediately followed our intervention (July 2014). Bold estimates indicate that the nominal p-values survive both Bonferroni and FDR corrections. The results show that most interventions have effects on web access whereas only interventions highlighting the punishment have an effect on actual payment. However, this estimated positive effect holds only for bad taxpayers, and its statistical significance does not survive the multiple comparison corrections).

Figure A20: **Field Experiment: Effects of the Information Treatment v. the Placebo Control (no IPW)**



The figure depicts the effects of providing taxpayers with a reminder of an upcoming property tax due date plus additional information about the existence of the tax holiday lottery (Treatment) relative to a group that received a reminder only (Placebo). In the left panel, Paid On Time measures whether the taxpayer paid the bill punctually in each payment period; on the right, Web Access measures whether the account holder accessed his or her online taxpayer account during the relevant period. Both outcomes are measured for the seven payment periods that followed our intervention (which took place in June-July 2014). Here we pool estimates for taxpayers who were eligible to win the lottery at the time of intervention and those who were ineligible but who could become eligible if they brought their payments up to date. Here we do not adjust for differing assignment probabilities across ineligible and eligible taxpayers (i.e. do not use inverse probability weights—no IPW), but we compare the treatment group to the placebo control, as specified in our original pre-analysis plan (PAP) in F.1 below; see Table 3.2 and section 5 especially. Unequal probabilities of assignment of eligible and ineligible taxpayers arise particularly with comparisons to the pure control group, which we added after registering the PAP. The results show no effects of the information on tax payment and short-lived but negative effects on Web Access of the treatment relative to the placebo.

Figure A21: **Field Experiment: Effects of Information About the Tax Holiday on Compliance. Heterogeneous Treatment Effects by Taxpayer Type.**



The figure depicts the effects of providing taxpayers with a reminder of an upcoming property tax due date (Placebo) or a reminder of an upcoming property tax due date plus additional information about the existence of the tax holiday lottery (Treatment), relative to each other and to a group that received no flyers (Pure Control). In the left panel, Paid On Time measures whether the taxpayer paid the bill punctually in each payment period; in the right panel, Web Access measures whether the account holder accessed his or her online taxpayer account during the relevant period. Both outcomes are measured for the seven payment periods that followed our intervention, which took place in June-July 2014. The final period ("pooled") shows the estimates for the seven-period averages. The estimates for taxpayers who were eligible to win the lottery at the time of intervention are in the top panel while those who were ineligible but who could become eligible if they brought their payments up to date are in the bottom panel. The results show no effects of the information on tax payment and short-lived effects on Web Access for both subgroups, with more pronounced effects for eligible taxpayers. The placebo had a stronger effect than the treatment, which is why estimated effects are negative in the first two post-payment periods for Web Access.

Table A6: **Field Experiment: Effects of Information About the Tax Holiday on Compliance (weighted averages of block-specific effects for eligible and ineligible taxpayers)**

ATE	SE	N	p-value	Comparison
Outcome: Paid on time				
-0.0012	0.0023	10728	0.6059	Treatment vs Placebo
0.0004	0.0016	17750	0.7893	Treatment vs Pure Control
0.0013	0.0020	14184	0.5081	Placebo vs Pure Control
Outcome: Web Access				
-0.0042	0.0010	10799	0.0000	Treatment vs Placebo
0.0016	0.0007	17842	0.0223	Treatment vs Pure Control
0.0066	0.0009	14267	0.0000	Placebo vs Pure Control

D Survey experiment: Additional Analyses

Table A7: **Survey Experiment: Pooled Lottery vs. Discretionary Benefit Conditions**

	Difference	SE Diff	N	p-value	FDR reject	Bonferroni reject
Rewards go to the same people as always	1.05	0.22	1542	0.00	reject null	reject null
Worth it to be up to date	-0.48	0.15	2266	0.00	reject null	reject null
Mun. taxes are just	0.04	0.04	2291	0.16	do not reject	do not reject
Mun.gov. does a good job	-0.17	0.14	2313	0.23	do not reject	do not reject
Rewards are waste of money	-0.02	0.18	2234	0.90	do not reject	do not reject

The table shows effects in our survey experiment on five measures of attitudes towards taxation (see Figure 8 in the paper). Respondents in the "lottery" group were informed about the reward lottery using language similar to that printed on our mailed flyers in the field experiment. Respondents in the "discretionary" group were instead told that the municipality "from time to time" selects good taxpayers and rewards them with a year free of tax payment. The first column shows the difference between these groups; the second column shows the estimated SE for the difference; the next two columns show the N and the p-value for the difference. The last two columns indicate whether the effects survive multiple comparison adjustments considering all the tests specified in the pre-analysis plan.

E Simulation code

Figure A22: **Simulation Code.**

```
set.seed(12345)
library(ggplot2)

## SIMULATION
z <- 2 #tax payment
p <- 1 #probability of punishment
c <- 1 #cost of punishment
b <- .05 #intrinsic benefit of compliance

# function for the decision to comply in a single period
sim_function <- function(z, b, p, c, theta, lambda0){

  lambda <- NA
  for(i in 1:50){
    v <- rnorm(1) #random noise
    vector_theta <- (theta^((i+49):1))
    vector_lambda <- if(i==1){lambda0}else{c(lambda0, lambda)}
    lambda[i] <- ifelse((b + (p*c) - z + (1/5000)*z +
                           vector_theta %*% vector_lambda - v) > 0, 1, 0)}

  return(lambda)
}

# complete iteration including shock
sim_function1000 <- function(z, b, p,  c, theta, history, N, treat){
  sims <- matrix(NA, N, 50)
  for(i in 1:N){
    if (history=="perfect"){lambda0 <- rbinom(50,1,prob=1)}
    if (history=="marginal"){lambda0 <- rbinom(50,1,prob=.4)}
    if (treat==1) {lambda0[48:50] <- c(0, 0, 0)}
    sims[i, ] <- sim_function(z=z, b=b, p=p, c=c, theta=theta, lambda0=lambda0)}
  sims <- apply(sims, 2, mean)
  return(sims)
}

paym <- rbind(
  cbind(sim_function1000(z=z, b=b, p=p, c=c, theta=0.7, history = "perfect",
                         treat=1, N=10000), "Perfect Past Complier"),
  cbind(sim_function1000(z=z, b=b, p=p, c=c, theta=0.7, history = "marginal",
                         treat=1, N=10000), "Imperfect Past Complier"))

paym <- as.data.frame(paym)
paym$t <- c(1:50, 1:50)
names(paym)[1:2] <- c("mean_payments","type")
paym$type <- as.factor(paym$type)
paym$mean_payments <- as.numeric(as.character(paym$mean_payments))


ggplot(paym, aes(x=t, y=mean_payments, group=type)) +
  geom_line(aes(linetype = type)) +
  theme_bw() + ylim(0, 1) +
  ylab("Average Rate of Compliance") +
  xlab(expression(paste("Payment Periods ", italic("(t)")))) +
  labs(shape = "Taxpayer Type") +
  theme(legend.position = "bottom")
```

F Reproduction code (for all results in main text and Online Appendix Sections B-E)

In this section, we include all code necessary to reproduce the analysis in our main paper and those previously presented in this online appendix.

Disrupting Compliance: The Impact of a Randomized Tax Holiday in Uruguay

2023-05-06

```
# Basic setup --
rm(list=ls())
set.seed(1234)
options(scipen=999, digits=5)

##############################################
message("required libraries and setwd")
```

```
## required libraries and setwd
```

```
# Load/install packages --
if (!require("pacman")) install.packages("pacman")
```

```
## Loading required package: pacman
```

```
pacman::p_load(
  plyr,
  ggplot2,
  reshape2,
  zoo,
  sandwich,
  AER,
  xtable,
  stats,
  tidyr,
  dplyr,
  weights,
  estimatr
)

if (grepl ("/Users/gtunon", getwd ()) == TRUE){
  home <- "~/Dropbox/Working_papers/Uruguay_state_capacity/JOP_replication"
} else {
  # home <- " "
}
setwd(home)

##############################################
message("load data, functions, etc.")
```

```
## load data, functions, etc.
```



```
load("data/panel_taxtime.Rda")
load("data/cross_naturalex.Rda")
load("data/panel_goodtaxpayer.Rda")
load("data/naturalex_debt_gtp.Rda")
load("data/fieldex_data.Rda")
load("data/survey_data.Rda")

source("code/t_test.R")

################################################################################
# PAPER Tables & Figures
################################################################################
taxes_panel$YEARMON_LOTT <- as.yearmon(taxes_panel$FECHA_SORTEO2)
taxes_panel$YEAR_LOTT <- as.numeric(format(taxes_panel$YEARMON_LOTT, "%Y"))

taxes_panel$missed_payment <- as.numeric(taxes_panel$en_fecha==0)
taxes_panel$nr_parmntsowed <- taxes_panel$cuotas_adeudadas
taxes_panel$compliance <- as.numeric(taxes_panel$cuotas_adeudadas==0)


fieldex$tpooled <- NA
fieldex$tpooled[fieldex$treatment==6] <- "Control"
fieldex$tpooled[fieldex$treatment==0] <- "Reminder"
fieldex$tpooled[fieldex$treatment %in% c(1,2,4)] <- "Reminder+Info"

holiday <- taxes_panel[taxes_panel$cuota_exonerada==1,]

CI <- c(min(holiday$t[holiday$TRIBUTO=="Contribucion Inmobiliaria"]),
        max(holiday$t[holiday$TRIBUTO=="Contribucion Inmobiliaria"]))
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Contribucion Inmobiliaria" &
                               taxes_panel$ES_BP==1 &
                               taxes_panel$t>=CI[1] & taxes_panel$t<=CI[2]),]
taxes_panel$st[taxes_panel$TRIBUTO=="Contribucion Inmobiliaria"] <- CI[2] - 3
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Contribucion Inmobiliaria" &
                               taxes_panel$ES_BP==0 &
                               taxes_panel$t>=4 & taxes_panel$t<=CI[2]),]

PR <- c(min(holiday$t[holiday$TRIBUTO=="Patente de Rodados"]),
        max(holiday$t[holiday$TRIBUTO=="Patente de Rodados"]))
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Patente de Rodados" &
                               taxes_panel$ES_BP==1 &
                               taxes_panel$t>=PR[1] & taxes_panel$t<=PR[2]),]
taxes_panel$st[taxes_panel$TRIBUTO=="Patente de Rodados"] <- 0

TS <- c(min(holiday$t[holiday$TRIBUTO=="Saneamiento"]),
        max(holiday$t[holiday$TRIBUTO=="Saneamiento"]))
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Saneamiento" &
                               taxes_panel$ES_BP==1 &
                               taxes_panel$t>=TS[1] & taxes_panel$t<=TS[2]),]
taxes_panel$st[taxes_panel$TRIBUTO=="Saneamiento"] <- TS[2] - 3
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Saneamiento" &
                               taxes_panel$ES_BP==0 &
                               taxes_panel$t>=4 & taxes_panel$t<=TS[2]),]
```

```
TD <- c(1,max(holiday$t[holiday$TRIBUTO=="Tributos Domiciliarios"]))
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Tributos Domiciliarios" &
                              taxes_panel$ES_BP==1 &
                              taxes_panel$t>=TD[1] & taxes_panel$t<=TD[2]),]
taxes_panel$st[taxes_panel$TRIBUTO=="Tributos Domiciliarios"] <- TD[2] - 3
taxes_panel <- taxes_panel[!(taxes_panel$TRIBUTO=="Tributos Domiciliarios" &
                              taxes_panel$ES_BP==0 &
                              taxes_panel$t>=4 & taxes_panel$t<=TD[2]),]

taxes_panel$t_st <- ifelse((taxes_panel$ES_BP==1 & taxes_panel$t > 0) |
                             (taxes_panel$ES_BP==0 & taxes_panel$t > 4),
                           taxes_panel$t - taxes_panel$st, taxes_panel$t)
rm(holiday, CI, TD, TS, PR)

taxes_panel$holiday_type <- 1
taxes_panel$holiday_type[taxes_panel$TRIBUTO=="Patente de Rodados"] <- 0


taxes_panel$time_yearmon <- as.yearmon(taxes_panel$time, "%Y-%b-%d")
taxes_panel$bill_month <- format(taxes_panel$time_yearmon, "%b")

# add tax names in english
taxes_panel$tax <- taxes_panel$TRIBUTO
taxes_panel$tax <- as.factor(taxes_panel$tax)
levels(taxes_panel$tax) <-  c("Property", "Vehicle", "Sewage", "Head")
```

```
##############################################
message("MAIN PAPER: Figure 1")
```

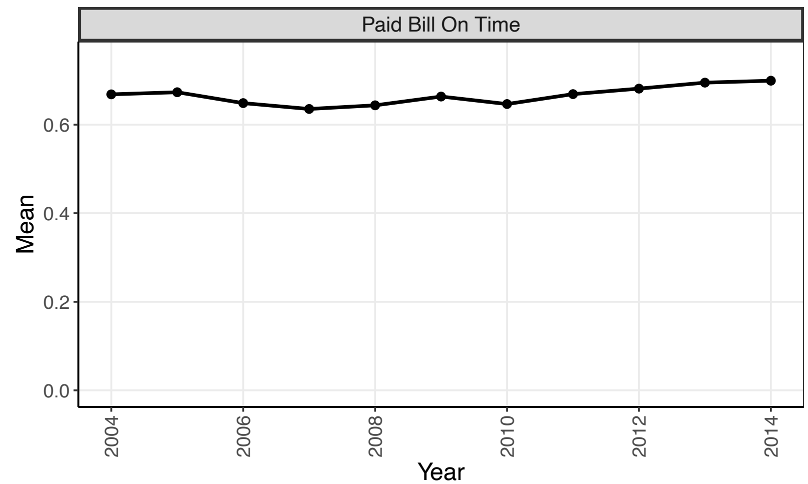
```
## MAIN PAPER: Figure 1
```

```
# Paid on time figure
taxes_panel %>%
  # We keep only observations in the control group
  filter(TREATMENT==0) %>%
  # group by year and take the mean of paid on time
  group_by(YEAR_LOTT) %>%
  dplyr::summarise(
  mean = mean(en_fecha, na.rm=T),
  outcome = "Paid Bill On Time"
) %>%
  ggplot(aes(YEAR_LOTT, mean)) +
  facet_wrap(~ outcome, scales="free_y") +
  geom_point(size=2) +
  geom_line(size=1) +
  xlab("Year") +
  ylab("Mean") +
  theme_bw() + ylim(c(0,.75)) +
  scale_colour_manual(values = c("black")) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
```



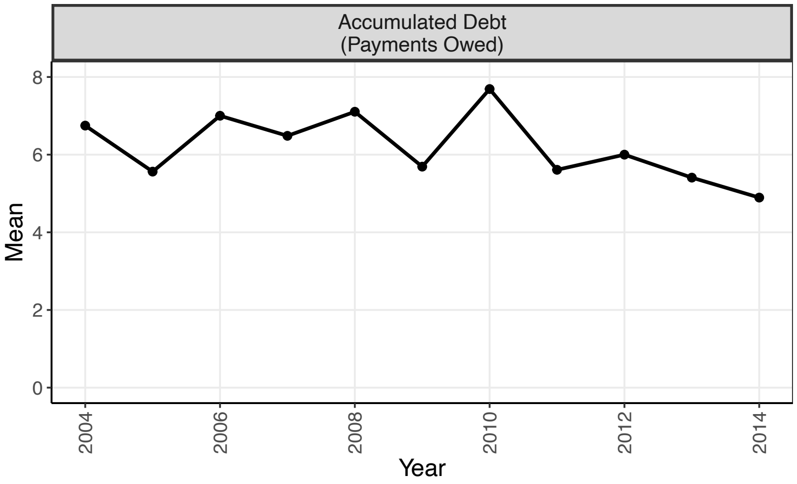
```
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
        axis.text.x = element_text(size = rel(1.25), angle = 90, vjust = 0.5, hjust=1),
        #axis.text.x = element_text(size = rel(1.1), hjust=.7),
        legend.position = "bottom",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## Warning: The `size` argument of `element_rect()` is deprecated as of ggplot2 3.4.0.
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

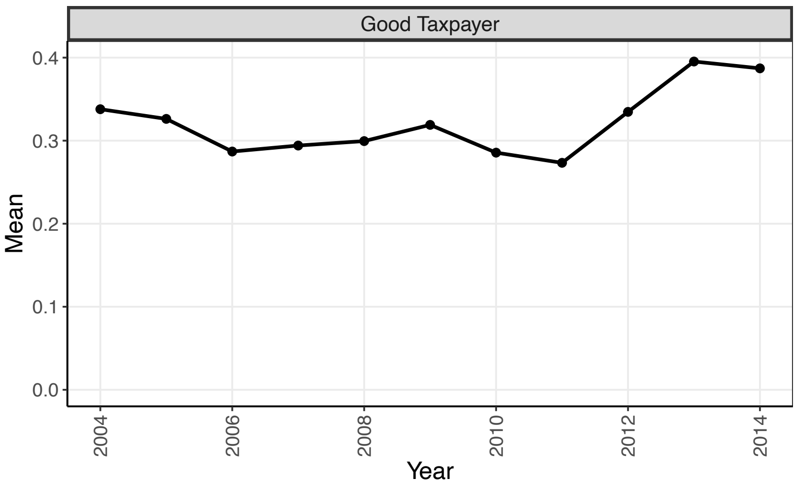



```
# Accumulated Debt figure
taxes_panel %>%
  # We keep only observations in the control group
```

```r
  filter(TREATMENT==0) %>%
  # group by year and take the mean of accumulated debt
  group_by(YEAR_LOTT) %>%
  dplyr::summarise(
    mean = mean(cuotas_adeudadas, na.rm=T),
    outcome = "Accumulated Debt\n(Payments Owed)"
  ) %>%
  ggplot(aes(YEAR_LOTT, mean)) +
  facet_wrap(~ outcome, scales="free_y") +
  geom_point(size=2) +
  geom_line(size=1) +
  xlab("Year") +
  ylab("Mean") +
  theme_bw() + ylim(c(0,8)) +
  scale_colour_manual(values = c("black")) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
        axis.text.x = element_text(size = rel(1.25), angle = 90, vjust = 0.5, hjust=1),
        legend.position = "bottom",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```r
# Good taxpayer figure
taxes_panel %>%
  # We keep only observations in the control group and observations for t=0
  filter(TREATMENT==0 & t==0) %>%
  # group by year and take the mean of accumulated debt
  group_by(YEAR_LOTT) %>%
  dplyr::summarise(
    mean = mean(ES_BP, na.rm=T),
    outcome = "Good Taxpayer"
  ) %>%
  ggplot(aes(YEAR_LOTT, mean)) +
  facet_wrap(~ outcome, scales="free_y") +
  geom_point(size=2) +
  geom_line(size=1) +
  xlab("Year") +
  ylab("Mean") +
  theme_bw() + ylim(c(0,.4)) +
  scale_colour_manual(values = c("black")) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
        axis.text.x = element_text(size = rel(1.25), angle = 90, vjust = 0.5, hjust=1),
        legend.position = "bottom",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```
################################################
message("MAIN PAPER: Table 3.1")
```

```
## MAIN PAPER: Table 3.1
```

```
buen_pagador_panel %>%
  # keep only the cross-section of observations at the time of treatment assignment
  filter(centered==0) %>%
  dplyr::group_by(ES_BP, TRIBUTO) %>%
  dplyr::summarize(
    winning_accounts = sum(TREATMENT==1),
    non_winning_accounts = sum(TREATMENT==0),
    study_group = n()
  )
```

```
## `summarise()` has grouped output by 'ES_BP'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 8 x 5
## # Groups:   ES_BP [2]
##   ES_BP TRIBUTO                winning_accounts non_winning_accounts study_group
##   <int> <fct>                             <int>                <int>       <int>
## 1     0 Contribucion Inmobili~             1211                 1225        2436
## 2     0 Patente de Rodados                 1899                 1924        3823
## 3     0 Saneamiento                         915                  939        1854
## 4     0 Tributos Domiciliarios             2083                 2062        4145
## 5     1 Contribucion Inmobili~             1339                 1354        2693
```



```
## 6     1 Patente de Rodados                  391                  375         766
## 7     1 Saneamiento                         452                  404         856
## 8     1 Tributos Domiciliarios             1007                 1041        2048
```

```
buen_pagador_panel %>%
  # keep only the cross-section of observations at the time of treatment assignment
  filter(centered==0) %>%
  dplyr::group_by(ES_BP) %>%
  dplyr::summarize(
    winning_accounts = sum(TREATMENT==1),
    non_winning_accounts = sum(TREATMENT==0),
    study_group = n()
  )
```

```
## # A tibble: 2 x 4
##   ES_BP winning_accounts non_winning_accounts study_group
##   <int>            <int>                <int>       <int>
## 1     0             6108                 6150       12258
## 2     1             3189                 3174        6363
```

```
################################################
message("MAIN PAPER: Table 3.2")
```

```
## MAIN PAPER: Table 3.2
```

```
# table by treatment and taxpayer type
table(fieldex$tpooled, fieldex$type)
```

```
##
##                  good taxpayer bad taxpayer
##   Control                 7243         3412
##   Reminder                1532         2080
##   Reminder+Info           3037         4150
```

```
# table by taxpayer type
table(fieldex$type)
```

```
##
## good taxpayer  bad taxpayer
##         14784         13862
```

```
################################################
message("MAIN PAPER: Figure 2")
```

```
## MAIN PAPER: Figure 2
```

```
table(taxes_panel$TREATMENT, taxes_panel$TIENE_EXO)
```

```
##
##          0      1
##   0 550896      0
##   1 421618 128276
```

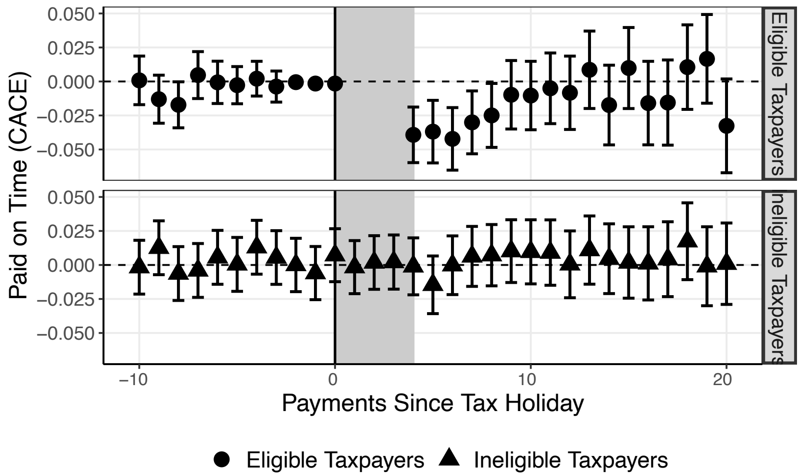
```r
t <- -10:20
btp <- NULL
gtp <- NULL

for (i in t){

  temp <- taxes_panel[taxes_panel$t_st == i, ]

  btp_est <- tidy(lm_robust(en_fecha ~ TREATMENT,
                            data = filter(temp, ES_BP==0)))[2,]
  btp <- rbind.data.frame(btp, cbind.data.frame(i, "Ineligible Taxpayers", btp_est))

  # For good taxpayers, skip periods under the tax holiday
  if (nrow(temp[temp$TREATMENT==1 & temp$ES_BP==1,])==0) next

  gtp_est <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                            data = filter(temp, ES_BP==1)))[2,]
  gtp <- rbind.data.frame(gtp, cbind.data.frame(i, "Eligible Taxpayers", gtp_est))

}
```

```
## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
```

```r
names(gtp)[2] <- names(btp)[2] <- "type"
plot <- rbind.data.frame(gtp, btp); rm(gtp, btp)

ggplot(plot, aes(x=i, y=estimate, shape=type)) +
facet_grid(type~.) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=i,
                    ymin=conf.low,
                    ymax=conf.high),
                width=.6, size=.8, position = position_dodge(width = 0.6)) +
  geom_point(size=4, position = position_dodge(width = 0.6)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Paid on Time (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  #scale_x_discrete(limit = t[!t %in% c(1,2,3)],
  #                 labels = as.character(c(t[t<1],
  #                                         t[!t%in%c(1,2,3) & t>0]-3))) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
```



```r
        axis.text.x = element_text(size = rel(1.1), hjust=.7),
        legend.position = "bottom",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```r
################################################
message("MAIN DOC: Table 4.3")
```

```
## MAIN DOC: Table 4.3
```

```r
# estimate for 1st period after holiday (t_st==4)
t1_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                         data = filter(taxes_panel, ES_BP==1 & t_st==4)))
t1_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                          data = filter(taxes_panel, ES_BP==1 & t_st==4)))
t1 <- c(t1_ace[1,2], t1_ace[2,2], t1_ace[2,3], t1_ace[2,5], t1_cace[2,2]);
rm(t1_ace, t1_cace)

# estimate for 5th period after holiday (t_st==8)
t5_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                         data = filter(taxes_panel, ES_BP==1 & t_st==8)))
t5_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                          data = filter(taxes_panel, ES_BP==1 & t_st==8)))
t5 <- c(t5_ace[1,2], t5_ace[2,2], t5_ace[2,3], t5_ace[2,5], t5_cace[2,2]);
rm(t5_ace, t5_cace)
```

```r
# estimate for 10th period after holiday (t_st==13)
t10_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                         data = filter(taxes_panel, ES_BP==1 & t_st==13)))
t10_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                          data = filter(taxes_panel, ES_BP==1 & t_st==13)))
t10 <- c(t10_ace[1,2], t10_ace[2,2], t10_ace[2,3], t10_ace[2,5], t10_cace[2,2]);
rm(t10_ace, t10_cace)

# estimate for periods 1-10 after holiday (t_st==4-13)
t1_10_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                          data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14)))
t1_10_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                           data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14)))
t1_10 <- c(t1_10_ace[1,2], t1_10_ace[2,2], t1_10_ace[2,3], t1_10_ace[2,5],
           t1_10_cace[2,2]);
rm(t1_10_ace, t1_10_cace)


# estimate for periods 1-10 after holiday (t_st==4-13) PROPERTY TAX
t1_10_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                            data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                          TRIBUTO == "Contribucion Inmobiliaria")))
t1_10_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                             data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                           TRIBUTO == "Contribucion Inmobiliaria")))
t1_10_property <- c(t1_10_ace[1,2], t1_10_ace[2,2], t1_10_ace[2,3], t1_10_ace[2,5],
                    t1_10_cace[2,2]);
rm(t1_10_ace, t1_10_cace)


# estimate for periods 1-10 after holiday (t_st==4-13) HEAD
t1_10_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                            data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                          TRIBUTO == "Tributos Domiciliarios")))
t1_10_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                             data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                           TRIBUTO == "Tributos Domiciliarios")))
t1_10_head <- c(t1_10_ace[1,2], t1_10_ace[2,2], t1_10_ace[2,3], t1_10_ace[2,5],
                t1_10_cace[2,2]);
rm(t1_10_ace, t1_10_cace)


# estimate for periods 1-10 after holiday (t_st==4-13) SEWAGE
t1_10_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                            data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                          TRIBUTO == "Saneamiento")))
t1_10_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                             data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                           TRIBUTO == "Saneamiento")))
t1_10_sewage <- c(t1_10_ace[1,2], t1_10_ace[2,2], t1_10_ace[2,3], t1_10_ace[2,5],
                  t1_10_cace[2,2]);
rm(t1_10_ace, t1_10_cace)
```



```r
# estimate for periods 1-10 after holiday (t_st==4-13) VEHICLE
t1_10_ace <- tidy(lm_robust(en_fecha ~ TREATMENT,
                            data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                          TRIBUTO == "Patente de Rodados")))
t1_10_cace <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                             data = filter(taxes_panel, ES_BP==1 & t_st>3 & t_st<14 &
                                           TRIBUTO == "Patente de Rodados")))
t1_10_vehicle <- c(t1_10_ace[1,2], t1_10_ace[2,2], t1_10_ace[2,3], t1_10_ace[2,5],
                   t1_10_cace[2,2]);
rm(t1_10_ace, t1_10_cace)



# Build table
table4.3 <- rbind.data.frame(t1,t5,t10,t1_10, t1_10_property, t1_10_head,
                             t1_10_sewage, t1_10_vehicle);
rm(t1,t5,t10,t1_10, t1_10_property, t1_10_head, t1_10_sewage, t1_10_vehicle)

rownames(table4.3) <- c("Post Tax Holiday Payment 1",
                     "Post Tax Holiday Payment 5",
                     "Post Tax Holiday Payment 10",
                     "Post Tax Holiday Payments 1-10",
                     "Post Tax Holiday Payments 1-10 (Property)",
                     "Post Tax Holiday Payments 1-10 (Head)",
                     "Post Tax Holiday Payments 1-10 (Sewage)",
                     "Post Tax Holiday Payments 1-10 (Vehicle)")

colnames(table4.3) <- c("Control Mean", "ACE", "SE_ACE", "p-value", "CACE")

table4.3
```

```
##                                            Control Mean        ACE    SE_ACE
## Post Tax Holiday Payment 1                      0.92802 -0.0286352 0.0075862
## Post Tax Holiday Payment 5                      0.90544 -0.0181432 0.0087125
## Post Tax Holiday Payment 10                     0.86822  0.0062235 0.0105257
## Post Tax Holiday Payments 1-10                  0.89809 -0.0155211 0.0028548
## Post Tax Holiday Payments 1-10 (Property)       0.92140 -0.0225692 0.0041321
## Post Tax Holiday Payments 1-10 (Head)           0.90578 -0.0119218 0.0046971
## Post Tax Holiday Payments 1-10 (Sewage)         0.93520 -0.0220321 0.0063792
## Post Tax Holiday Payments 1-10 (Vehicle)        0.75081  0.0029792 0.0115982
##                                                   p-value       CACE
## Post Tax Holiday Payment 1                 0.000161942551 -0.0392398
## Post Tax Holiday Payment 5                 0.037355649073 -0.0249482
## Post Tax Holiday Payment 10                0.554376596143  0.0085857
## Post Tax Holiday Payments 1-10             0.000000054477 -0.0212912
## Post Tax Holiday Payments 1-10 (Property)  0.000000047673 -0.0296383
## Post Tax Holiday Payments 1-10 (Head)      0.011153772130 -0.0173794
## Post Tax Holiday Payments 1-10 (Sewage)    0.000556227175 -0.0332471
## Post Tax Holiday Payments 1-10 (Vehicle)   0.797289559027  0.0035880
```

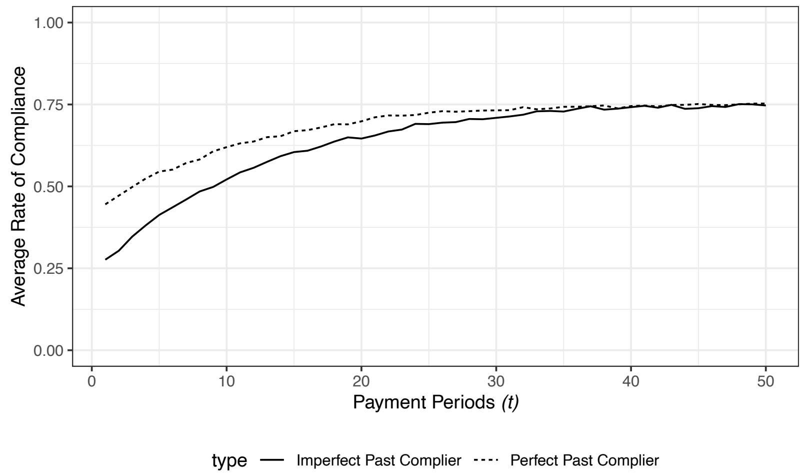
```r
##############################################
message("MAIN PAPER: Figure 3")
```

MAIN PAPER: Figure 3

```
set.seed(12345)

## SIMULATION
z <- 2 #tax payment
p <- 1 #probability of punishment
c <- 1 #cost of punishment
b <- .05 #intrinsic benefit of compliance

# function for the decision to comply in a single period
sim_function <- function(z, b, p, c, theta, lambda0){

  lambda <- NA
  for(i in 1:50){
    v <- rnorm(1) #random noise
    vector_theta <- (theta^((i+49):1))
    vector_lambda <- if(i==1){lambda0}else{c(lambda0, lambda)}
    lambda[i] <- ifelse((b + (p*c) - z + (1/5000)*z +
                       vector_theta %*% vector_lambda - v) > 0, 1, 0)}

  return(lambda)
}

# complete iteration including shock
sim_function1000 <- function(z, b, p,  c, theta, history, N, treat){
  sims <- matrix(NA, N, 50)
  for(i in 1:N){
    if (history=="perfect"){lambda0 <- rbinom(50,1,prob=1)}
    if (history=="marginal"){lambda0 <- rbinom(50,1,prob=.4)}
    if (treat==1) {lambda0[48:50] <- c(0, 0, 0)}
    sims[i, ] <- sim_function(z=z, b=b, p=p, c=c, theta=theta, lambda0=lambda0)}
  sims <- apply(sims, 2, mean)
  return(sims)
}

paym <- rbind(
  cbind(sim_function1000(z=z, b=b, p=p, c=c, theta=0.7, history = "perfect",
                         treat=1, N=10000), "Perfect Past Complier"),
  cbind(sim_function1000(z=z, b=b, p=p, c=c, theta=0.7, history = "marginal",
                         treat=1, N=10000), "Imperfect Past Complier"))

paym <- as.data.frame(paym)
paym$t <- c(1:50, 1:50)
names(paym)[1:2] <- c("mean_payments","type")
paym$type <- as.factor(paym$type)
paym$mean_payments <- as.numeric(as.character(paym$mean_payments))


ggplot(paym, aes(x=t, y=mean_payments, group=type)) +
  geom_line(aes(linetype = type)) +
  theme_bw() + ylim(0, 1) +
  ylab("Average Rate of Compliance") +
  xlab(expression(paste("Payment Periods ", italic("(t)")))) +
```



```
  labs(shape = "Taxpayer Type") +
  theme(legend.position = "bottom")
```




```
################################################
message("MAIN PAPER: Figure 4")
```

MAIN PAPER: Figure 4

```
gtp_taxes <- taxes_panel[taxes_panel$ES_BP==1, ]
t <- unique(gtp_taxes$t_st)
t <- t[order(t)]
t <- t[t>-11 & t<=21]
gtp_plot <- NULL

for (i in 1:length(t)){

  temp <- gtp_taxes[gtp_taxes$t_st == t[i], ]

  if (nrow(temp[temp$TREATMENT==1 & temp$holiday_type==0,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$holiday_type==0,])==0) next

  on_time_noholiday <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                                      data=temp[temp$holiday_type==0,]))
  on_time_noholiday <- c(on_time_noholiday$estimate[2], on_time_noholiday$std.error[2])

  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], on_time_noholiday, "No Holiday"))
```

```
  )

}

for (i in 1:length(t)){

  temp <- gtp_taxes[gtp_taxes$t_st == t[i], ]

  if (nrow(temp[temp$TREATMENT==1 & temp$holiday_type==1,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$holiday_type==1,])==0) next

  on_time_holiday <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                          data=temp[temp$holiday_type==1,]))
  on_time_holiday <- c(on_time_holiday$estimate[2], on_time_holiday$std.error[2])


  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], on_time_holiday, "Holiday"))
  )

}
```

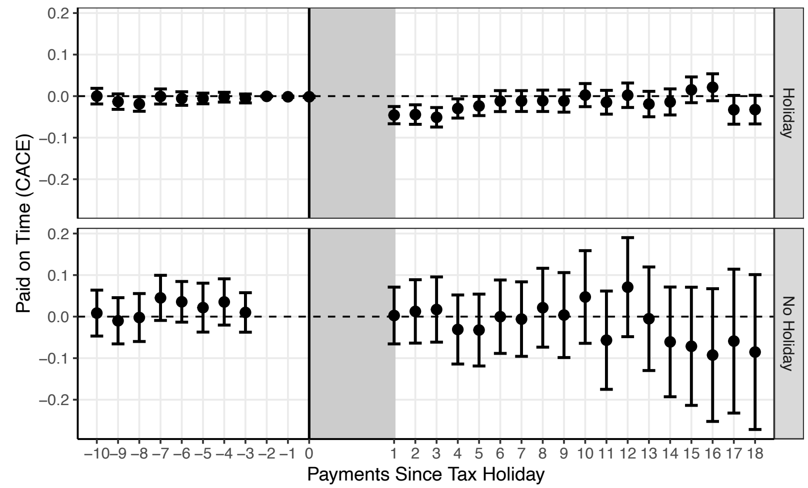
```
## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
```

```
gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "CACE", "SE", "lottery_type")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))
gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$SE <- as.numeric(as.character(gtp_plot$SE))
gtp_plot$upper <- gtp_plot$CACE + qnorm(.975) * gtp_plot$SE
gtp_plot$lower <- gtp_plot$CACE - qnorm(.975) * gtp_plot$SE

p <- ggplot(gtp_plot, aes(x=t, y=CACE))
p + facet_grid(lottery_type~.) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.5)) +
  geom_point(size=2.5, position = position_dodge(width = 0.5)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Paid on Time (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  scale_alpha_manual(values = c("FALSE"=0.8, "TRUE"=1), guide='none') +
  scale_x_discrete(limit = t[!t%in%c(1,2,3)],
                   labels = as.character(c(t[t<1],
```



```
                                           t[!t%in%c(1,2,3) & t>0]-3))) +
  theme(legend.position = "none",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```

```
## Warning: Continuous limits supplied to discrete scale.
## i Did you mean 'limits = factor(...)' or 'scale_*_continuous()'?
```




```
#################################################
message("MAIN PAPER: Figure 5")
```

```
## MAIN PAPER: Figure 5
```

```
t <- unique(taxes_panel$t_st)
t <- t[order(t)]
t <- t[t>-11 & t<=28]
gtp_plot <- NULL

gtp_taxes <- taxes_panel[taxes_panel$ES_BP==1, ]
gtp_taxes <- gtp_taxes[gtp_taxes$TRIBUTO!="Patente de Rodados",]

for (i in 1:length(t)){

  temp <- gtp_taxes[gtp_taxes$t_st == t[i], ]

  if (nrow(temp[temp$TREATMENT==1,])==0) next
```

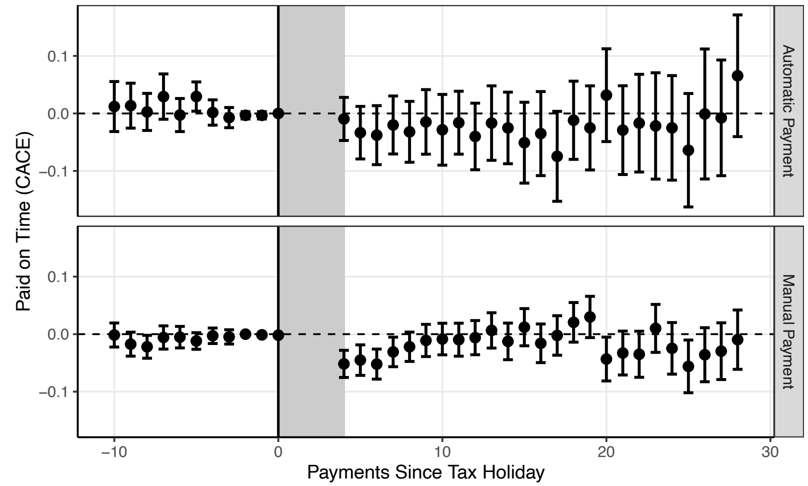
```
  if (nrow(temp[temp$TREATMENT==0,])==0) next


  on_time_auto <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                                 data=temp[temp$autopay_win==1,]))
  on_time_auto <- c(on_time_auto$estimate[2], on_time_auto$std.error[2])

   on_time_manual <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                                 data=temp[temp$autopay_win==0,]))
  on_time_manual <- c(on_time_manual$estimate[2], on_time_manual$std.error[2])


  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], on_time_auto,
                                "Automatic Payment")),
                    as.vector(c("Paid on Time", t[i], on_time_manual,
                                "Manual Payment"))
  )

}
```

```
## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
```

```
gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "CACE", "SE", "sample")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))
gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$SE <- as.numeric(as.character(gtp_plot$SE))
gtp_plot$upper <- gtp_plot$CACE + qnorm(.975) * gtp_plot$SE
gtp_plot$lower <- gtp_plot$CACE - qnorm(.975) * gtp_plot$SE

p <- ggplot(gtp_plot, aes(x=t, y=CACE))
p + facet_grid(sample~.) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.5)) +
  geom_point(size=2.5, position = position_dodge(width = 0.5)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Paid on Time (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  scale_alpha_manual(values = c("FALSE"=0.8, "TRUE"=1), guide='none') +
  #scale_x_discrete(limit = t[!t%in%c(1,2,3)],
  #                 labels = as.character(c(t[t<1],
  #                                         t[!t%in%c(1,2,3) & t>0]-3))) +
  theme(legend.position = "none",
        legend.title=element_blank(),
```



```
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```
###############################################
message("MAIN PAPER: Figure 6")
```

```
## MAIN PAPER: Figure 6
```

```
taxpayer_type <- ddply(gtp_taxes[gtp_taxes$t_st<=0 & gtp_taxes$t_st>-15,],
                       "CUENTA", summarise,
                       type = mean(en_fecha, na.rm=T))

gtp_taxes <- merge(gtp_taxes, taxpayer_type, by="CUENTA", all.x=T)

t <- unique(gtp_taxes$t_st)
t <- t[order(t)]
t <- t[t>-11 & t<=21]
gtp_plot <- NULL

for (i in 1:length(t)){

  temp <- gtp_taxes[gtp_taxes$t_st == t[i], ]

  if (nrow(temp[temp$TREATMENT==1 & temp$type!=1,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$type!=1,])==0) next

  ivest <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
```

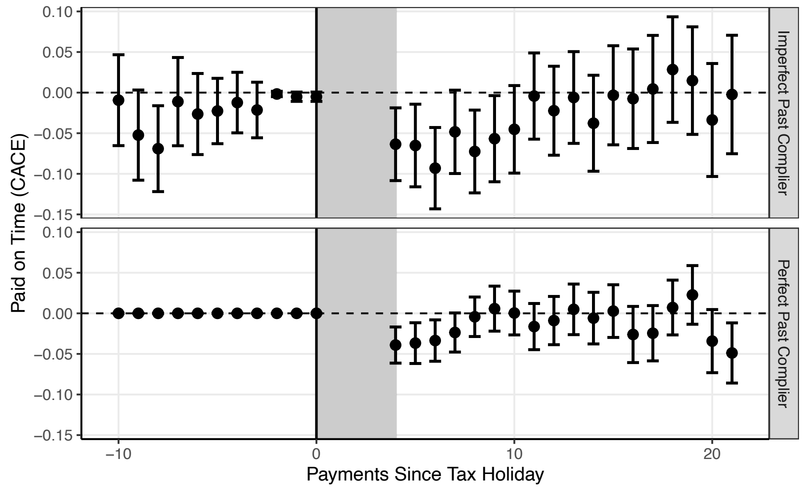
```r
                             data=temp[temp$type!=1,]))
  on_time_marginal <- c(ivest$estimate[2], ivest$std.error[2])


  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i],
                                on_time_marginal, "Imperfect Past Complier"))
  )

}
```

```
## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
```

```r
for (i in 1:length(t)){

  temp <- gtp_taxes[gtp_taxes$t_st == t[i], ]

  if (nrow(temp[temp$TREATMENT==1 & temp$type==1,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$type==1,])==0) next

  ivest <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                          data=temp[temp$type==1,]))
  on_time_compliant <- c(ivest$estimate[2], ivest$std.error[2])


  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i],
                                on_time_compliant, "Perfect Past Complier")))

}

gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "CACE", "SE", "lottery_type")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))
gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$SE <- as.numeric(as.character(gtp_plot$SE))
gtp_plot$upper <- gtp_plot$CACE + qnorm(.975) * gtp_plot$SE
gtp_plot$lower <- gtp_plot$CACE - qnorm(.975) * gtp_plot$SE

p <- ggplot(gtp_plot, aes(x=t, y=CACE))
p + facet_grid(lottery_type~.) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.5)) +
  geom_point(size=2.5, position = position_dodge(width = 0.5)) +
```



```r
  xlab("Payments Since Tax Holiday") +
  ylab("Paid on Time (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  scale_alpha_manual(values = c("FALSE"=0.8, "TRUE"=1), guide='none') +
  #scale_x_discrete(limit = t[!t%in%c(1,2,3)],
  #                 labels = as.character(c(t[t<1],
  #                                         t[!t%in%c(1,2,3) & t>0]-3))) +
  theme(legend.position = "none",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```r
###############################################
message("MAIN PAPER: Figure 7")
```

```
## MAIN PAPER: Figure 7
```

```r
# # Weigths for treatment - placebo comparison
table(fieldex$type, fieldex$pooled_124_0)
```

```
##
##                     0    1
##   good taxpayer 1532 3037
##   bad taxpayer  2080 4150
```

```r
table.weights <- 1/prop.table(table(fieldex$type, fieldex$pooled_124_0), 1)
table.weights <- melt(table.weights)
names(table.weights) <- c("type", "pooled_124_0", "pooled_124_0_wts")
fieldex <- fieldex %>% left_join(table.weights)
```

```
## Joining with `by = join_by(type, pooled_124_0)`
```

```r
# # Weights for treatment versus pure control
table(fieldex$type, fieldex$pooled_124_6)
```

```
##
##                      0    1
##   good taxpayer 7243 3037
##   bad taxpayer  3412 4150
```

```r
table.weights <- 1/prop.table(table(fieldex$type, fieldex$pooled_124_6), 1)
table.weights <- melt(table.weights)
names(table.weights) <- c("type", "pooled_124_6", "pooled_124_6_wts")
fieldex <- fieldex %>% left_join(table.weights)
```

```
## Joining with `by = join_by(type, pooled_124_6)`
```

```r
# # Weights for placebo versus pure control
table(fieldex$type, fieldex$pooled_0_6)
```

```
##
##                      0    1
##   good taxpayer 7243 1532
##   bad taxpayer  3412 2080
```

```r
table.weights <- 1/prop.table(table(fieldex$type, fieldex$pooled_0_6), 1)
table.weights <- melt(table.weights)
names(table.weights) <- c("type", "pooled_0_6", "pooled_0_6_wts")
fieldex <- fieldex %>% left_join(table.weights)
```

```
## Joining with `by = join_by(type, pooled_0_6)`
```

```r
# Add variable to summarize compliance between 2014-2016
fieldex <-  fieldex %>% mutate(
  compliance_1416 = (JUL_2014_ontime + NOV_2014_ontime +
                     MAR_2015_ontime + JUL_2015_ontime + NOV_2015_ontime +
                     MAR_2016_ontime + JUL_2016_ontime)/12,
  intended_1416 <- (JUL_2014_WEBACCESS + NOV_2014_WEBACCESS +
                    MAR_2015_WEBACCESS + JUL_2015_WEBACCESS + NOV_2015_WEBACCESS +
                    MAR_2016_WEBACCESS + JUL_2016_WEBACCESS)/12
)


#### TREATMENT VERSUS PURE CONTROL
# a) compliance
```



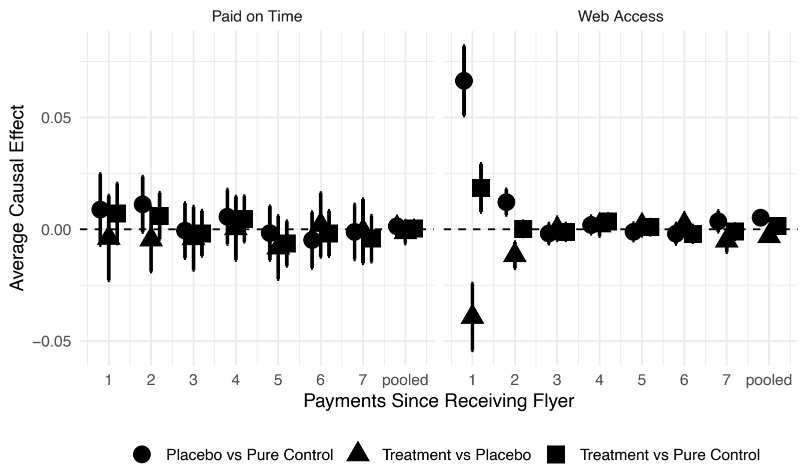
```r
comp_control <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(NOV_2014_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(MAR_2015_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(JUL_2015_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(NOV_2015_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(MAR_2016_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(JUL_2016_ontime ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(compliance_1416 ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex)
)
comp_control$outcome <- "Paid on Time"
comp_control$control <- "Treatment vs Pure Control"
# b) Intended compliance
intcomp_control <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex),
  difference_in_means(intended_1416 ~ pooled_124_6, weights = pooled_124_6_wts,
                      data = fieldex)
)
intcomp_control$outcome <- "Web Access"
intcomp_control$control <- "Treatment vs Pure Control"


#### TREATMENT VERSUS PLACEBO
# a) Compliance
comp_placebo <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                      data = fieldex),
  difference_in_means(NOV_2014_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                      data = fieldex),
  difference_in_means(MAR_2015_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                      data = fieldex),
  difference_in_means(JUL_2015_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
```

```r
                    data = fieldex),
  difference_in_means(NOV_2015_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                    data = fieldex),
  difference_in_means(MAR_2016_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                    data = fieldex),
  difference_in_means(JUL_2016_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                    data = fieldex),
  difference_in_means(compliance_1416 ~ pooled_124_0, weights = pooled_124_0_wts,
                    data = fieldex)
)
comp_placebo$outcome <- "Paid on Time"
comp_placebo$control <- "Treatment vs Placebo"
# b) Intended compliance
intcomp_placebo <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex),
  difference_in_means(intended_1416 ~ pooled_124_0,
                    weights = pooled_124_0_wts, data = fieldex)
)
intcomp_placebo$outcome <- "Web Access"
intcomp_placebo$control <- "Treatment vs Placebo"


#### PURE CONTROL VERSUS PLACEBO
# a) Compliance
comp_control_pla <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(NOV_2014_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(MAR_2015_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(JUL_2015_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(NOV_2015_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(MAR_2016_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(JUL_2016_ontime ~ pooled_0_6, weights = pooled_0_6_wts,
                    data = fieldex),
  difference_in_means(compliance_1416 ~ pooled_0_6, weights = pooled_0_6_wts,
```



```r
                    data = fieldex)
)
comp_control_pla$outcome <- "Paid on Time"
comp_control_pla$control <- "Placebo vs Pure Control"
# b) Intended compliance
intcomp_control_pla <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex),
  difference_in_means(intended_1416 ~ pooled_0_6,
                    weights = pooled_0_6_wts, data = fieldex)
)
intcomp_control_pla$outcome <- "Web Access"
intcomp_control_pla$control <- "Placebo vs Pure Control"

## Combine and plot
plotdata <- rbind.data.frame(comp_placebo, comp_control, comp_control_pla,
                             intcomp_placebo, intcomp_control,
                             intcomp_control_pla)
rm(comp_placebo, comp_control,
   intcomp_placebo, intcomp_control,
   comp_control_pla, intcomp_control_pla)

plotdata$time <- rep(1:8, 6)
#plotdata <- plotdata[plotdata$time!=8,]

plotdata$control <- as.factor(plotdata$control)

pd <- position_dodge(width = 0.6)
ggplot(plotdata, aes(x=time, y=coefficients, group = control, shape = control)) +
  facet_wrap( ~ outcome) + #, scales="free"
  geom_point(size=4.5, position=pd) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  geom_errorbar(aes(x=time,
                    ymin=conf.low,
                    ymax=conf.high),
                width=.15, size=1, position=pd) +
  xlab("Payments Since Receiving Flyer") + ylab("Average Causal Effect") +
  theme_minimal() +
  scale_colour_manual(values = c("black", "black")) +
  scale_x_continuous(breaks=1:8,
                     labels=c(as.character(1:7), "pooled")) +
```

```
theme(legend.position = "bottom",
      legend.title=element_blank())
```



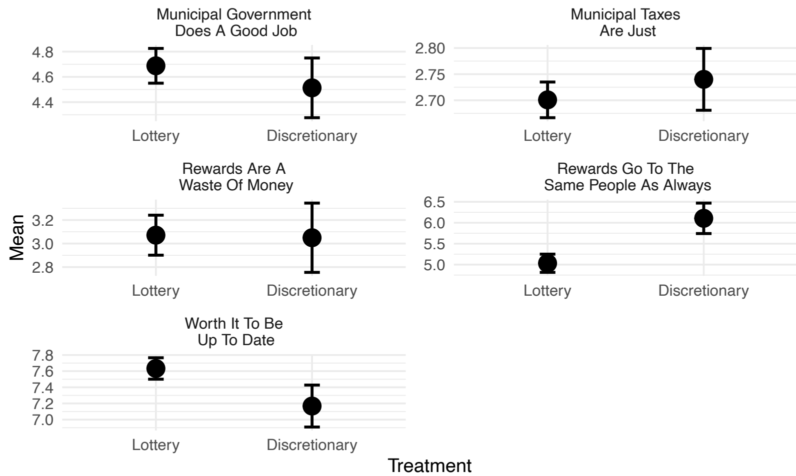

```
################################################
message("MAIN PAPER: Figure 8")
```

```
## MAIN PAPER: Figure 8
```

```
discretion1 <- ddply(survey_data, "treat_discretion",
                     summarise,
                     mean=mean(S1p4,na.rm=T),
                     N=length(na.omit(S1p4,na.rm=T)),
                     se=sd(S1p4,na.rm=T)/sqrt(N))

discretion2 <- ddply(survey_data, "treat_discretion",
                     summarise,
                     mean=mean(S1p1,na.rm=T),
                     N=length(na.omit(S1p1,na.rm=T)),
                     se=sd(S1p1,na.rm=T)/sqrt(N))

discretion3 <-  ddply(survey_data, "treat_discretion",
                      summarise,
                      mean=mean(S1p3,na.rm=T),
                      N=length(na.omit(S1p3,na.rm=T)),
                      se=sd(S1p3,na.rm=T)/sqrt(N))

discretion4 <- ddply(survey_data, "treat_discretion",
```



```
                     summarise,
                     mean=mean(S1p2,na.rm=T),
                     N=length(na.omit(S1p2,na.rm=T)),
                     se=sd(S1p2, na.rm=T)/sqrt(N))

discretion5 <- ddply(survey_data, "treat_discretion",
                     summarise,
                     mean=mean(S1p5,na.rm=T),
                     N=length(na.omit(S1p5,na.rm=T)),
                     se=sd(S1p5, na.rm=T)/sqrt(N))

discretion <- as.data.frame(rbind(discretion1, discretion2, discretion3,
                                  discretion4, discretion5))

discretion$outcome <- rep(c("Rewards Go To The\n Same People As Always",
                            "Rewards Are A\n Waste Of Money",
                            "Worth It To Be\n Up To Date",
                            "Municipal Government\n Does A Good Job",
                            "Municipal Taxes\n Are Just"), each=2)

discretion$upper <- discretion$mean + 1.96*discretion$se
discretion$lower <- discretion$mean - 1.96*discretion$se
discretion$treat_discretion <- as.factor(discretion$treat_discretion)
levels(discretion$treat_discretion) <- c("Lottery", "Discretionary")

discretion$treatment <- "Discretionary vs Lottery Rewards"

p <- ggplot(discretion, aes(x=treat_discretion, y=mean))
p + facet_wrap(~ outcome, ncol=2, scales="free") +
  geom_point(size=4.5) +
  xlab("Treatment") +
  ylab("Mean") +
  geom_errorbar(aes(x=treat_discretion,
                    ymin=lower,
                    ymax=upper),
                width=.1, size=.8) +
  theme_minimal() +
  theme(legend.position = "none")
```




```r
###############################################################################
# APPENDIX Tables & Figures
###############################################################################

##############################################
message("APPENDIX: Table A1, page 10")
```

```
## APPENDIX: Table A1, page 10
```

```r
naturalex0 <- naturalex
naturalex0$rental <- as.numeric(naturalex0$PADRON_ENV!=naturalex0$PADRON)

gtp_panel <- buen_pagador_panel
rm("buen_pagador_panel")

gtp_panel <- gtp_panel[abs(gtp_panel$centered)<58,]
taxes_panel$t[taxes_panel$TRIBUTO=="TS" | taxes_panel$TRIBUTO=="TD"] <- taxes_panel$t/2


balance <- rbind(
  ttest(naturalex0$JUBILADO, naturalex0$TREATMENT),
  ttest(naturalex0$TODOPAGO, naturalex0$TREATMENT),
  ttest(naturalex0$VALOR_CAT2004, naturalex0$TREATMENT),
  ttest(naturalex0$VALOR_CATASTRALACTUAL, naturalex0$TREATMENT),
  ttest(naturalex0$rental, naturalex0$TREATMENT),
  with(gtp_panel[gtp_panel$centered==0,], ttest(BP_time, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-1,], ttest(BP_time, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-2,], ttest(BP_time, TREATMENT)),
```



```r
  with(gtp_panel[gtp_panel$centered==-3,], ttest(BP_time, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-4,], ttest(BP_time, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-5,], ttest(BP_time, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==0,], ttest(autopay_win, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-1,], ttest(autopay_win, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-2,], ttest(autopay_win, TREATMENT)),
  with(gtp_panel[gtp_panel$centered==-3,], ttest(autopay_win, TREATMENT)))

balance <- round(balance, 3)
balance <- as.data.frame(balance)
rownames(balance) <-  c("Retiree", "Paid year in full", "2004 Property value",
                        "Current property value", "Rented Property",
                        "Good taxpayer at t=0",
                        "Good taxpayer at t=-1",
                        "Good taxpayer at t=-2",
                        "Good taxpayer at t=-3",
                        "Good taxpayer at t=-4",
                        "Good taxpayer at t=-5",
                        "Automatic debit at t=0",
                        "Automatic debit at t=-1",
                        "Automatic debit at t=-2",
                        "Automatic debit at t=-3"
)

balance$sample <- c(rep("All Taxes", 6),
                    rep("Property, head & sewage", 7),
                    rep("Property", 2))

balance[,c(2:4,6,8,9)]
```

```
##                           Mean 0  Difference     SE_Diff     N p-value
## Retiree                    0.008      -0.001       0.002  5129   0.655
## Paid year in full          0.287      -0.010       0.013  5129   0.441
## 2004 Property value     1259241.960 -199730.648 164141.207 13462   0.224
## Current property value  2747841.019 -305151.051 345883.758 13998   0.378
## Rented Property            0.243      -0.009       0.007 13998   0.229
## Good taxpayer at t=0       0.346       0.003       0.007 18621   0.651
## Good taxpayer at t=-1      0.347       0.001       0.007 18309   0.934
## Good taxpayer at t=-2      0.345       0.002       0.007 17991   0.823
## Good taxpayer at t=-3      0.346       0.002       0.007 17670   0.769
## Good taxpayer at t=-4      0.345       0.005       0.007 17350   0.514
## Good taxpayer at t=-5      0.345       0.003       0.007 17021   0.689
## Automatic debit at t=0     0.106      -0.003       0.005 14027   0.506
## Automatic debit at t=-1    0.106      -0.003       0.005 13807   0.503
## Automatic debit at t=-2    0.106      -0.004       0.005 13580   0.498
## Automatic debit at t=-3    0.106      -0.004       0.005 13350   0.447
##                                       sample
## Retiree                            All Taxes
## Paid year in full                  All Taxes
## 2004 Property value                All Taxes
## Current property value             All Taxes
## Rented Property                    All Taxes
## Good taxpayer at t=0               All Taxes
```

```
## Good taxpayer at t=-1   Property, head & sewage
## Good taxpayer at t=-2   Property, head & sewage
## Good taxpayer at t=-3   Property, head & sewage
## Good taxpayer at t=-4   Property, head & sewage
## Good taxpayer at t=-5   Property, head & sewage
## Automatic debit at t=0  Property, head & sewage
## Automatic debit at t=-1 Property, head & sewage
## Automatic debit at t=-2               Property
## Automatic debit at t=-3               Property
```

```
##############################################
message("APPENDIX: Table A2")
```

```
## APPENDIX: Table A2
```

```
btp <- gtp_panel %>% filter(ES_BP==0)

placebo <- rbind.data.frame(
  with(btp %>% filter(centered==4), ttest(BP_time, TREATMENT)),
  with(btp %>% filter(centered==5), ttest(BP_time, TREATMENT)),
  with(btp %>% filter(centered==6), ttest(BP_time, TREATMENT)),
  with(btp %>% filter(centered==7), ttest(BP_time, TREATMENT)),
  with(btp %>% filter(centered==8), ttest(BP_time, TREATMENT)),
  with(btp %>% filter(centered %in% 4:8) %>% group_by(CUENTA, TREATMENT) %>%
         dplyr::summarize(BP_time = mean(BP_time)),
       ttest(BP_time, TREATMENT))
)
```

```
## `summarise()` has grouped output by 'CUENTA'. You can override using the
## `.groups` argument.
```

```
names(placebo) <- c("Mean 1", "Mean 0", "Difference", "SE_Diff",
                    "t-stat","N","df","p-value")

placebo$outcome <- c("Good taxpayer t=1",
                     "Good taxpayer t=2",
                     "Good taxpayer t=3",
                     "Good taxpayer t=4",
                     "Good taxpayer t=5",
                     "Good taxpayer t=1-5")

placebo
```

```
##     Mean 1   Mean 0 Difference   SE_Diff   t-stat     N    df p-value
## 1 0.090345 0.084775  0.0055700 0.0052404 1.06290 11639 11622 0.28785
## 2 0.102573 0.095222  0.0073515 0.0055708 1.31964 11489 11465 0.18698
## 3 0.116531 0.114190  0.0023402 0.0060008 0.38998 11339 11332 0.69656
## 4 0.118282 0.115829  0.0024528 0.0060785 0.40352 11192 11185 0.68658
## 5 0.119650 0.115447  0.0042035 0.0061281 0.68594 11052 11041 0.49277
## 6 0.109675 0.104818  0.0048571 0.0052155 0.93129 11639 11622 0.35172
##               outcome
## 1   Good taxpayer t=1
```



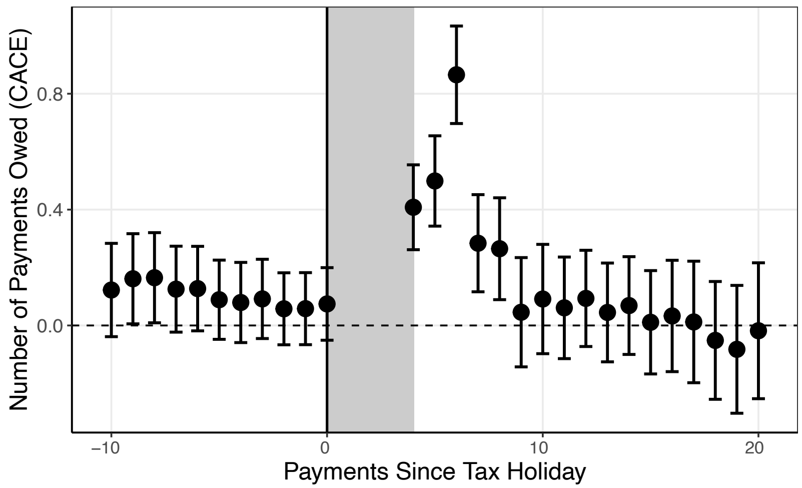
```
## 2   Good taxpayer t=2
## 3   Good taxpayer t=3
## 4   Good taxpayer t=4
## 5   Good taxpayer t=5
## 6 Good taxpayer t=1-5
```

```
# ------------------------------------------------------------------------- #
message("APPENDIX: Figure A4(a)")
```

```
## APPENDIX: Figure A4(a)
```

```
t <- unique(taxes_panel$t_st)
t <- t[order(t)]
t <- t[t %in% -10:20]
gtp_plot <- NULL

for (i in 1:length(t)){

  temp <- taxes_panel[taxes_panel$t_st == t[i], ]
  if (nrow(temp[temp$TREATMENT==1 & temp$ES_BP==0,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$ES_BP==0,])==0) next

  on_time <- t.test(cuotas_adeudadas ~ TREATMENT,
                    data=temp[temp$ES_BP==0,])
  on_time <- c(on_time$estimate[2]-on_time$estimate[1],
               -on_time$conf.int[1:2])

  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], 0, on_time)))

  if (nrow(temp[temp$TREATMENT==1 & temp$ES_BP==1,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$ES_BP==1,])==0) next

  ivest <- tidy(iv_robust(cuotas_adeudadas ~ TIENE_EXO | TREATMENT,
                          data=temp[temp$ES_BP==1,]))
   on_time <- c(ivest$estimate[2], ivest$std.error[2])



  on_time <- c(on_time[1], on_time[1]-1.96*on_time[2], on_time[1]+1.96*on_time[2])

  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], 1, on_time)))

}

gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "ES_BP", "CACE", "upper", "lower")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))
gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$upper <- as.numeric(as.character(gtp_plot$upper))
gtp_plot$lower <- as.numeric(as.character(gtp_plot$lower))
```

```r
ggplot(gtp_plot[gtp_plot$ES_BP==1,], aes(x=t, y=CACE)) +
geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
          fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.6)) +
  geom_point(size=4, position = position_dodge(width = 0.6)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Number of Payments Owed (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
        axis.text.x = element_text(size = rel(1.1), hjust=.7),
        legend.position = "bottom",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```

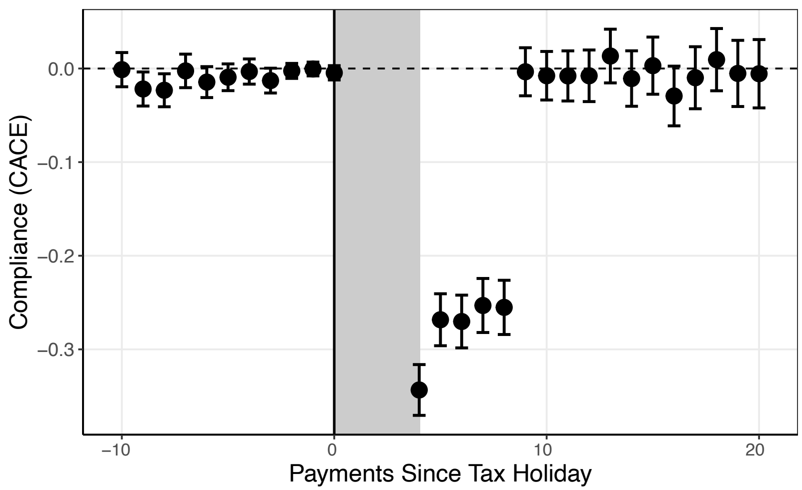



```r
################################################
message("APPENDIX: Figure A4(b)")
```

```
## APPENDIX: Figure A4(b)
```

```r
gtp_plot <- NULL

for (i in 1:length(t)){

  temp <- taxes_panel[taxes_panel$t_st == t[i], ]
  if (nrow(temp[temp$TREATMENT==1 & temp$ES_BP==0,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$ES_BP==0,])==0) next

  on_time <- t.test(compliance ~ TREATMENT,
                    data=temp[temp$ES_BP==0,])
  on_time <- c(on_time$estimate[2]-on_time$estimate[1],
               -on_time$conf.int[1:2])

  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], 0, on_time)))

  if (nrow(temp[temp$TREATMENT==1 & temp$ES_BP==1,])==0) next
  if (nrow(temp[temp$TREATMENT==0 & temp$ES_BP==1,])==0) next

  ivest <- tidy(iv_robust(compliance ~ TIENE_EXO | TREATMENT,
                          data=temp[temp$ES_BP==1,]))
  on_time <- c(ivest$estimate[2], ivest$std.error[2])

  on_time <- c(on_time[1], on_time[1]-1.96*on_time[2], on_time[1]+1.96*on_time[2])

  gtp_plot <- rbind(gtp_plot,
                    as.vector(c("Paid on Time", t[i], 1, on_time)))

}

gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "ES_BP", "CACE", "upper", "lower")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))
gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$upper <- as.numeric(as.character(gtp_plot$upper))
gtp_plot$lower <- as.numeric(as.character(gtp_plot$lower))

ggplot(gtp_plot[gtp_plot$ES_BP==1,], aes(x=t, y=CACE)) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.6)) +
  geom_point(size=4, position = position_dodge(width = 0.6)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Compliance (CACE)") +
```

```
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
        axis.text.x = element_text(size = rel(1.1), hjust=.7),
        legend.position = "bottom",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```
#################################################
message("APPENDIX: Figure A4(c)")
```

```
## APPENDIX: Figure A4(c)
```

```
difference_in_means(debt_amount ~ won_lottery,
                    data = naturalex_debt_gtp)
```

```
## Design:  Standard
```



```
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper     DF
## won_lottery   183.64     296.68 0.61898  0.53596  -397.98   765.26 4947.9
```

```
#################################################
message("APPENDIX: Figure A5")
```

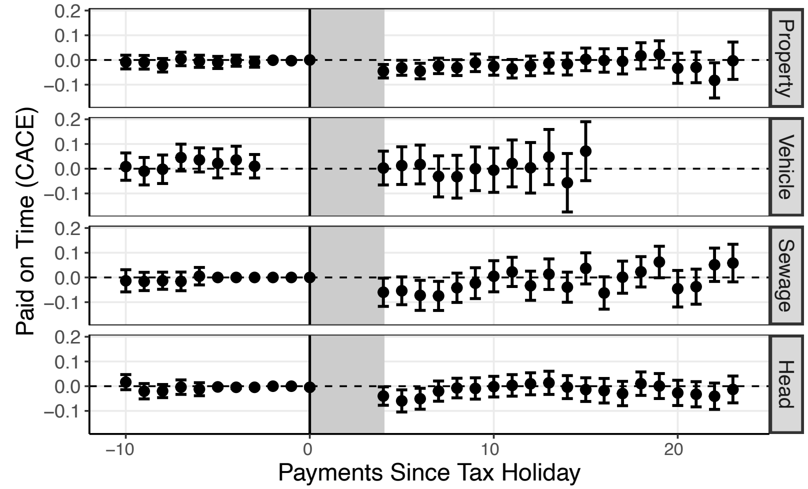
```
## APPENDIX: Figure A5
```

```
gtp_taxes <- taxes_panel[taxes_panel$ES_BP==1, ]

t <- unique(gtp_taxes$t_st[gtp_taxes$TREATMENT==1])
t <- t[order(t)]
t <- t[t>-11 & t<24]
taxes <- as.character(unique(taxes_panel$TRIBUTO))
gtp_plot <- NULL

for (i in 1:length(t)){

  for (j in 1:4){

    temp <- gtp_taxes[gtp_taxes$t_st == t[i] & gtp_taxes$TRIBUTO==taxes[j], ]
    temp$en_fecha[temp$TREATMENT==1]

    if (nrow(temp[temp$TREATMENT==1,])==0) next
    if (nrow(temp[temp$TREATMENT==0,])==0) next

    ivest <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                            data=temp))
    on_time <- c(ivest$estimate[2], ivest$std.error[2])

    ivest <- tidy(iv_robust(cuotas_adeudadas ~ TIENE_EXO | TREATMENT,
                            data=temp))
    bills_owed <- c(ivest$estimate[2], ivest$std.error[2])


    gtp_plot <- rbind(gtp_plot,
                      as.vector(c("Paid on Time", t[i], on_time, taxes[j])),
                      as.vector(c("Nr. of Bills Owed", t[i], bills_owed, taxes[j]))
    )

  }
}
```

```
## Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
```

```
gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "CACE", "SE", "tax")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))
gtp_plot <- gtp_plot[!(gtp_plot$tax=="Patente de Rodados" &
                         gtp_plot$t>15),]
gtp_plot <- gtp_plot[!((gtp_plot$tax=="Contribucion Inmobiliaria" |
```

```
                          gtp_plot$tax=="Patente de Rodados") &
                         gtp_plot$t>25),]

gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$SE <- as.numeric(as.character(gtp_plot$SE))
gtp_plot$upper <- gtp_plot$CACE + qnorm(.975) * gtp_plot$SE
gtp_plot$lower <- gtp_plot$CACE - qnorm(.975) * gtp_plot$SE

gtp_plot$tax <- as.factor(gtp_plot$tax)
levels(gtp_plot$tax) <- c("Property", "Vehicle", "Sewage", "Head")

ggplot(gtp_plot[gtp_plot$outcome=="Paid on Time",], aes(x=t, y=CACE)) +
  facet_grid(tax~.) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.5)) +
  geom_point(size=2.5, position = position_dodge(width = 0.5)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Paid on Time (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.x = element_text(size = rel(1.1), hjust=.7),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
        strip.background = element_rect(size = 1.5),
        legend.position = "none",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```


```
################################################
message("APPENDIX: Figure A6")
```

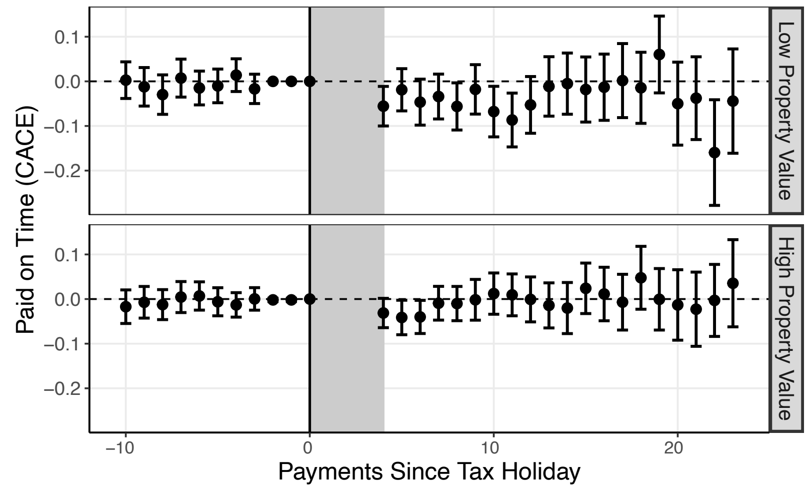
```
## APPENDIX: Figure A6
```

```
gtp_taxes <- taxes_panel[taxes_panel$ES_BP==1 &
                           taxes_panel$TRIBUTO=="Contribucion Inmobiliaria", ]

gtp_taxes$high_propvalue <- ifelse(gtp_taxes$VALOR_CAT2004 >
                                     median(gtp_taxes$VALOR_CAT2004, na.rm=T),
                                   1, 0)

t <- unique(gtp_taxes$t_st[gtp_taxes$TREATMENT==1])
t <- t[order(t)]
t <- t[t>-11 & t<24]
pay_month <- as.character(unique(gtp_taxes$bill_month))
gtp_plot <- NULL

gtp_taxes <- gtp_taxes[!is.na(gtp_taxes$high_propvalue),]

for (i in 1:length(t)){

  for (j in c(1,0)){

    temp <- gtp_taxes[gtp_taxes$t_st == t[i] &
                        gtp_taxes$high_propvalue==j, ]

    if (nrow(temp[temp$TREATMENT==1,])==0) next
```

```r
    if (nrow(temp[temp$TREATMENT==0,])==0) next

    ivest <- tidy(iv_robust(en_fecha ~ TIENE_EXO | TREATMENT,
                            data=temp))
    on_time  <- c(ivest$estimate[2], ivest$std.error[2])

    gtp_plot <- rbind(gtp_plot,
                      as.vector(c("Paid on Time", t[i], on_time, j)))


  }

}

gtp_plot <- as.data.frame(gtp_plot)
names(gtp_plot) <- c("outcome", "t", "CACE", "SE", "prop_value")

gtp_plot$t <- as.numeric(as.character(gtp_plot$t))

gtp_plot$CACE <- as.numeric(as.character(gtp_plot$CACE))
gtp_plot$SE <- as.numeric(as.character(gtp_plot$SE))
gtp_plot$upper <- gtp_plot$CACE + qnorm(.975) * gtp_plot$SE
gtp_plot$lower <- gtp_plot$CACE - qnorm(.975) * gtp_plot$SE

gtp_plot$prop_value <- as.factor(gtp_plot$prop_value)
levels(gtp_plot$prop_value) <- c("Low Property Value", "High Property Value")

ggplot(gtp_plot, aes(x=t, y=CACE)) +
  facet_grid(prop_value ~ .) +
  geom_rect(data=NULL,aes(xmin=0, xmax=4, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80") +
  geom_errorbar(aes(x=t,
                    ymin=lower,
                    ymax=upper),
                width=.6, size=.8, position = position_dodge(width = 0.5)) +
  geom_point(size=2.5, position = position_dodge(width = 0.5)) +
  xlab("Payments Since Tax Holiday") +
  ylab("Paid on Time (CACE)") +
  geom_vline(aes(xintercept=0), size=.7) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  theme_bw() +
  scale_colour_manual(values = c("black","blue")) +
  #scale_x_discrete(limit = t[!t%in%c(1,2,3)],
  #                 labels = as.character(c(t[t<1],
  #                                         t[!t%in%c(1,2,3) & t>0]-3))) +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text.x = element_text(size = rel(1.1), hjust=.7),
        axis.text.y = element_text(size = rel(1.25)),
        axis.title.y = element_text(size = rel(1.3)),
        axis.title.x = element_text(size = rel(1.3)),
        legend.text = element_text(size = rel(1.2)),
        strip.text.x = element_text(size = rel(1.4)),
        strip.text.y = element_text(size = rel(1.4)),
```



```r
        strip.background = element_rect(size = 1.5),
        legend.position = "none",
        legend.title=element_blank(),
        panel.grid.minor = element_blank(),
        axis.line = element_line(colour = "black"))
```




```r
# Function to test the difference of the differences
comp.eff <- function(dm1, dm2){

  print("Difference in Means 1")
  print(dm1)

  print("Difference in Means 2")
  print(dm2)

  print("#####. Difference in Effects")

  diff <- dm1$coefficients - dm2$coefficients
  se.diff <- sqrt((dm1$std.error^2)+(dm2$std.error^2))
  t.val.diff <- diff/se.diff
  df <- dm1$nobs + dm2$nobs -2
  # Calculate the p-value
  p_val <- 2 * (1 - pt(abs(t.val.diff), df=df))

  res <- c(diff,se.diff,t.val.diff, p_val)
names(res) <- c("Diff in effects", "SE", "t", "p-value")
print(res)
return(res)
```

```
}

comp.eff(difference_in_means(en_fecha ~ TREATMENT,
                             data = gtp_taxes[gtp_taxes$t_st == 10 &
                                                gtp_taxes$high_propvalue==1,]),
         difference_in_means(en_fecha ~ TREATMENT,
                             data = gtp_taxes[gtp_taxes$t_st == 10 &
                                                gtp_taxes$high_propvalue==0,]))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|)  CI Lower CI Upper     DF
## TREATMENT 0.010163   0.019644 0.51734  0.60505 -0.028393 0.048719 873.83
## [1] "Difference in Means 2"
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|)  CI Lower   CI Upper     DF
## TREATMENT -0.047965   0.020257 -2.3679 0.018117 -0.087725 -0.0082055 838.94
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        0.058128        0.028218        2.059993        0.039545

## Diff in effects              SE               t         p-value
##        0.058128        0.028218        2.059993        0.039545
```

```
comp.eff(difference_in_means(en_fecha ~ TREATMENT,
                             data = gtp_taxes[gtp_taxes$t_st == 11 &
                                                gtp_taxes$high_propvalue==1,]),
         difference_in_means(en_fecha ~ TREATMENT,
                             data = gtp_taxes[gtp_taxes$t_st == 11 &
                                                gtp_taxes$high_propvalue==0,]))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|)  CI Lower CI Upper     DF
## TREATMENT 0.0077676   0.019869 0.39093  0.69595 -0.031232 0.046767 839.27
## [1] "Difference in Means 2"
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower  CI Upper     DF
## TREATMENT -0.061075   0.021408 -2.8529 0.004446  -0.1031 -0.019051 783.66
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        0.068842        0.029208        2.356993        0.018538

## Diff in effects              SE               t         p-value
##        0.068842        0.029208        2.356993        0.018538
```

```
################################################
# APPENDIX TABLE A3: Multiple comparison adjustments for the Natural Experiment

# Main test: Diff-in-diff (three outcomes): 1 year DiD for all taxes pooled
```



```
# rescaling the time variable to account for the taxes that have twice as
# many payments per year
taxes_panel$t_st_2 <- ifelse(taxes_panel$tax=="Sewage" | taxes_panel$tax=="Head",
                             taxes_panel$t_st/2, taxes_panel$t_st)


# 1 year diff in diff setup
dd_data <-  rbind.data.frame(taxes_panel %>% filter(ES_BP==1) %>%
                               group_by(CUENTA, TRIBUTO, TREATMENT) %>% dplyr::summarise(
                                 compliance_mean_DiD_1yr =
                                   mean(compliance[t_st>3 & t_st<=6], na.rm=T) -
                                   mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                 missed_payment_mean_DiD_1yr =
                                   mean(missed_payment[t_st>3 & t_st<=6], na.rm=T)-
                                   mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                 nr_missed_payments_mean_DiD_1yr =
                                   mean(cuotas_adeudadas[t_st> 3 & t_st<=6], na.rm=T)-
                                   mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T),
                                 compliance_mean_DiD_1yr.yr2 =
                                   mean(compliance[t_st>6 & t_st<=9], na.rm=T) -
                                   mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                 missed_payment_mean_DiD_1yr.yr2 =
                                   mean(missed_payment[t_st>6 & t_st<=9], na.rm=T)-
                                   mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                 nr_missed_payments_mean_DiD_1yr.yr2 =
                                   mean(cuotas_adeudadas[t_st> 6 & t_st<=9], na.rm=T)-
                                   mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T)
                               )
)
```

```
## `summarise()` has grouped output by 'CUENTA', 'TRIBUTO'. You can override using
## the `.groups` argument.
```

```
DiD_1yr_compliance <- difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data)
DiD_1yr_compliance
```

```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT  -0.2061  0.0094951 -21.706
##
## TREATMENT 0.00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper     DF
## TREATMENT -0.22471 -0.18748 4772.5
```

```
DiD_1yr_missed_payment <- difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data)
DiD_1yr_missed_payment
```

```
## Design:  Standard
##           Estimate Std. Error t value    Pr(>|t|) CI Lower CI Upper     DF
## TREATMENT  0.02783  0.0067317  4.1341 0.000036166 0.014633 0.041026 5413.9
```

```r
DiD_1yr_nrmissed_payments <- difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data)
DiD_1yr_nrmissed_payments
```

```
## Design:  Standard
##           Estimate Std. Error t value                   Pr(>|t|) CI Lower
## TREATMENT  0.35802   0.039006  9.1787 0.0000000000000000000065324  0.28155
##           CI Upper     DF
## TREATMENT  0.43449 4445.8
```

```r
main <- rbind.data.frame(c("Main 1yr DiD - Missed Payment", DiD_1yr_missed_payment$p.value),
c("Main 1yr DiD - Nr Missed Payments", DiD_1yr_nrmissed_payments$p.value),
c("Main 1yr DiD - Compliance", DiD_1yr_compliance$p.value))

# Persistence of effects, heterogeneous effects (three outcomes)
# Difference between years 1 and 2 for the three outcomes.

persistence_missed <- comp.eff(difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data),
difference_in_means(missed_payment_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value    Pr(>|t|) CI Lower CI Upper     DF
## TREATMENT  0.02783  0.0067317  4.1341 0.000036166 0.014633 0.041026 5413.9
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|)  CI Lower CI Upper     DF
## TREATMENT 0.017565  0.0070116  2.5051 0.012272 0.0038191 0.031311 5014.2
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##       0.0102649       0.0097199       1.0560633       0.2909636
```

```r
persistence_nrowed <- comp.eff(difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data),
difference_in_means(nr_missed_payments_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value                   Pr(>|t|) CI Lower
## TREATMENT  0.35802   0.039006  9.1787 0.0000000000000000000065324  0.28155
##           CI Upper     DF
## TREATMENT  0.43449 4445.8
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  0.11354   0.040471  2.8054 0.0050456 0.034195  0.19288 4946
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##      0.24448466      0.05620823      4.34962394      0.00001376
```



```r
persistence_compliance <- comp.eff(difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data),
difference_in_means(compliance_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT  -0.2061  0.0094951 -21.706
##
## TREATMENT 0.00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper     DF
## TREATMENT -0.22471 -0.18748 4772.5
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT -0.12163  0.0088733 -13.708
##                                                      Pr(>|t|) CI Lower CI Upper
## TREATMENT 0.000000000000000000000000000000000000000000051821 -0.13903 -0.10424
##                  DF
## TREATMENT 5019.8
## [1] "#####. Difference in Effects"
##    Diff in effects                SE                  t             p-value
## -0.084466442099955  0.012995886315108 -6.499475299484610  0.000000000084091
```

```r
persistence <- rbind.data.frame(c("Persistence yr 1 vs 2  - Missed Payment", persistence_missed[4]),
                                c("Persistence yr 1 vs 2  - Nr Payments Missed", persistence_nrowed[4])
                                c("Persistence yr 1 vs 2 - Compliance", persistence_compliance[4])
                )



# Heterogeneous effects, by cost of payment

# 1 year diff in diff setup
dd_data_vc <-  rbind.data.frame(taxes_panel %>% filter(ES_BP==1 & TRIBUTO=="Contribucion Inmobiliaria")
                               group_by(CUENTA, TREATMENT, VALOR_CAT2004) %>% dplyr::summarise(
                                 compliance_mean_DiD_1yr =
                                   mean(compliance[t_st>3 & t_st<=6], na.rm=T) -
                                   mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                 missed_payment_mean_DiD_1yr =
                                   mean(missed_payment[t_st>3 & t_st<=6], na.rm=T)-
                                   mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                 nr_missed_payments_mean_DiD_1yr =
                                   mean(cuotas_adeudadas[t_st> 3 & t_st<=6], na.rm=T)-
                                   mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T)
                               )
)
```

```
## `summarise()` has grouped output by 'CUENTA', 'TREATMENT'. You can override
## using the `.groups` argument.
```

```r
dd_data_vc$high_propvalue <- ifelse(dd_data_vc$VALOR_CAT2004 >
                                  median(dd_data_vc$VALOR_CAT2004, na.rm=T),
                                1, 0)

income_missed <- comp.eff(difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 1)),
         difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper     DF
## TREATMENT 0.034689   0.010999  3.1538 0.0016582 0.013106 0.056273 1031.9
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|)   CI Lower CI Upper     DF
## TREATMENT 0.018827   0.013383  1.4067   0.1598 -0.0074338 0.045087 1069.1
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        0.015863        0.017323        0.915687        0.359930
```

```r
income_nrmissed <- comp.eff(difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 1)),
         difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT  0.94033   0.047994  19.592
##                                                                               Pr(>|t|)
## TREATMENT 0.0000000000000000000000000000000000000000000000000000000000000000000064371
##           CI Lower CI Upper     DF
## TREATMENT   0.8461   1.0346 724.79
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value                               Pr(>|t|)
## TREATMENT  0.81161   0.066725  12.164 0.000000000000000000000000000000044187
##           CI Lower CI Upper     DF
## TREATMENT  0.68069  0.94253 1132.5
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        0.128721        0.082193        1.566085        0.117464
```

```r
income_compliance <- comp.eff(difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 1)),
         difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
```



```
##           Estimate Std. Error t value
## TREATMENT  -0.5051   0.021872 -23.093
##                                                                                                     P
## TREATMENT 0.0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper    DF
## TREATMENT -0.54803 -0.46217 820.7
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT -0.41839   0.022015 -19.005
##                                                                             Pr(>|t|)
## TREATMENT 0.00000000000000000000000000000000000000000000000000000000000000000018416
##           CI Lower CI Upper     DF
## TREATMENT  -0.4616 -0.37518 855.95
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##      -0.0867107       0.0310331      -2.7941387       0.0052463
```

```r
income <- rbind.data.frame(c("Income HTE - Missed Payment", income_missed[4]),
                           c("Income HTE - Nr Payments Missed", income_nrmissed[4]),
                           c("Income HTE - Compliance", income_compliance[4]))


# Comparing natural and field experiments
natvsfield <- comp.eff(difference_in_means(missed_payment ~ TREATMENT,
                             data = filter(taxes_panel, ES_BP==1 & t_st==4)),
         difference_in_means(JUL_2014_ontime ~ pooled_124_0, weights = pooled_124_0_wts,
                             data = filter(fieldex, type=="good taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value   Pr(>|t|) CI Lower CI Upper     DF
## TREATMENT 0.028635  0.0075862  3.7746 0.00016197 0.013763 0.043507 5362.7
## [1] "Difference in Means 2"
## Design:  Standard (weighted)
##                Estimate Std. Error t value Pr(>|t|)  CI Lower CI Upper   DF
## pooled_124_0 0.00079172  0.0075116  0.1054  0.91606 -0.013935 0.015518 4565
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        0.027844        0.010676        2.608069        0.009119
```

```r
natvsfield <- rbind.data.frame(c("Natural vs Field Experiment", natvsfield[4]))


names(main) <- names(persistence) <- names(income) <- names(natvsfield) <- c("H","p")
all <- rbind(main, persistence, income, natvsfield)
all$p <- as.numeric(all$p)
all$p.bonferroni <- p.adjust(all$p, method = "bonferroni")
all$p.fdr <- p.adjust(all$p, method = "fdr")

# Round the numeric columns to 3 decimal places
all[, 2:4] <- round(all[, 2:4], 5)

all
```

```
##                                            H       p p.bonferroni   p.fdr
## 1                Main 1yr DiD - Missed Payment 0.00004      0.00036 0.00007
## 2            Main 1yr DiD - Nr Missed Payments 0.00000      0.00000 0.00000
## 3                    Main 1yr DiD - Compliance 0.00000      0.00000 0.00000
## 4      Persistence yr 1 vs 2  - Missed Payment 0.29096      1.00000 0.32329
## 5  Persistence yr 1 vs 2  - Nr Payments Missed 0.00001      0.00014 0.00003
## 6          Persistence yr 1 vs 2 - Compliance 0.00000      0.00000 0.00000
## 7                 Income HTE - Missed Payment 0.35993      1.00000 0.35993
## 8             Income HTE - Nr Payments Missed 0.11746      1.00000 0.14683
## 9                     Income HTE - Compliance 0.00525      0.05246 0.00874
## 10                Natural vs Field Experiment 0.00912      0.09119 0.01303
```

```
##############################################
# APPENDIX TABLE A4: Field Experiment: Pre-Treatment Balance

out <- cbind.data.frame(
  c("MAR_2010_ontime", "JUL_2010_ontime", "NOV_2010_ontime", "MAR_2011_ontime",
    "JUL_2011_ontime", "NOV_2011_ontime", "MAR_2012_ontime", "JUL_2012_ontime",
    "NOV_2012_ontime", "MAR_2013_ontime", "JUL_2013_ontime", "NOV_2013_ontime",
    "MAR_2014_ontime"),
  c("Paid on Time MAR 2010", "Paid on Time JUL 2010", "Paid on Time NOV 2010",
    "Paid on Time MAR 2011", "Paid on Time JUL 2011", "Paid on Time NOV 2011",
    "Paid on Time MAR 2012", "Paid on Time JUL 2012", "Paid on Time NOV 2012",
    "Paid on Time MAR 2013", "Paid on Time JUL 2013", "Paid on Time NOV 2013",
    "Paid on Time MAR 2014")
)

treat <- rbind.data.frame(
  c("pooled_124_6", "pooled_124_6_wts", "Treatment versus Pure Control"),
  c("pooled_124_0", "pooled_124_0_wts", "Treatment versus Placebo"),
  c("pooled_0_6", "pooled_0_6_wts", "Placebo versus Pure Control")
)

bal <- merge(treat, out)
names(bal) <- c("treat_var", "wts", "treat_label", "out_var", "out_label")

balance <- NULL
for (i in 1:nrow(bal)){

  out <- fieldex[,bal$out_var[i]]
  treat <- fieldex[,bal$treat_var[i]]
  wts <- fieldex[,bal$wts[i]]
  balance <- rbind.data.frame(balance, difference_in_means(out ~ treat,
                                                           weights = wts))
  rm(out, treat, wts)
}

balance <- cbind(bal[,c("treat_label", "out_label")],
                 balance[,c("coefficients", "std.error", "nobs", "p.value")])

balance
```

```
##                      treat_label             out_label coefficients std.error
## 1  Treatment versus Pure Control Paid on Time MAR 2010   0.00144937 0.0071966
```



```
## 2       Treatment versus Placebo Paid on Time MAR 2010   0.00204241 0.0100170
## 3    Placebo versus Pure Control Paid on Time MAR 2010  -0.00123572 0.0087109
## 4  Treatment versus Pure Control Paid on Time JUL 2010   0.00449265 0.0072252
## 5       Treatment versus Placebo Paid on Time JUL 2010   0.00193801 0.0100599
## 6    Placebo versus Pure Control Paid on Time JUL 2010   0.00499282 0.0086563
## 7  Treatment versus Pure Control Paid on Time NOV 2010   0.01951457 0.0072176
## 8       Treatment versus Placebo Paid on Time NOV 2010  -0.00180374 0.0100054
## 9    Placebo versus Pure Control Paid on Time NOV 2010   0.02171193 0.0086319
## 10 Treatment versus Pure Control Paid on Time MAR 2011   0.00844649 0.0071248
## 11      Treatment versus Placebo Paid on Time MAR 2011  -0.00240279 0.0099315
## 12   Placebo versus Pure Control Paid on Time MAR 2011   0.01056837 0.0085213
## 13 Treatment versus Pure Control Paid on Time JUL 2011   0.00759820 0.0070888
## 14      Treatment versus Placebo Paid on Time JUL 2011  -0.00219970 0.0098723
## 15   Placebo versus Pure Control Paid on Time JUL 2011   0.00820028 0.0085152
## 16 Treatment versus Pure Control Paid on Time NOV 2011   0.00091900 0.0070400
## 17      Treatment versus Placebo Paid on Time NOV 2011  -0.00184469 0.0098361
## 18   Placebo versus Pure Control Paid on Time NOV 2011   0.00117599 0.0084731
## 19 Treatment versus Pure Control Paid on Time MAR 2012   0.01151766 0.0069545
## 20      Treatment versus Placebo Paid on Time MAR 2012  -0.00583782 0.0097552
## 21   Placebo versus Pure Control Paid on Time MAR 2012   0.01193612 0.0083481
## 22 Treatment versus Pure Control Paid on Time JUL 2012  -0.00140435 0.0070199
## 23      Treatment versus Placebo Paid on Time JUL 2012   0.00081256 0.0098653
## 24   Placebo versus Pure Control Paid on Time JUL 2012  -0.00567193 0.0085353
## 25 Treatment versus Pure Control Paid on Time NOV 2012  -0.00237368 0.0069520
## 26      Treatment versus Placebo Paid on Time NOV 2012  -0.00711907 0.0097933
## 27   Placebo versus Pure Control Paid on Time NOV 2012   0.00425777 0.0082334
## 28 Treatment versus Pure Control Paid on Time MAR 2013   0.00310250 0.0069097
## 29      Treatment versus Placebo Paid on Time MAR 2013  -0.00737321 0.0097620
## 30   Placebo versus Pure Control Paid on Time MAR 2013   0.00966121 0.0081435
## 31 Treatment versus Pure Control Paid on Time JUL 2013  -0.00703562 0.0067432
## 32      Treatment versus Placebo Paid on Time JUL 2013  -0.00836457 0.0098008
## 33   Placebo versus Pure Control Paid on Time JUL 2013  -0.00083002 0.0077181
## 34 Treatment versus Pure Control Paid on Time NOV 2013   0.00425667 0.0068100
## 35      Treatment versus Placebo Paid on Time NOV 2013   0.00146086 0.0098633
## 36   Placebo versus Pure Control Paid on Time NOV 2013   0.00291111 0.0078577
## 37 Treatment versus Pure Control Paid on Time MAR 2014  -0.00167924 0.0068263
## 38      Treatment versus Placebo Paid on Time MAR 2014  -0.00314534 0.0098393
## 39   Placebo versus Pure Control Paid on Time MAR 2014   0.00031359 0.0078903
##     nobs   p.value
## 1  16531 0.8403899
## 2  10022 0.8384393
## 3  13241 0.8871936
## 4  16582 0.5340776
## 5  10060 0.8472387
## 6  13282 0.5640932
## 7  16616 0.0068632
## 8  10091 0.8569390
## 9  13297 0.0119041
## 10 16896 0.2358369
## 11 10245 0.8088339
## 12 13507 0.2149139
## 13 16937 0.2838001
## 14 10270 0.8236834
## 15 13541 0.3355571
```

```
## 16 16984 0.8961412
## 17 10295 0.8512382
## 18 13583 0.8896167
## 19 17234 0.0977111
## 20 10447 0.5495631
## 21 13777 0.1527959
## 22 17367 0.8414413
## 23 10494 0.9343578
## 24 13883 0.5063650
## 25 17378 0.7327759
## 26 10503 0.4672810
## 27 13895 0.6050703
## 28 17567 0.6534322
## 29 10598 0.4500852
## 30 14065 0.2354949
## 31 17660 0.2967938
## 32 10654 0.3934254
## 33 14134 0.9143599
## 34 17686 0.5319366
## 35 10669 0.8822584
## 36 14155 0.7110303
## 37 17808 0.8056894
## 38 10778 0.7492231
## 39 14244 0.9682983
```

```
################################################
## APPENDIX TABLE A6: Field Experiment: Treatment Conditions and Sample
## Sizes (Full Experimental Design)
table(fieldex$treatment, fieldex$type)
```

```
##
##     good taxpayer bad taxpayer
##   0          1532         2080
##   1           767         1050
##   2           751         1043
##   3          1465         2109
##   4          1519         2057
##   5          1507         2111
##   6          7243         3412
```

```
table(fieldex$type)
```

```
##
## good taxpayer  bad taxpayer
##         14784         13862
```

```
################################################
## APPENDIX FIGURE A15: Field Experiment: Complete Results.

# Filter into eligible and ineligible datasets
fieldexE <- fieldex %>% filter(type=="good taxpayer")
fieldexN <- fieldex %>% filter(type=="bad taxpayer")
```



```
## Eligibles
web <- as.data.frame(rbind(
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==0),],
       ttest(july_web_access,(treatment==0)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==1),],
       ttest(july_web_access,(treatment==1)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==3),],
       ttest(july_web_access,(treatment==3)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==4),],
       ttest(july_web_access,(treatment==4)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==5),],
       ttest(july_web_access,(treatment==5)))[c(3:4,8)]))
web$outcome <- "Intended Compliance (Accessed Web Account)"

missed <- as.data.frame(rbind(
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==0),],
       ttest(JUL_2014_ontime,(treatment==0)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==1),],
       ttest(JUL_2014_ontime,(treatment==1)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==3),],
       ttest(JUL_2014_ontime,(treatment==3)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==4),],
       ttest(JUL_2014_ontime,(treatment==4)))[c(3:4,8)],
  with(fieldexE[(fieldexE$treatment==6 | fieldexE$treatment==5),],
       ttest(JUL_2014_ontime,(treatment==5)))[c(3:4,8)]))
missed$outcome <- "Paid Bill On Time"

fieldex_plotE <- rbind(missed, web)
fieldex_plotE$type <- "Eligible Taxpayers"

## Noneligibles
web <- as.data.frame(rbind(
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==0),],
       ttest(july_web_access,(treatment==0)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==1),],
       ttest(july_web_access,(treatment==1)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==3),],
       ttest(july_web_access,(treatment==3)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==4),],
       ttest(july_web_access,(treatment==4)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==5),],
       ttest(july_web_access,(treatment==5)))[c(3:4,8)]))
web$outcome <- "Intended Compliance (Accessed Web Account)"

missed <- as.data.frame(rbind(
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==0),],
       ttest(JUL_2014_ontime,(treatment==0)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==1),],
       ttest(JUL_2014_ontime,(treatment==1)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==3),],
       ttest(JUL_2014_ontime,(treatment==3)))[c(3:4,8)],
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==4),],
       ttest(JUL_2014_ontime,(treatment==4)))[c(3:4,8)],
```

```r
  with(fieldexN[(fieldexN$treatment==6 | fieldexN$treatment==5),],
       ttest(JUL_2014_ontime,(treatment==5)))[c(3:4,8)]))
missed$outcome <- "Paid Bill On Time"

fieldex_plotN <- rbind(missed, web)
fieldex_plotN$type <- "Ineligible Taxpayers"

fieldex_plot <- rbind(fieldex_plotE,fieldex_plotN)

names(fieldex_plot) <- c("mean", "se", "p-value", "outcome", "type")

fieldex_plot$upper <- fieldex_plot$mean + qnorm(.975)*(fieldex_plot$se)
fieldex_plot$lower <- fieldex_plot$mean - qnorm(.975)*(fieldex_plot$se)

fieldex_plot$treatment <- rep(c("Reminder",
                                "Individual\n Reward",
                                "Individual\n Punishment",
                                "Social\n Reward",
                                "Social\n Punishment"), 4)

fieldex_plot$treatment <- as.factor(fieldex_plot$treatment)
fieldex_plot$treatment <- factor(fieldex_plot$treatment,
                                 levels = c("Reminder", "Individual\n Reward",
                                            "Individual\n Punishment",
                                            "Social\n Reward",
                                            "Social\n Punishment"))
class(fieldex_plot$treatment)
```

```
## [1] "factor"
```

```r
fieldex_plot$type <- as.factor(fieldex_plot$type)



# FDR and Bonferroni corrections

#Threshold for FDR correction
# get and order the nominal p-values
ordered.ps <- fieldex_plot$`p-value`[order(fieldex_plot$`p-value`,decreasing=F)]
ordered.ps
```

```
##  [1] 0.00000000000043949 0.00000000095673935 0.00001579396908720
##  [4] 0.00002988982333668 0.00004163861098085 0.00177493405890230
##  [7] 0.00560698939283640 0.00649614798294486 0.00992842550242703
## [10] 0.07646818678917872 0.08554740761848820 0.14340588290909964
## [13] 0.14508250319634866 0.21101199882511876 0.23634696267256564
## [16] 0.42456695205420925 0.63467888512286907 0.76737312043598704
## [19] 0.83246937184424596 0.83418565969326619
```

```r
comp <- (1:length(ordered.ps)/length(ordered.ps))*(.05)

FDR <- cbind(ordered.ps,comp,ordered.ps<=comp)
```



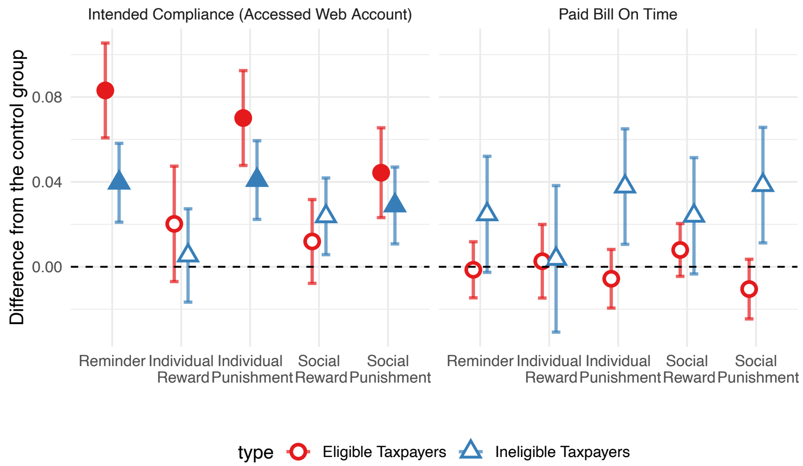
```r
fdr <- max(FDR[,1][FDR[,3]==1])
fdr
```

```
## [1] 0.0099284
```

```r
#Threshold for Bonferroni correction
bonf <- 0.05/length(ordered.ps)
bonf
```

```
## [1] 0.0025
```

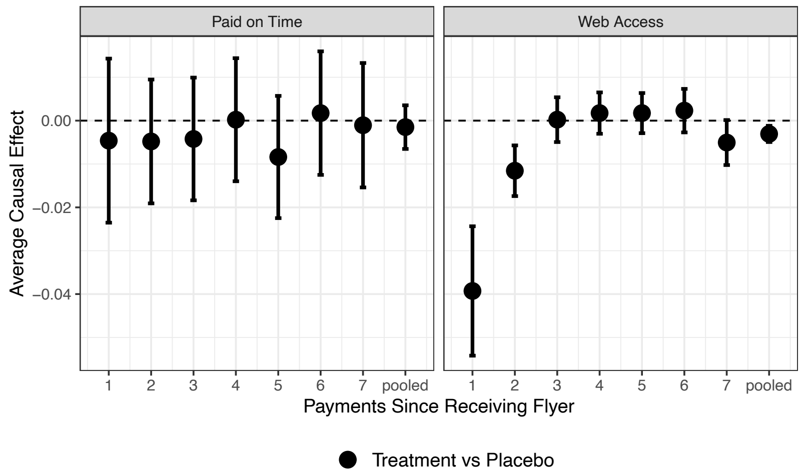
```r
fieldex_plot$Bonf_reject <- NA
fieldex_plot$FDR_reject <- NA

fieldex_plot$Bonf_reject[fieldex_plot$`p-value`<=bonf] <- "yes"
fieldex_plot$Bonf_reject[fieldex_plot$`p-value`>bonf]  <- "no"

fieldex_plot$FDR_reject[fieldex_plot$`p-value`<=fdr] <- "yes"
fieldex_plot$FDR_reject[fieldex_plot$`p-value`>fdr]  <- "no"

fieldex_plot$bonf_fdr <- as.numeric(fieldex_plot$Bonf_reject=="yes" &
                                      fieldex_plot$FDR_reject=="yes")

pd <- position_dodge(width = 0.4)
ggplot(fieldex_plot, aes(x=treatment, y=mean, group=type,
                         color=type, shape=type)) +
  facet_wrap(~outcome) +
    geom_errorbar(aes(x=treatment,
                    ymin=lower, ymax=upper),
                width=.25, size=.9, alpha=.7,
                position=pd) +
  geom_point(size=4.5, position=pd) +
  geom_point(size=2.5, position=pd,
             data=fieldex_plot[fieldex_plot$bonf_fdr==0,],
             aes(x=treatment, y=mean, group=type,
                shape=type), color = "white") +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  xlab(" ") +
  ylab("Difference from the control group") +
  scale_colour_brewer(palette="Set1") +
  theme_minimal() +
  theme(legend.position = "bottom")
```




```
###############################################
## APPENDIX FIGURE A16: Field Experiment: Effects of the Information Treatment
# versus the Placebo Control (no IPW)

# a) Compliance
comp_placebo <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(NOV_2014_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(MAR_2015_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(JUL_2015_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(NOV_2015_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(MAR_2016_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(JUL_2016_ontime ~ pooled_124_0, data = fieldex),
  difference_in_means(compliance_1416 ~ pooled_124_0, data = fieldex)
)
comp_placebo$outcome <- "Paid on Time"
comp_placebo$control <- "Treatment vs Placebo"
# b) Intended compliance
intcomp_placebo <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_0, data = fieldex),
  difference_in_means(intended_1416 ~ pooled_124_0, data = fieldex)
)
intcomp_placebo$outcome <- "Web Access"
```



```
intcomp_placebo$control <- "Treatment vs Placebo"


## Combine and plot
plotdata <- rbind.data.frame(comp_placebo, intcomp_placebo)
rm(comp_placebo, intcomp_placebo)

plotdata$time <- rep(1:8, 2)
#plotdata <- plotdata[plotdata$time!=8,]

plotdata$control <- as.factor(plotdata$control)

pd <- position_dodge(width = 0.6)
#pdf(file=paste0(save, "/fieldex_appendix.pdf"), height=6, width=9)
ggplot(plotdata, aes(x=time, y=coefficients, group = control, shape = control)) +
  facet_wrap( ~ outcome) + #, scales="free"
  geom_point(size=4.5, position=pd) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  geom_errorbar(aes(x=time,
                    ymin=conf.low,
                    ymax=conf.high),
                width=.15, size=1, position=pd) +
  xlab("Payments Since Receiving Flyer") + ylab("Average Causal Effect") +
  theme_bw() +
  scale_colour_manual(values = c("black", "black")) +
  scale_x_continuous(breaks=1:8,
                     labels=c(as.character(1:7), "pooled")) +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text.x = element_text(size = rel(1)),
        axis.text.y = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        legend.text = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1)),
        strip.text.y = element_text(size = rel(1)),
        legend.position = "bottom",
        legend.title=element_blank())
```




```r
#dev.off()


###############################################
## APPENDIX FIGURE A17: Field Experiment: Effects of Information About the
# Tax Holiday on Compliance. Heterogeneous Treatment Effects by Taxpayer Type.

#### Eligibles
#### TREATMENT VERSUS PURE CONTROL
# a) compliance
comp_controlE <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(NOV_2014_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(MAR_2015_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(JUL_2015_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(NOV_2015_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(MAR_2016_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(JUL_2016_ontime ~ pooled_124_6, data = fieldexE),
  difference_in_means(compliance_1416 ~ pooled_124_6, data = fieldexE)
)
comp_controlE$outcome <- "Paid on Time"
comp_controlE$control <- "Treatment vs Pure Control"
comp_controlE$type <- "Eligibles"
# b) Intended compliance
intcomp_controlE <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_6, data = fieldexE),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_6, data = fieldexE),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_6, data = fieldexE),
```



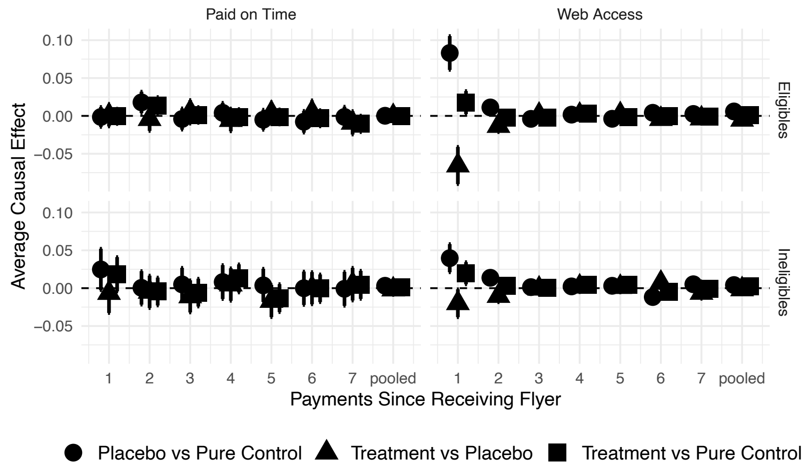
```r
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_6, data = fieldexE),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_6, data = fieldexE),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_6, data = fieldexE),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_6, data = fieldexE),
  difference_in_means(intended_1416 ~ pooled_124_6, data = fieldexE)
)
intcomp_controlE$control <- "Treatment vs Pure Control"
intcomp_controlE$outcome <- "Web Access"
intcomp_controlE$type <- "Eligibles"

#### TREATMENT VERSUS PLACEBO
# a) Compliance
comp_placeboE <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(NOV_2014_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(MAR_2015_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(JUL_2015_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(NOV_2015_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(MAR_2016_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(JUL_2016_ontime ~ pooled_124_0, data = fieldexE),
  difference_in_means(compliance_1416 ~ pooled_124_0, data = fieldexE)
)
comp_placeboE$control <- "Treatment vs Placebo"
comp_placeboE$outcome <- "Paid on Time"
comp_placeboE$type <- "Eligibles"
# b) Intended compliance
intcomp_placeboE <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_0, data = fieldexE),
  difference_in_means(intended_1416 ~ pooled_124_0, data = fieldexE)
)
intcomp_placeboE$outcome <- "Web Access"
intcomp_placeboE$control <- "Treatment vs Placebo"
intcomp_placeboE$type <- "Eligibles"
#### PURE CONTROL VERSUS PLACEBO
# a) Compliance
comp_control_plaE <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(NOV_2014_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(MAR_2015_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(JUL_2015_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(NOV_2015_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(MAR_2016_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(JUL_2016_ontime ~ pooled_0_6, data = fieldexE),
  difference_in_means(compliance_1416 ~ pooled_0_6, data = fieldexE)
)
comp_control_plaE$outcome <- "Paid on Time"
comp_control_plaE$control <- "Placebo vs Pure Control"
```

```r
comp_control_plaE$type <- "Eligibles"
# b) Intended compliance
intcomp_control_plaE <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_0_6, data = fieldexE),
  difference_in_means(intended_1416 ~ pooled_0_6, data = fieldexE)
)
intcomp_control_plaE$outcome <- "Web Access"
intcomp_control_plaE$control <- "Placebo vs Pure Control"
intcomp_control_plaE$type <- "Eligibles"

#### Ineligibles
#### TREATMENT VERSUS PURE CONTROL
# a) compliance
comp_controlN <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(NOV_2014_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(MAR_2015_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(JUL_2015_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(NOV_2015_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(MAR_2016_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(JUL_2016_ontime ~ pooled_124_6, data = fieldexN),
  difference_in_means(compliance_1416 ~ pooled_124_6, data = fieldexN)
)
comp_controlN$outcome <- "Paid on Time"
comp_controlN$control <- "Treatment vs Pure Control"
comp_controlN$type <- "Ineligibles"
# b) Intended compliance
intcomp_controlN <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_6, data = fieldexN),
  difference_in_means(intended_1416 ~ pooled_124_6, data = fieldexN)
)
intcomp_controlN$outcome <- "Web Access"
intcomp_controlN$control <- "Treatment vs Pure Control"
intcomp_controlN$type <- "Ineligibles"
#### TREATMENT VERSUS PLACEBO
# a) Compliance
comp_placeboN <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_124_0, data = fieldexN),
  difference_in_means(NOV_2014_ontime ~ pooled_124_0, data = fieldexN),
  difference_in_means(MAR_2015_ontime ~ pooled_124_0, data = fieldexN),
  difference_in_means(JUL_2015_ontime ~ pooled_124_0, data = fieldexN),
```



```r
  difference_in_means(NOV_2015_ontime ~ pooled_124_0, data = fieldexN),
  difference_in_means(MAR_2016_ontime ~ pooled_124_0, data = fieldexN),
  difference_in_means(JUL_2016_ontime ~ pooled_124_0, data = fieldexN),
  difference_in_means(compliance_1416 ~ pooled_124_0, data = fieldexN)
)
comp_placeboN$outcome <- "Paid on Time"
comp_placeboN$control <- "Treatment vs Placebo"
comp_placeboN$type <- "Ineligibles"
# b) Intended compliance
intcomp_placeboN <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_124_0, data = fieldexN),
  difference_in_means(intended_1416 ~ pooled_124_0, data = fieldexN)
)
intcomp_placeboN$outcome <- "Web Access"
intcomp_placeboN$control <- "Treatment vs Placebo"
intcomp_placeboN$type <- "Ineligibles"
#### PURE CONTROL VERSUS PLACEBO
# a) Compliance
comp_control_plaN <- rbind.data.frame(
  difference_in_means(JUL_2014_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(NOV_2014_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(MAR_2015_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(JUL_2015_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(NOV_2015_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(MAR_2016_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(JUL_2016_ontime ~ pooled_0_6, data = fieldexN),
  difference_in_means(compliance_1416 ~ pooled_0_6, data = fieldexN)
)
comp_control_plaN$outcome <- "Paid on Time"
comp_control_plaN$control <- "Placebo vs Pure Control"
comp_control_plaN$type <- "Ineligibles"
# b) Intended compliance
intcomp_control_plaN <- rbind.data.frame(
  difference_in_means(JUL_2014_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(NOV_2014_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(MAR_2015_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(JUL_2015_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(NOV_2015_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(MAR_2016_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(JUL_2016_WEBACCESS ~ pooled_0_6, data = fieldexN),
  difference_in_means(intended_1416 ~ pooled_0_6, data = fieldexN)
)
intcomp_control_plaN$outcome <- "Web Access"
intcomp_control_plaN$control <- "Placebo vs Pure Control"
intcomp_control_plaN$type <- "Ineligibles"
```

```r
## Combine and plot
plotdata <- rbind.data.frame(comp_placeboE, comp_controlE,
                             intcomp_placeboE, intcomp_controlE,
                             comp_control_plaE, intcomp_control_plaE,
                             comp_placeboN, comp_controlN,
                             intcomp_placeboN, intcomp_controlN,
                             comp_control_plaN, intcomp_control_plaN)
rm(comp_placeboE, comp_controlE,
   intcomp_placeboE, intcomp_controlE,
   comp_control_plaE, intcomp_control_plaE,
   comp_placeboN, comp_controlN,
   intcomp_placeboN, intcomp_controlN,
   comp_control_plaN, intcomp_control_plaN)

plotdata$time <- rep(1:8, 6)
#plotdata <- plotdata[plotdata$time!=8,]

plotdata$control <- as.factor(plotdata$control)

pd <- position_dodge(width = 0.6)
#pdf(file=paste0(save, "/fieldex_appendix.pdf"), height=6, width=9)
ggplot(plotdata, aes(x=time, y=coefficients, group = control, shape = control)) +
  facet_grid(type ~ outcome) + #, scales="free"
  geom_point(size=4.5, position=pd) +
  geom_hline(aes(yintercept=0), size=.5, linetype="dashed") +
  geom_errorbar(aes(x=time,
                    ymin=conf.low,
                    ymax=conf.high),
                width=.15, size=1, position=pd) +
  xlab("Payments Since Receiving Flyer") + ylab("Average Causal Effect") +
  theme_minimal() +
  scale_colour_manual(values = c("black", "black")) +
  scale_x_continuous(breaks=1:8,
                     labels=c(as.character(1:7), "pooled")) +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text.x = element_text(size = rel(1)),
        axis.text.y = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        legend.text = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1)),
        strip.text.y = element_text(size = rel(1)),
        legend.position = "bottom",
        legend.title=element_blank())
```


```r
#dev.off()

################################################
## APPENDIX TABLE A6: Field Experiment: Effects of Information About the Tax
# Holiday on Compliance (no IPW, taxpayer type blocks)

blockedATE <- rbind.data.frame(
  # Treatment versus placebo: compliance
  difference_in_means(compliance_1416 ~ pooled_124_0, blocks = type,
                      se_type = "default", data = fieldex),
  # Treatment versus pure control: compliance
  difference_in_means(compliance_1416 ~ pooled_124_6, blocks = type,
                      se_type = "default", data = fieldex),
  # Placebo versus pure control: compliance
  difference_in_means(compliance_1416 ~ pooled_0_6, blocks = type,
                      se_type = "default", data = fieldex),
  # Treatment versus placebo: intended compliance
  difference_in_means(intended_1416 ~ pooled_124_0, blocks = type,
                      se_type = "default", data = fieldex),
  # Treatment versus pure control: intended compliance
  difference_in_means(intended_1416 ~ pooled_124_6, blocks = type,
                      se_type = "default", data = fieldex),
  # Placebo versus pure control: intended compliance
  difference_in_means(intended_1416 ~ pooled_0_6, blocks = type,
                      se_type = "default", data = fieldex)

)
blockedATE$outcome <- rep(c("Paid on Time", "Web Access"), each=3)
blockedATE$control <- rep(c("Treatment vs Placebo","Treatment vs Pure Control",
```

```
                             "Placebo vs Pure Control"), 2)

blockedATE[,c(1,2,4,6,11,18)]
```

```
##   coefficients  std.error  nobs         p.value      outcome
## 1  -0.00119254 0.00231141 10728 0.6059101491772 Paid on Time
## 2   0.00044087 0.00164977 17750 0.7892923665515 Paid on Time
## 3   0.00130733 0.00197545 14184 0.5081156375579 Paid on Time
## 4  -0.00306700 0.00095492 10799 0.0013230605725   Web Access
## 5   0.00151662 0.00068966 17842 0.0278834744340   Web Access
## 6   0.00518427 0.00089141 14267 0.0000000061641   Web Access
##                   control
## 1      Treatment vs Placebo
## 2 Treatment vs Pure Control
## 3   Placebo vs Pure Control
## 4      Treatment vs Placebo
## 5 Treatment vs Pure Control
## 6   Placebo vs Pure Control
```

```
##############################################
## APPENDIX TABLE A7: Survey Experiment: Pooled Lottery vs. Discretionary
# Benefit Conditions

survexp.results <- rbind.data.frame(
  difference_in_means(S1p4 ~ treat_discretion, data = survey_data),
  difference_in_means(S1p1 ~ treat_discretion, data = survey_data),
  difference_in_means(S1p3 ~ treat_discretion, data = survey_data),
  difference_in_means(S1p2 ~ treat_discretion, data = survey_data),
  difference_in_means(S1p5 ~ treat_discretion, data = survey_data)
)

survexp.results$outcome <- c("Rewards Go To The Same People As Always",
                             "Rewards Are A Waste Of Money",
                             "Worth It To Be Up To Date",
                             "Municipal Government Does A Good Job",
                             "Municipal Taxes Are Just")

survexp.results
```

```
##   coefficients std.error      df nobs statistic       p.value   conf.low
## 1     1.073239  0.216350  700.84 1595   4.96067 0.00000088249  0.648468
## 2    -0.021692  0.173490 1000.58 2325  -0.12503 0.90052396446 -0.362139
## 3    -0.465227  0.148811  924.95 2359  -3.12628 0.00182559391 -0.757274
## 4    -0.174690  0.139656 1047.47 2402  -1.25086 0.21126397602 -0.448728
## 5     0.039279  0.034865 1032.24 2382   1.12661 0.26017111621 -0.029135
##   conf.high alpha             term                                 outcome
## 1  1.498011  0.05 treat_discretion Rewards Go To The Same People As Always
## 2  0.318755  0.05 treat_discretion            Rewards Are A Waste Of Money
## 3 -0.173179  0.05 treat_discretion               Worth It To Be Up To Date
## 4  0.099347  0.05 treat_discretion    Municipal Government Does A Good Job
## 5  0.107693  0.05 treat_discretion                Municipal Taxes Are Just
##   condition2 condition1      vcov   design
## 1          1          0 0.0468072 Standard
```



```
## 2          1          0 0.0300989 Standard
## 3          1          0 0.0221448 Standard
## 4          1          0 0.0195038 Standard
## 5          1          0 0.0012156 Standard
```

```
#Threshold for FDR correction
# get and order the nominal p-values
ordered.ps <- survexp.results$p.value[order(survexp.results$p.value,decreasing=F)]
ordered.ps
```

```
## [1] 0.00000088249 0.00182559391 0.21126397602 0.26017111621 0.90052396446
```

```
comp <- (1:length(ordered.ps)/length(ordered.ps))*(.05)

FDR <- cbind(ordered.ps,comp,ordered.ps<=comp)
FDR
```

```
##         ordered.ps comp
## [1,] 0.00000088249 0.01 1
## [2,] 0.00182559391 0.02 1
## [3,] 0.21126397602 0.03 0
## [4,] 0.26017111621 0.04 0
## [5,] 0.90052396446 0.05 0
```

```
fdr <- max(FDR[,1][FDR[,3]==1])
fdr
```

```
## [1] 0.0018256
```

```
#Threshold for Bonferroni correction
bonf <- 0.05/length(ordered.ps)
bonf
```

```
## [1] 0.01
```

```
survexp.results$Bonf_reject <- NA
survexp.results$FDR_reject <- NA

survexp.results$Bonf_reject[survexp.results$p.value<=bonf] <- "yes"
survexp.results$Bonf_reject[survexp.results$p.value>bonf]  <- "no"

survexp.results$FDR_reject[survexp.results$p.value<=fdr] <- "yes"
survexp.results$FDR_reject[survexp.results$p.value>fdr]  <- "no"

survexp.results
```

```
##   coefficients std.error      df nobs statistic       p.value   conf.low
## 1     1.073239  0.216350  700.84 1595   4.96067 0.00000088249  0.648468
## 2    -0.021692  0.173490 1000.58 2325  -0.12503 0.90052396446 -0.362139
## 3    -0.465227  0.148811  924.95 2359  -3.12628 0.00182559391 -0.757274
## 4    -0.174690  0.139656 1047.47 2402  -1.25086 0.21126397602 -0.448728
```

```
## 5     0.039279  0.034865 1032.24 2382   1.12661 0.26017111621 -0.029135
##   conf.high alpha             term                              outcome
## 1  1.498011  0.05 treat_discretion Rewards Go To The Same People As Always
## 2  0.318755  0.05 treat_discretion            Rewards Are A Waste Of Money
## 3 -0.173179  0.05 treat_discretion               Worth It To Be Up To Date
## 4  0.099347  0.05 treat_discretion    Municipal Government Does A Good Job
## 5  0.107693  0.05 treat_discretion                Municipal Taxes Are Just
##   condition2 condition1      vcov   design Bonf_reject FDR_reject
## 1          1          0 0.0468072 Standard         yes        yes
## 2          1          0 0.0300989 Standard          no         no
## 3          1          0 0.0221448 Standard         yes        yes
## 4          1          0 0.0195038 Standard          no         no
## 5          1          0 0.0012156 Standard          no         no
```

G Pre-analysis plan (PAP)

G.1 Original PAP (registered July 23, 2014)

PRE-ANALYSIS PLAN

"Positive vs. Negative Incentives for Compliance: Evaluating a Randomized Tax Holiday"

July 23, 2014

This pre-analysis plan is registered (with a different title) to [NAME OF STUDY REGISTRIES]. The study has been approved by the Institutional Review Board [AT AUTHORS' HOME INSTITUTION].

Abstract

Can positive rather than negative incentives boost tax compliance in developing countries? We study a unique randomized policy innovation in Montevideo, Uruguay, in which the municipal government raffles tax holidays to good taxpayers who are current on past payments. Using unusual access to over-time tax payment records as well as survey data, we assess the impact of holidays on subsequent tax compliance, as well as citizens' attitudes towards taxation and governance. We also use field and survey experiments to study the effects of informing eligible and ineligible taxpayers about the rebate lottery—which has not been effectively advertised by the government. Our informational treatments allow us to compare the influence of priming negative incentives for tax compliance, such as fines and punishment for non-payment, with the positive inducement provided by the lottery.

Keywords: Tax compliance, developing countries, state capacity, positive vs. negative incentives; field experiment, natural experiment, information, lottery

Contents

List of Figures

List of Tables

1 Introduction

Developing countries face enduring obstacles to engendering tax revenue, a key facet of state capacity. To explain these obstacles, researchers often focus on the difficulty of sanctioning non-payment of taxes. According to this reasoning, states can boost tax compliance using *negative* incentives, such as fines and other punishments for non-payment; however, enforcement is difficult for states with poor administrative capacity. Thus, the extent of compliance is a function not only of the severity of sanctions but also the probability of their application. In this view, weak state institutions are responsible for persistent non-compliance in developing countries.

Yet, governments have occasionally used *positive* rather than negative incentives to generate tax compliance.[1] In this project, we study an unusual policy innovation in Montevideo, Uruguay, in which tax holidays are randomly assigned to eligible taxpayers. Since 2004, and across four kinds of taxes (property, vehicle, sewage, and head), the municipality of Montevideo has randomly selected taxpayers and—conditional on a recent history of good taxpaying—rewarded them with a year free of tax payments. Rather than being punished for poor tax-paying behavior, citizens are therefore rewarded for their history of tax compliance.

Insights from psychology and behavioral economics suggest that promised rewards may have substantially different effects than threatened punishments. Yet the impact of positive incentives for tax compliance has not to our knowledge been reliably assessed. Nor have researchers systematically compared the effectiveness of positive and negative encouragements to pay taxes. Indeed, given severe selection problems in non-randomized observational studies, it is typically challenging to answer questions about the effects of positive and negative inducements.

In the research we register with this document, we use the natural experiment provided by Montevideo's randomized tax holiday, and our unusual access to a panel of administrative tax payment records, to study the impact of winning a tax holiday. The lottery sets up a straightforward comparison, among eligible taxpayers—those with good tax-payment records—between winners and a randomly selected control group of eligible non-winners. Thus, we can assess whether winning the lottery, which not only provides a year free of tax payments but also informs many taxpayers of the existence of the program,

[1]For example, countries such as Argentina and Uruguay have raffled prizes to taxpayers who turn in receipts for value-added taxes. The objective of such policies is to increase the reporting of sales transactions by vendors.

influences subsequent tax compliance. Varying the period of elapsed time between winning the lottery and outcome measurement allows us to assess the persistence of effects on tax compliance. We also use household survey data for a sub-sample of winners and eligible non-winners to compare attitudes towards the equity and fairness of the tax system, as well as broader political attitudes, support for political incumbents, and so forth. For the natural experiment, we focus on the effects of winning any of the four types of taxes subject to tax holidays, as we have administrative tax payment data available for all four types of taxes.

However, the impact of the program is likely greater than these comparisons will suggest, because knowledge of the program may influence citizens to bring their tax accounts up to date (to become good taxpayers) in order to gain eligibility for tax holidays. We therefore also utilize a supplementary field experiment to assess whether informing citizens, including bad taxpayers, about the existence of the lottery affects subsequent tax compliance. In particular, in collaboration with the municipal government, we mail flyers stamped with the municipal logo—which appear very much like tax bills themselves—reminding taxpayers of the due date for taxes and providing experimentally varied messages to taxpayers. For these informational interventions, we focus on property (real estate) taxes, as these are the most visible and important of the four taxes we study.

In more detail, we experimentally vary the content of messages to study the effects of negative and positive inducements for tax compliance. Thus, we compare taxpayers who are 1. simply reminded of an impending due date for real estate taxes (placebo control group) to: 2. those who are reminded/informed of the existence of the tax holiday lottery and 3. those who are reminded/informed of the existence of punishments for non-payment of taxes.[2] For group 2., we further subdivide the treatment group into those who receive 2A. no information about the probability of winning the lottery and 2B. those who are told their probability of winning (which we compare to taxpayers' priors about the likelihood of winning, estimated from follow-up survey data on the control group). Finally, in addition to these treatments priming individual rewards and punishments for payment or non-payment of taxes, we use alternative interventions that additional prime either 4. the social rationale for the lottery or 5. the social rationale for punishment of non-compliance with taxes. Our social treatments may manipulate the individual normative/expressive benefits of paying taxes and thus may influence

[2]See subsection 3.2.1 for the text of the informational interventions.

compliance in distinct ways. The results of our informational experiment will allow us to compare the impact of manipulating negative and positive incentives for tax payment, thus shedding light on basic motivations for tax compliance, and also evaluate the best means of boosting future impact through informational interventions that effectively advertise the policy. The findings may therefore suggest how to maximize impact when the program is scaled up or transported to other settings.

Our study has several key advantages. First and foremost, the randomization of the natural and field experimental interventions sets up straightforward, unconfounded comparisons between the treatment groups, aiding inferences about the causal impact of the interventions. Second, we make use of detailed, unobtrusive outcome measures—namely, individualized administrative records on tax compliance, extent of tax debt, and so forth. These records are akin to turnout data in voter mobilization experiments, but unusually they allow unobtrusive measurement of *tax payment*. Third, the experimental realism of our interventions is quite high, since we designed flyers that match the format of tax bills and arrive bearing the municipal logo—just as would real interventions designed to boost compliance by publicizing negative or positive incentives for tax payment.

Finally, our factorial experimental design and supplementary measurement using household surveys will shed tentative light on the mechanisms through which positive and negative inducements generate compliance. For example, winning the lottery might shape compliance by boosting income or altering habits (which could conceivably lead to a negative impact on compliance), by notifying uninformed taxpayers of the existence of the lottery, or by shaping beliefs about the equity and transparency of the tax system as well as broader political attitudes. Though distinguishing such mechanisms is tricky, as a recent literature on causal inference highlights, we use several strategies to explore possible *reasons* for the impact of promised benefits or threatened punishments. For example, we compare the effect of winning a tax holiday lottery to the effect of merely informing eligible non-winners about it (and use data on the proportion of uninformed taxpayers from our household surveys to generate instrumental-variables estimators). We also test for heterogeneous effects that shed light on the impact of rewards, for example, we assess whether effects depend on the size of past tax arrears, on the theory that manipulating material incentives for payment are greatest for "taxpayers at risk" (e.g., good taxpayers who have not always been good, or bad taxpayers who have not always been bad). We also use survey as well as field and natural experiments to compare the effects of negative and positive inducements on beliefs about the tax system as well as broader political attitudes; effects on such outcomes are of interest because beliefs and attitudes may in turn influence compliance, but also because they represent independently interesting and important consequences of the tax holiday lottery.

In the rest of this document, we sketch the theory that motivates our study, particularly our comparison of negative vs. positive incentives for tax compliance (Section 2); describe our empirical strategy, including our natural and field experimental designs and the timing and nature of data collection (Section 3); discuss our outcome measures (Section 4); and outline our hypotheses and statistical tests, including adjustments for multiple comparisons (Section 5).

2 Theory

2.1 What hinders tax compliance?

Engendering tax compliance is a routine problem, especially in developing countries. According to most accounts, states with weak administrative capacity fail to "penetrate" society sufficiently, thereby allowing citizens to avoid payment of taxes by not imposing credible penalties for non-payment. Thus, scholars typically see information and monitoring problems at the heart of the problem, in which lack of state capacity explains failures to elicit compliance.

This account relies on an underlying behavioral theory: the decision not to pay taxes is driven by the benefit of evasion, minus the cost of punishment discounted by the probability of detection. The problem, according to this theory, is thus that in developing countries the probability of punishment for non-compliance is negligible. As a corollary, collecting taxes may be politically unpopular, which may provide an additional incentive for elected governments to opt not to enforce penalties for non-payment of taxes. In sum, failure to generate tax revenue is seen as a problem of enforcement that is due in part to weak state capacity.

Yet, in developed and even developing countries alike, many people do pay taxes—even when penalties and chances of punishment for evasion are quite low. This is puzzling from the perspective of a focus on negative incentives. Indeed, the simple decision framework we develop in this section suggests that if compliance were only driven by the benefit of evasion minus the expected cost of punishment, nobody would pay taxes at all. Some lab experimental evidence also suggests that compliance behavior can be insensitive to the cost of penalties or the probability of punishment. What then explains why in many settings people *do* pay taxes?

We investigate in this study the impact of *benefits* for tax compliance—in particular, the positive incentive provided by the randomized tax holiday for good taxpayers in Montevideo. Individual benefits play a key role in theories of voting over tax policy, but the implications for compliance are underexplored. Note that benefits of paying taxes may be *material* (e.g. a direct individual or group return to taxes, in the form of public spending) or *expressive/normative*, e.g., due to social preferences, equity considerations, altruism, or burden sharing. A tax holiday lottery such as the one we study might shape both the perceived material and expressive benefits of paying taxes; moreover, different ways of framing the existence of the lottery (for instance, emphasizing its social rationale) might influent the individual expressive/normative benefits of payment. The impact of this lottery has not been widely studied, and thus our research question is empirical: do individual rewards influence tax *compliance* behavior, and if so, why?

Note that the tax setting we study is somewhat different than in the classic compliance problem. Here we study municipal real estate, vehicle, sewage, and head taxes, in which the tax due is known by the taxing authority, e.g. because the value of assets is appraised). Thus, information and monitoring problems are plausibly less severe than in many settings, such as the payment of income tax.[3] Nonetheless, enforcement remains a central issue in this context: the municipal government can decide to pursue delinquent taxpayers more or less aggressively—sometimes failing to pursue bad taxpayers for a period of many years and sometimes using the courts to expropriate the property of bad taxpayers—and the outcome of any individual renegotiation is ex-ante uncertain for taxpayers. As we discuss below, Montevideo's city government has frequently issued general amnesties for bad tax payers, and it often renegotiates debt with individual taxpayers as well. For this reason, punishment for non-compliance is very far from certain, which tends to make tax payment mysterious from the point of view of the classic decision calculus we sketch next—in which the choice not to pay taxes is driven by the benefit of non-payment or delay in payment, minus the expected cost of punishment. Our setting therefore provides a useful vehicle for comparing the impact of positive vs. negative incentives for tax payment: in particular, how the probability of punishment and the possibility of rewards condition tax compliance; and how manipulation of other factors—such as the *benefits* of paying taxes—may engender greater tax compliance.

[3] In this sense, the problem facing the state is akin to that of a credit-card company faced with non-payment by consumers: the value of debt is known, and hiding the amount owed is not an option, but the company can still face substantial barriers to collecting taxes.

2.2 The role of positive incentives

A simple formalization of the decision problem highlights the difficulty of encouraging tax compliance. Let y be an asset value, t be the annual tax rate, and z be the unpaid annual amount of taxes due; with full nonpayment, $z = ty$. The expected utility of full nonpayment in any year is thus

$$z - pc, \tag{1}$$

where c is the penalty for nonpayment and p the probability of punishment. In the setting we study, the cost of punishment c could include (1) fines and interest charges for delayed payments, and ultimately (2) losing one's house or other property.[4] However, a taking of property by the city government is not certain, even for the worst taxpayers, given the legal costs involved. Moreover, the possibility of a future amnesty for bad taxpayers, which happen frequently as we discuss below, means that the probability of punishment p is very far from one in this context.[5]. The fine for nonpayment c is often small.

Under the policy we study, good taxpayers win a year free of tax payments with probability 1/5,000 in any tax payment period. Thus, the expected utility of paying the full year's taxes this year is (without discounting)

$$\frac{1}{5,000}z - z. \tag{2}$$

Note that the direct material benefit provided by the lottery is limited, relative to the cost of compliance here: in order to gain eligibility for the lottery, one has to pay a year's worth of taxes z, while gaining a year free of tax payments in the following year only with probability $1/5,000$. Indeed, comparing (2)

[4] Property owners also cannot sell their houses until clearing their property tax accounts.
[5] Note that $(1 - p)$ is the (unknown) probability of amnesties and individual renegotiations

with (1), we see that paying taxes is only optimal when

$$pc > \frac{9,999}{5,000}z. \tag{3}$$

In words, the cost of the penalty discounted by the probability of punishment must be essentially as large as two full years of taxes due on the asset.

A natural question that arises is: does manipulation of the expressive or perceived benefit of paying taxes further influence compliance? Let b be the perceived benefit of paying taxes, in terms of the social return or legitimacy of taxation. The expected utility of paying taxes is then $(1/5,000)z - z + b$, so tax payment occurs whenever

$$pc > \frac{9,999}{5,000}z - b, \tag{4}$$

which is satisfied more easily than is (3). Depending on the size of b—the perceived benefit of paying taxes—it could be incentive compatible for taxpayers to comply, even when pc is small.

Our theory is thus that beyond providing a direct benefit—which is vanishly small in expectation—Montevideo's lottery may also affect b, e.g. by influencing perceptions of the fairness and equity of taxes and/or the individual normative benefit of paying taxes. In some sense, this must be true if the lottery policy affects compliance, at least if people make payment decisions according to (4). Among property owners, the average annual value of property taxes is over US\$1,000 (24,000 Uruguayan pesos), which is non-trivial[6]; but the *expected* value of the lottery US\$1,000/5,000, or about twenty US cents. We therefore expect the material benefits of the lottery to influence compliance only if (1) people misperceive the probability of winning the lottery or (2) the benefits provided by the lottery exceed the expected material payoff, e.g. because rewards influence behavior differently from punishments, or because the transparency and legitimacy of the lottery policy influence expressive benefits b.

This decision framework also suggests that manipulating benefits may have a bigger effect on certain kinds of taxpayers, in terms of pushing them over the threshold to compliance:

- Taxpayers who owe less (e.g. for whom the amount owed $z < ty$);
- Those with higher income or capacity to pay, relative to asset values; and

[6] This is an estimate from an interview with the IM; we will verify this using our probability sample of tax records.

- Those with high subjective p, i.e., who overestimate the probability of winning the lottery).

We discuss these hypotheses about heterogeneous effects for such marginal taxpayers—those with small amounts of indebtedness or imperfect but not terrible records of compliance—further in Section 5.

3 Empirical Strategy and Design

The innovative tax holiday policy developed by Montevideo's City Hall (Intendencia de Montevideo—IM) seeks to improve tax compliance by providing positive incentives for good taxpaying; it also counteracts negative perceptions among citizens of forgiveness for non-compliance. The policy was initiated by the center-left government of the Frente Amplio in the context of an amnesty for many delinquent taxpayers following the economic crisis of 2002. The idea was to counteract perceived negative incentives of the tax amnesty.[7] As officials at the IM have told us, the economic crisis generated a dilemma: how to lower the burden for those under dire circumstances while at the same time continuing to promote compliance. The lotteries were their answer.[8] After almost ten years, however, no evaluation of the program's effectiveness exists.

To select taxpayers for holidays, the government uses the results of Uruguay's National Lottery, which posts online five random digits that indicate winning lottery numbers. Taxpayers whose four-digit IDs correspond to the final four digits of the winning National Lottery numbers are selected as the provisional winners of tax holidays.[9] The National Lottery frequently posts lottery results, which the municipal government uses to select winners of rebates for each type of tax. The municipality sends a letter to eligible winners whose tax accounts are paid up in the previous fiscal year (i.e., "good" taxpayers) indicating that they should register for a year free of tax payments; registration allows the

[7] For discussion of the initiation of the lottery, see http://historico.elpais.com.uy/Suple/LaSemanaEnElPais/04/10/29/lasem_ciud_118264.asp; http://www.montevideo.com.uy/notnoticias_66228_1.html; and http://www.180.com.uy/articulo/14284.

[8] In October 2013, the municipality announced a renewed amnesty for certain bad taxpayers, underscoring the difficulties of cracking down on non-compliance. There have been amnesties in 2004, 2008, and 2010, among other years. See http://www.montevideo.com.uy/notnoticias_66228_1.html and http://www.180.com.uy/articulo/14284.

[9] The randomization occurs through the selection from balls from an urn, as described in Spanish at http://www.loteria.gub.uy/Juego_Loteria.php. For an example of posted lottery results, see http://www.loteria.gub.uy/ver_resultados.php?vdia=21vmes=6vano=2013. Winning taxpayer numbers are posted at http://www.montevideo.gub.uy/sorteosBP/pages/sorteosBuenosPagadores.xhtml.

city government to screen winners to ensure that they are physical persons and not, e.g., corporations. For good taxpayers, the probability of winning any given rebate lottery is 1/10,000; however, some taxpayers who would have won a tax lottery are not good taxpayers or physical persons or do not present themselves to the city government after being notified. Also, the municipality grants such holidays six times a year (February, April, June, August, October, and December) for head and sewage taxes, and three times a year (March, July, and November) for vehicle and property (real estate) tax. Because the real estate and vehicle taxes are paid three times a year, with two lotteries in the interim (and winning either lottery grants a year free of tax payments), for those taxes the probability of winning a year free of taxes in any tax period is 1/5,000. We focus in the project mainly on the impact of the tax holiday for property taxes, though we also estimate effects of winning the lottery for the other three taxes.

We will use both administrative data (tax payment records) provided by the municipality and household survey data (both described further in Section 4) to study the effects of the lottery as well as the effects of our informational interventions. Due to the greater cost involved in data collection for the household surveys, our household survey will be administered to a random sample of individuals for whom we have administrative/tax payment data. Tables 3.1 and 3.2 in the next sub-sections give the sample sizes for each of the two datasets, which we justify using power calculations in Section 3.4. In the next subsections, we describe how we will use comparisons between the cells of Tables 3.1 and 3.2 to estimate different kinds of causal effects.

In more detail, we asked the municipality for taxpayer records for all eligible lottery winners since 2004 (we currently believe that is circa 7,200 taxpayers but may be fewer), and a random sample of eligible ("good") taxpayers who have not won the lottery of approximately the same size (see Table 3.1)); below, we describe the sampling procedure for the latter group. These samples comprise the treatment and control groups in the natural experiment. For the study group for the field experiment, we asked for a random sample of eligible ("good") taxpayers (N=14,250) and a random sample of ineligible ("bad") taxpayers (N=14,250) (see Table 3.2). We then draw a sub-sample of size 8,000 (N=2,000 from the natural experimental study group, N=6,000 from the field experimental study group) for household surveys. The administrative data list taxpayer names, addresses and in some cases phone numbers; our survey firm will use those data to track down sampled individuals. In cases where sampled individuals cannot be found, are not physical persons, or refuse to participate, we sample replacement households

at random from the taxpayer records included in the study group. The survey will be conducted by the Uruguayan firm CIFRA.

3.1 Natural experiment: the effect of winning a tax holiday

The design of the lottery allows us readily to estimate the effects of winning a tax holiday, among good taxpayers. In particular, we will use a time-series panel of administrative data (2004-2013) to assess the effects of winning the lottery on subsequent tax payments, comparing the payment history at $t+1$, $t+2$, $t+3\ldots$ of lottery winners to a control group of eligible non-winners, where t is the year (or portion of year) in which winners won the lottery.

Constructing appropriate treatment and especially control groups requires some care. For the treatment group, we will have access to tax payment data for all winners of each tax lottery since 2004. We will check to ensure that the last four digits of winners' taxpayer IDs (current account numbers) match the last four digits of the winning National Lottery number in the corresponding lottery. Municipal officials estimate the number of winners of tax holidays across all lotteries at 7,200 (though we believe the number may be smaller); we will receive the data on approximately July 25, 2014, shortly after filing this pre-analysis plan (see sub-section 3.3.1). We will request data for both eligible and ineligible taxpayers whose IDs correspond to the four-digit numbers of lottery winners.[10]

For the control group, note first that the right counterfactual group for winners of a particular lottery are taxpayers who were eligible to win as of the date of *that* lottery, based on being current on their tax payments over the previous year.[11] To select a random sample of eligible lottery non-winners for *each* lottery since 2004, we randomly generate a single four-digit number for each lottery since 2004, replacing at random any numbers that coincide with the IDs of lottery winners.[12] The municipality will then locate all taxpayer IDs that end in each random four-digit number and identify those who were eligible to win the corresponding lottery, based on their tax compliance status at the time the

[10]The treatment and control groups are comprised only of eligible taxpayers, but data on ineligible taxpayers will be useful for a number of purposes, such as placebo tests

[11]If we used a control group of currently eligible taxpayers, we would risk bias: the treatment group would include only taxpayers who were eligible to win (good taxpayers) as of the date of each lottery, while the control group would include a mix of taxpayers who were eligible and who were ineligible as of the date of each lottery.

[12]These data will be provided to us free of charge, yet recovering specific taxpayer records requires effort from the municipal bureaucracy; thus, we cannot use a census of administrative data in the control group, as we do in the treatment group.

lottery took place. Thus, our random procedure for constructing the control group exactly mimics the random process that created the treatment group of lottery winners. The average number of winners per lottery is approximately 20, and thus so is the average size of our control group for each lottery, given the manner in which we construct it. Both the treatment and control groups are comprised of random samples from the population of good taxpayers in Montevideo, as of the date of each lottery.

Table 3.1: **Natural Experiment: Sample Sizes and Data Sources**

	Lottery winners (Good taxpayers)	**Lottery non-winners (Good taxpayers)**
Sample Size and Data Sources	Admin. Data, $N \approx 7,200$* (+ Surveys, N=1,000)	Admin. Data, $N \approx 7,200*$ (+ Surveys, N=1,000)

Total N=2,000 (Survey data); $N \approx 14,400$ (Administrative data). * The municipality estimates there are approximately 7,200 winners of the lottery since 2004 but we believe the number may be smaller. Our sampling procedure will ensure that the number of lottery winners approximately equals the number of eligible non-winners.

Our approach has the advantage that the treatment and control groups will be approximately the same size (balanced design), which is typically the most efficient design conditional on the overall size of the study group.[13] Moreover, the procedure naturally distributes the study group across the four types of taxes in proportion to the prevalence of winners of lotteries for each tax. Here we in fact have a series of mini-natural experiments, in which each lottery generates a treatment group of winners and a control group of non-winners. Thus this is effectively blocked random assignment, where the blocks are individual lotteries; however, the probability of winning any lottery is the same in every block (1/10,000), so we will not need to adjust for blocked assignment in our analysis. Finally, because our treatment and control groups are both random samples from the population of eligible ("good") taxpayers in each corresponding lottery, we can use our natural experiment to estimate population average treatment effects (PATEs) for the population of good taxpayers.

This design allows us to use straightforward comparisons to estimate the effects of winning a lottery, among good taxpayers. For example, mean differences between the columns of Table 3.1 estimate the average causal effect of winning the lottery.[14] Varying the number of included tax years subsequent to

[13]Due to our sampling method, the size of the treatment and control groups are random variables; however, this will not lead to bias in treatment effect estimators due to independence of the denominator and the ratio of the numerator to the denominator, in estimators of treatment effects such as the average causal effect.

[14]We may also estimate spillover effects by comparing average payments of winners' and non-winners' neighbors. This is

the date of each lottery allows us to assess the persistence of effects. We discuss the analysis of the natural experimental data further in Section 5.

3.2 Field experiment: positive vs. negative incentives

The comparison of winners and losers likely underestimates program impact, because the lottery may induce some bad taxpayers to bring their accounts up to date to gain eligibility for the lottery. Moreover, the lottery has apparently not been effectively advertised by the municipal government, which raises the question: if Montevideo—or another municipal government—were to use an informational campaign to tell citizens about the existence of the rebate lottery, what sort of interventions would be most effective in boosting tax payments? Finally, what mechanisms may explain any effect we find of winning the lottery on tax payment?

To answer these questions, and to probe basic motivations for tax compliance, we use a field experiment in which we provide varied information to a random sample of taxpaying households. Our informational experiments allow us a unique opportunity to compare the effects of positive incentives provided by the lottery to the effects of negative incentives: for example, messages about sanctions such as fines for non-payment?

3.2.1 Text of informational treatments

In more detail, we collaborated with the municipal government to design and mail to households flyers printed with messages that correspond to the following treatment conditions:

1. *Placebo control* (reminder that tax bill is due);

2A. *Individual reward 1* (informing citizens of existence of lottery);

2B. *Individual reward 2* (also priming probability of winning);

3. *Individual sanction* (existence of fines/ punishment for non-payment);

attractive as some of the informational effect of the lottery could conceivably work through neighbor-to-neighbor communication. However, this introduces non-trivial logistical obstacles because it involves linking a large number of geo-located physical addresses to taxpayer records—rather than sampling taxpayer records that are linked to physical mailing addresses, as we do here. We are uncertain whether the inferential benefit would outweigh the cost, as we have other ways (described next) to estimate the effects of learning about the lottery's existence on future compliance. If we decide to estimate spillover effects to neighbors, we will file an amendment to this pre-analysis plan in advance of the additional data collection and analysis.

4. *Social reward* (emphasizing social rationale for lottery); and

5. *Social sanction* (emphasizing social rationale for fines/punishment).

The experimental realism of our treatments is very high: when folded for mailing, the logo of the municipality is visible, and upon delivery to households the flyers appear *identical* to municipal tax bills.[15] The experience of receiving and opening a tax bill on which the municipality prints encouragements to pay taxes would thus be very similar to the experience of receiving these flyers stamped with the municipal government's logo. Figures 2-7 show the messages on the Spanish-language flyers, while Figure 8 shows the back side of the flyer with the municipal logo.

The translated text of the flyers is as follows:

PLACEBO CONTROL (Figure 2):

Dear neighbor:

We want to remind you that **the second payment of property taxes is due in July**. If you have not received your bill, you can obtain a duplicate copy on our web site (www.montevideo.gub.uy).

For questions: Phone the tax department (1950 300)

INDIVIDUAL REWARD 1 (Figure 3):

Dear neighbor:

We want to remind you that **the second payment of property taxes is due in July**. If you have not received your bill, you can obtain a duplicate copy on our web site (www.montevideo.gub.uy).

The municipal government of Montevideo wants to reward good taxpayers. **If you pay on time, you will be automatically entered in a lottery to win a year free of property tax payments.**

Lotteries occur every other month of the year in conjunction with the National Lottery. The winners will be duly informed and the results of the lottery will be published on the web site of the city government.

You can be the next winner!

For questions: Phone the tax department (1950 300)

[15] For example, our survey interviewers exclaimed that the folder flyers appeared to them to be tax bills.

INDIVIDUAL REWARD 2 (Figure 4):

Dear neighbor:

We want to remind you that **the second payment of property taxes is due in July**. If you have not received your bill, you can obtain a duplicate copy on our web site (www.montevideo.gub.uy).

The municipal government of Montevideo wants to reward good taxpayers. **If you pay on time, you will participate automatically in a lottery to win a year free of property tax payments.**

In each lottery, 1 of every 5,000 households receives this benefit.

Lotteries occur every other month of the year in conjunction with the National Lottery. The winners will be duly informed and the results of the lottery will be published on the web site of the city government.

You can be the next winner!

For questions: Phone the tax department (1950 300)

INDIVIDUAL SANCTION (Figure 5):

Dear neighbor:

We want to remind you that **the second payment of property taxes is due in July**. If you have not received your bill, you can obtain a duplicate copy on our web site (www.montevideo.gub.uy).

Those who do not pay on time may be subject to fines and charges. The municipal government of Montevideo may take legal and administrative actions to enforce the rules where applicable.

Pay on time, avoid fines and charges!

For questions: Phone the tax department (1950 300)

SOCIAL REWARD (Figure 6):

Dear neighbor:

We want to remind you that **the second payment of property taxes is due in July**. If you have not received your bill, you can obtain a duplicate copy on our web site (www.montevideo.gub.uy).

The municipal government of Montevideo wants to reward good taxpayers. **If you pay on time, you will participate automatically in a lottery to win a year free of property tax payments.**

Lotteries occur every other month of the year in conjunction with the National Lottery. The winners will be duly informed and the results of the lottery will be published on the web site of the city government.

The municipal government of Montevideo conducts this lottery to recognize good taxpayers for their contribution to constructing a city that is more just and better for all.

For questions: Phone the tax department (1950 300)

SOCIAL SANCTION (Figure 7):

Dear neighbor:

We want to remind you that **the second payment of property taxes is due in July**. If you have not received your bill, you can obtain a duplicate copy on our web site (www.montevideo.gub.uy).

Those who do not pay on time may be subject to fines and charges. The municipal government of Montevideo may take legal and administrative actions to enforce the rules where applicable.

Fines and charges are a sanction to those who do not pay their taxes and do not contribute to constructing a city that is more just and better for all.

For questions: Phone the tax department (1950 300)

The varied messages printed on our flyers allow us to study the following topics, to which we return in Section 5, Hypotheses and Tests.

Positive vs. negative incentives. How do the effects of the positive incentives provided by the lottery compare to the effects of negative incentives/sanctions? One previous informational experiment found that emphasizing fines and other legal consequences of non-payment, in a message similar to our Individual Sanction condition, increased tax compliance in Argentine municipalities by more than 4 percentage points.[16] As discussed in the next sub-section, our design replicates this intervention and compares it to the effect of information about the rebate lottery, among both good and bad taxpayers. In a developing-country context in which tax enforcement is routinely difficult, the effects of information about sanctions should provide a useful benchmark for comparing the effects of information about rewards. (For both treatments, we will use instrumental-variables analysis to study the effects on those who *learn* about sanctions and rewards through our interventions—see Section 5).

Informational effects. Winning the lottery provides a temporary income boost, and it may shape broader attitudes towards the fairness and social benefits of taxes; yet, it may also affect future tax compliance by shaping knowledge of the lottery and perceptions of the likelihood of future rebates.[17] To estimate pure informational effects, we compare good (eligible) non-winners who receive a flyer informing them of the existence of the lottery to a control group of eligible non-winners. We can then compare this to the effect of winning the lottery among eligible taxpayers, which combines informational effects with other factors that may shape compliance behavior. (Again, an instrumental-variables estimator gives an estimate of the effect of learning about the lottery for uninformed good taxpayers influenced by our informational treatment—see Section 5).

Measuring effects among bad taxpayers. A limitation of the government's program, from an inferential point of view, is that the lottery is restricted to good taxpayers. To assess the broader impact of the program—e.g., the effect of giving bad taxpayers greater incentives on the margin to pay their taxes and thus gain eligibility for the lottery—we use our informational intervention to estimate the effect of informing ineligible ("bad") taxpayers of the existence of the lottery, compared to the placebo control group (using both intent-to-treat and instrumental-variables analysis). While "bad" taxpayers may constitute a relatively small group during good economic times (say, 10-15% of taxpayers, though

[16] Carlos Scartascini and Lucio Castro. 2013. Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment. Manuscript. Abstract at http://www.cscartascini.org/work-in-progress.

[17] Winning lotteries may also engender self-reinforcing beliefs about individual merit, which might have broader political implications. See e.g. Di Tella, Rafael, Sebastian Galiani, and Ernesto Schargrodsky, 2007, The Formation of Beliefs: Evidence from the Allocation of Land Titles to Squatters, Quarterly Journal of Economics, 122: 209-41.

we will have more precise estimates later), during crises the size of this group can grow substantially; moreover, assessing affects for bad taxpayers may give some insight into likely effects of similar interventions in settings with larger numbers of bad taxpayers than Montevideo.

Our data analysis also helps assess other hypotheses, as discussed in Section 5.

3.3 Treatment assignment

To create our study group for the informational experiment, we asked the government to draw a random sample of 28,500 administrative records, including 14,250 taxpayers who are eligible to win the lottery for exoneration of payment starting in July 2014 ("good taxpayers"), and 14,250 who are ineligible based on not being up to date on payments over the past year ("bad taxpayers").[18] We then randomized these eligible and ineligible taxpayers with equal probability to one of the six treatment groups. The sample sizes shown in Table 3.1.[19]

Due to cost considerations, we will only survey a random sub-sample of households in each treatment group, rather than a census of the 28,500 households; thus, in the Placebo Control group among good taxpayers, we sent flyers to 2,850 households, but will gather household survey data for a random sample of 1,000 of these households. We will assess effects of our treatments on tax compliance using administrative data on all 28,500 households, while we assess effects on attitudes, knowledge of the program, and other outcomes using data from our household survey.

There are two important wrinkles. First, the list of addresses provided by the municipality includes non-physical persons (e.g., corporations) as well as physical persons living in households. The municipality does not have a ready way to distinguish these persons (which is why they must screen lottery-winning taxpayer accounts before awarding a year free of tax payments). Thus, our survey firm will also screen these addresses, interviewing only physical heads of household and/or physical persons who have responsibility for paying household taxes. We initially gave our survey firm a list of 8,000 randomly selected households, which it will use to fill the quota of 6,000 households in the field exper-

[18]We worked with the government on the technical requirements of drawing the sample; when we have the requisite data, we will compare covariates of our random sample of taxpayers to covariates in the population of taxpayer records, adjusting for the oversampling of ineligibles.

[19]Note that we sometimes conceptualize treatments 2A and 2B as the same treatment (i.e., we often pool across these treatment conditions), so the overall size of the 2A plus 2B group is kept equal to the other four groups.

Table 3.2: **Field Experiment: Treatment Conditions and Sample Sizes**

	Lottery non-winners (Good taxpayers)	**Sample of ineligibles (Bad taxpayers)**
1. Placebo Control	Admin. Data, N=2,850 (+ Surveys, N=1,000)	Admin. Data, N=2,850 (+ Surveys, N=1,000)
2A. Individual Reward/ Lottery	Admin. Data, N=1,425) (+ Surveys, N=500)	Admin. Data, N=1,425 (+ Surveys, N=500)
2B.(+ Probability of Winning)	Admin. Data, N=1,425) (+ Surveys, N=500)	Admin. Data, N=1,425 (+ Surveys, N=500)
3. Individual Sanction	Admin. Data, N=2,850 (+ Surveys, N=1000)	Admin. Data, N=2,850 (+ Surveys, N=1000)
4. Social Reward*	N=2,850 (Admin. Only)	N=2,850 (Admin. Only)
5. Social Sanction*	N=2,850 (Admin. Only)	N=2,850 (Admin. Only)
TOTAL N	Admin. Data, N=14,250 (+ Surveys, N=3,000)	Admin. Data, N=14,250 (+ Surveys, N=3,000)

Total N=6,000 (Survey data); N=28,500 (Administrative data). * For these conditions, only administrative outcome data will be gathered. See sub-section (3.2.1) for text of flyers in the information experiment.

iment, after eliminating non-physical persons and accounting for non-response; should this list not be sufficient to reach the desired sample size, the firm will return to us for additional randomly-selected households. We will follow the same procedure for household surveys in the natural experiment. Control over the process of random substitutions from the list of 28,500 taxpayer addresses will allow us to calculate the real rate of non-response as well as to estimate the proportion of physical persons among the population of taxpayer accounts.

Second, we learned only after obtaining contact data (household addresses) for our 28,500 sampled households and mailing our flyers to them in late June 2014 that an important number of taxpayers in our sample (6,789 accounts, or around 25%) pay their taxes in entirety for the full year, or for other reasons do not pay taxes in July.[20] Data analysis shows these are predominantly good taxpayers—and moreover they cannot be readily influenced by the treatments in our field experiment because they will not pay taxes in our initial period of outcome measurement. We therefore plan to exclude these households from both our household survey and administrative data analysis. Thus, the population

[20]For example, pensioners pay taxes in November.

from which our household survey and administrative data samples are drawn should be conceptualized as "all taxpaying households with bills due in July 2014."[21] Note that these are to a greater extent "taxpayers at risk" who may be affected by our intervention, in that the excluded group that does not pay in July consists disproportionately of good taxpayers.

One important issue here is a loss of statistical power relative to our planned design (see section 3.4), given that some of the households in our original sample households do not pay bills in July and thus cannot be affected by our intervention (in the short run, at least). We therefore amended our protocol to collect another round of administrative data to fill out our 28,500 cases with additional accounts paying taxes in July. We then randomized these cases to treatment groups and mailed flyers in two additional sub-phases, a Phase II.A with flyers mailed beginning July 5, 2014 and arriving by July 10, and a Phase II.B mailed from July 11 to July 26 (see sub-section 3.3.1). Because our flyers must arrive in advance of the tax due date, in this Phase II we screened records to include only those taxpayers who receive their tax bills after July 10, 2014. Thus, for Phase II the population from which we sampled taxpayer records should be conceptualized as "all taxpaying households with bills due after July 10, 2014." Assignment of due date is not randomized; however, qualitative evidence suggests it is haphazard. For example, the zones used by the municipality to order mailings of tax bills do not correspond to postal or other jurisdictional demarcations, and the mailing does not appear to follow a strong geographic logic. We will use geo-coding of the zones to map the order in which tax bills are sent, and compare covariates of individuals who receive tax bills earlier and later; we can also compare estimates from Phase I data (in particular, households with bills due prior to July 10 and those due later) to see if these estimates are statistically distinguishable. We think that obtaining internally valid estimates for a larger group of taxpayers is a first-order concern and this prompts our collection of these additional Phase II data.

3.3.1 Timing of intervention and data collection

Figure 1 describes the timing of our field experimental intervention, the municipality's mailing of tax bills, and data collection. Our flyers/informational treatments were generated, addressed, and then distributed by a company we hired beginning on June 27 and continuing until July 1 (Phase I). On

[21] From this population, we draw a stratified random sample with an oversample of ineligible/bad taxpayers.

July 7, we received from the postal service a list of addresses where these flyers were not received (so we could replace these addresses for purposes of the household survey). As discussed in the previous sub-section, we began the Phase II (A and B) mailings of flyers on July 5, targeting households with intermediate tax due dates for Phase II.A and the latest tax due dates for Phase II.B.

We can be confident that the households in our study group received our informational treatments before the due date for paying taxes—and in most cases, before the physical receipt of tax bills. Mailing of tax bills by the municipality began on July 1 and continues until July 21 in a staggered fashion, according to which different zones of the city receive the tax bill at different times. The due date for tax payments is also staggered and follows the order of delivery of tax bills, with the due data approximately 8 days after the receipt of bills. We were not initially aware of how the staggering would occur, and the city zones used by the municipality do not correspond to postal codes; thus, we did not time the mailing of our informational treatments to coincide exactly with the arrival of the tax bill. However, all households in the Phase I mailings of informational treatments should have received flyers before July 8, the earliest possible due date. Moreover, for Phases II.A and II.B, we gathered information on due dates for each household from the municipality and grouped households into Phases II.A and II.B accordingly, to ensure that flyers would arrive in advance of due dates and (usually) tax bills. The greater risk is that in some cases (especially Phase I mailings) our flyers could arrive several days or even weeks in advance of tax bills and due dates; this lag between receipt of our informational treatments and receipt of tax bills and due date—and our even later measurement of attitudinal outcomes using household surveys—could weaken the effect of our treatment, for households where the lag is long. However, we will have data on the date at which bills were received and the due date for payments; so in principal we can compare households who received informational treatments closer to the due date with those where the lag was greater. [22]

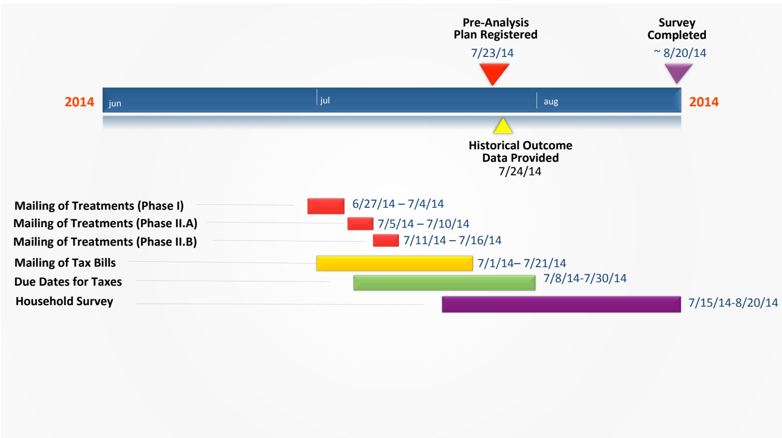
Historical outcome data for the natural experiment will be received from the municipality on approximately July 24. Our survey was fielded beginning on July 15 and will continue through approximately August 20, at some point after which we will receive survey outcome data from CIFRA. This pre-analysis plan is registered effective July 23, 2014.

Thus, we are filing this pre-analysis plan after the beginning of our field experimental intervention

[22] See discussion of analysis of heterogeneous effects in Section 5.

and household survey, but before any outcome data are collated or analyzed.

Figure 1: Timing of intervention and data collection



3.4 Power calculations

In this sub-section and the associated Figures 9-12, we present formal justification for the sample sizes in our natural and field experiments (Tables 3.1 and 3.2). For several of our power analyses, we take as a benchmark the informational experiment of Castro and Scartascini (2013), who estimate effects of informational treatments on tax compliance of over 4 percentage points using *negative* incentives for compliance (reminding taxpayers of fines for non-compliance, as in our Individual Sanction

treatment).[23] However, we calculate the probability of rejecting the null hypothesis of no effect, given various true effect sizes. This effect could be, e.g., the difference in subsequent tax compliance rates for lottery winners and eligible non-winners, or differences in various graded outcomes measured through administrative or survey data.

Binary outcomes (e.g. tax compliance): There are N units with n_T units assigned to treatment and $n_C = N - n_T$ to control; we assume equal numbers assigned to treatment and control ($n_C = n_T$), as in our natural experiment per Table 3.1 and as in most of of the pairwise comparisons in Table 3.2 for the field experiment. We suppose average tax compliance is around 70%, thus the variance of this binary outcome is $0.7 \times (1 - 0.7)$, pooling across treatment and control groups. Thus, the standard error for the difference of tax compliance rates across treatment and control groups is[24]

$$\sqrt{\frac{0.7 \times 0.3}{n_T} + \frac{0.7 \times 0.3}{n_C}}, \tag{5}$$

or, using $n_T = n_C = \frac{N}{2}$,

$$\text{SE} = \frac{2\sqrt{0.7 \times 0.3}}{\sqrt{N}}. \tag{6}$$

For each effect size, we calculate power under a two-tailed test as

$$1 - \Phi(2 - \frac{\textbf{effect}}{\text{SE}}), \tag{7}$$

where Φ is the normal cumulative distribution function, SE is given by equation (6), and **effect** is the true effect size.[25] Equation (7) gives the approximate area above the normal curve centered over **effect** that is more than two standard errors away from 0, the effect size under the null hypothesis.[26] For a one-tailed test, we use 1.65 in place of 2 in equation 7; a one-tailed test is more appropriate for many of our unidirectional hypotheses discussed in Section 5 (e.g., knowledge of the lottery increases

[23]Lucio Castro and Carlos Scartascini. 2013. "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment." Manuscript. Abstract at http://www.cscartascini.org/work-in-progress.

[24]We use the "conservative" formula for the standard error in randomized experiments, which is the same as for the difference of proportions of two independent samples; for formal justification, see Appendix notes 31, 33 of David Freedman, Roger Pisani, and Roger Purves, 2007, *Statistics*, W.W. Norton Co., 4th edition.

[25]We switch the signs in (7) to give the area *greater* than two standard errors above zero.

[26]To be conservative, here we use 2 in place of 1.96, though we can rely on the central limit theorems and use normal approximations for most hypothesis tests; with smaller n, one might want to use the t-distribution or permutation tests.

tax compliance, among lottery losers or bad taxpayers, but does not decrease it).

Figures 9 and 10 show power for two-tailed and one-tailed tests, respectively, assuming true effects of 4, 6, 8, and 10 percentage points, e.g., **effect** $\in \{0.04, 0.06, 0.08, 0.10\}$. In each figure, the vertical line shows the study group size, pooled across treatment and control groups, that is needed for 80% power given each effect size. For $N = 2,000$, we have slightly more than 80% power given a true effect size of 6 percentage points, using a two-tailed test; for a one-tailed test, we have 80% power against a true effect size of 5 percentage points ($N = 2,000$). With a one-tailed test, we also have 80% power for an effect size of 6 percentage points when $N = 1,500$. These calculations suggest reasonable power to measure moderate effects with binary outcomes, using our survey data. However, to measure the binary outcome of tax compliance, we will use cheaper administrative data and thus a larger N, so our power will be substantially greater.

Graded outcomes: Power is greater with graded measures rather than binary outcomes. Our household survey will measure attitudes towards the tax system, often using scales instead of binary outcomes (e.g., degree of agreement with statements about the fairness of the tax system); and we also construct graded measures of indebtedness, as discussed in Section 4. In Figures 11 and 12, we measure effect sizes in relation to the unknown standard deviation of this outcome variable. Thus, for pooled $N = 2,000$, we will have power of just over 80% against a true effect size of 0.13 standard deviations (two-tailed test).

Estimating control-group parameters: One important role of the household survey is to allow us to estimate the proportion of taxpayers in Montevideo who are uninformed about the existence of the lottery. It is critical that we estimate this proportion precisely, as this estimated proportion is the denominator in some of our instrumental-variables analyses. We have several sources of data to estimate the proportion of uninformed taxpayers: (1) survey data on the placebo control group in our field experiment ($N = 2,000$, pooling across eligibles and ineligibles); (2) survey data on the "individual sanction" group in our field experiment, who we also do not inform of the existence of the lottery ($N = 2,000$, pooling across eligibles and ineligibles); and (3) survey data on the control group in the natural experiment ($N = 1,000$).[27]. If 50% of good taxpayers who have never won the lottery

[27]The latter group is a random sample of the population of eligible taxpayers, but eligibility is as of the date of different lotteries, so this is a stratified random sample where we do not know the probability of being in each strata; we may therefore opt to use only (1) and (2) to estimate the proportion of uninformed taxpayers.

are unaware of its existence, the standard error for our estimate of this population percentage is 0.79% pooling across good and bad taxpayers and using data from (1) and (2) ($N = 4,000$). (We will need to weight the estimates when pooling, to account for our oversampling of ineligible taxpayers). With $N = 5,000$, the standard error is 0.71%.

Justification for sample size. These power calculations justify our sample size for the household surveys, as depicted in Tables 3.2 and 3.1. Our power is about 80% against effect sizes for tax compliance comparable to those estimated in previous research, in the case of negative incentives. Our sample size gives us similar power against movements of around 0.15 standard deviations in attitudinal dependent variables measured as scales. Finally, our sample of households who have not won the lottery allows us to estimate the proportion of taxpayers who are uninformed about the lottery with fairly good precision; these estimates are important for assessing overall program impact as well as the likely effects of more effectively promoting knowledge of the tax rebate lottery.

With respect to the administrative data collection, our sample size balances our desire for more data against the cost in time and effort to the municipality.[28] One issue is that some of the tax payer records are for juridical not physical persons (i.e., they are companies). In our surveys, we have the ability to filter juridical persons, ultimately by visiting households; this will allow us to assess the overall proportion of physical persons in the population of taxpayers. However, this will certainly diminish our true power, relative to these calculations. We therefore want to err as much as possible on the side of a large sample size for the administrative data.

4 Outcome Measures

4.1 Administrative (tax payment) data

The policy we study originated in 2004, creating a rich time series of tax payment data for lottery winners and non-winners. We have permission from the municipal government to access historical records for our sampled taxpayers from 2000 to 2014. In addition, we will estimate treatment effects in our field experiment using data on tax compliance posterior to the mailing of our informational

[28]Supplying data is not costless for the municipality, as will involve manual extraction of records from municipal databases using the four-digit IDs we generate, and the municipality will only grant access to a sample of the data.

treatments. Our study period will continue until at least December 2015 to allow us to assess the persistence of effects of our informational interventions using administrative records.

The administrative records allow us to define three main measures of tax compliance as outcome measures:[29]

1. *Compliance (0-1)*: This is a dichotomous indicator for whether a given taxpayer account is fully paid as of the due date. We measure this outcome at each payment date, i.e., three times a year for property and vehicle tax and six times a year for the sewage and head tax.

2. *Missed Payments*: This variable measures the total number of missed payments as of a given tax due date. It varies among taxpayers who are not fully paid up and can increase or decrease during each tax payment period.

3. *Total Debt*: This variable measures the total amount of debt, including unpaid principal, interest, and charges, at each tax payment due date.[30]

In addition, we will collaborate with the municipal government to obtain an additional outcome measure. Our informational treatments (including the placebo control) prompt tax payers to log on to the municipality's website to obtain duplicate copies of their tax bills if needed. We will use the web log and records of the municipality to define the following measure:

1. *Web bill request*: This measures whether a taxpayer logged on to view his or her tax account, change information such as the mailing address associated with the account, or print a duplicate copy of the bill.

The municipality will provide us only with the date or dates at which account information was accessed in a given time period (not the content of the activity).[31] Accessing the web site can be viewed as a measure of *intended tax compliance*, or at least interest in tax records, and it is therefore an interesting outcome variable for a number of our informational treatments.

[29]We may be able to define additional outcome measures prior to analysis, in which case we will amend the pre-analysis plan.

[30]Due to complications involved in calculating historical debts (e.g., the interest and charges that applied at a date in the past), this measure will be only available for our field experiment and only in the period immediately following our intervention.

[31]Web access requires entering a tax account number; thus, the municipality can track date of access by account number.

4.2 Survey data

Our survey instrument gathers data on individual covariates and the main attitudinal dependent variables in our analysis: for example, perceived equity and fairness of the tax system; perceptions of the benefits provided by taxation; and a host of attitudinal (e.g., belief in individual merit) and political (e.g. support for the incumbent Frente Amplio) variables that may be affected by winning the lottery or by knowledge of the lottery. These data on beliefs and perceptions will help us assess mechanisms that may explain any effects we estimate, as discussed in the proposal. The survey instrument will also measure knowledge of the rebate lottery, whether respondents know anyone who has won the lottery, and related variables, which will allow us to assess the likely effects of advertising the tax rebate program more widely.

Here we register eight main outcomes:

1. Trust in municipal government (question C.4.1 in our survey instruments);[32]
2. Trust in civil servants (question C.4.3);
3. Evaluation of the mayor (question C.4.4) (and/or Performance in Office of Mayor, C.9);
4. Fairness of municipal taxes in general (question D. 6); and
5. Fairness of the property tax (question D.7.1).
6. Attitudes towards tax amnesties (question D.8)
7. Agreement that sometimes taxes are not worth paying (question D.9)
8. Party vote intention (question E.5), or change in vote intention from previous election.

Most of these survey questions are all measured on a 0-10 scale. In our instrument, there are a host of other outcomes related to these items as well as to the performance of government and quality of public services. We do not register those secondary outcomes but will certainly explore them descriptively.

4.2.1 Survey experiments

We also will evaluate several outcomes in connection with our survey experiments. First, for the survey experiment about fines and charges, we register the following outcomes, where respondents are asked

[32]Note that the order of questions varies across versions of the questionnaire, due to our survey experiments, but the question identifier/number does not.

for their degree of agreement on a 0-10 scale with the following statements:

1. "People only pay their taxes on time when there are substantial fines and charges" (survey question M.1.1);
2. "In Montevideo, punishments don't apply to the privileged" (question M.1.4); and
3. "Fines and charges for bad taxpayers are pointless" (question M.1.5).[33]

Next, for the survey experiment about the benefit of tax holidays, we register these outcomes (again, as survey questions that ask for degree of agreement on a 0-10 scale):

1. "Policies that reward good taxpayers are a waste of money" (question S.1.1);
2. "In Montevideo, benefits for good taxpayers always go to the same people (question S.1.4)."[34]

Finally, we also asked outcome questions that are identical for both the "fines and charges" treatments and the "benefit of tax holidays" treatments. As for other outcome measures, we will compare the effects of variation in the treatments to assess effects. However, these questions, which are repeated in the corresponding sections of the survey instrument, will also allow us to use the survey experiment to compare the effects of perceptions of negative vs. positive incentives directly:

1. "In general, the municipal government does a good job" (questions M.1.3 and S.1.2);
2. "In Montevideo, it is worth it to be up to date on ones taxes" (question M.1.2 and S.1.3)'[35]
3. "How would you classify the taxes that the municipal government charges, in general: very just, fairly just, a little just, or not just at all?" (questions M.1.6 and S.1.5)[36]

5 Hypotheses and Tests

In this section, we discuss general hypotheses derived from our theoretical discussion and describe our operationalization of these hypotheses, measurement of key outcome variables, and key statistical tests. For a number of our hypotheses, we also describe "mechanisms"—which we understand here as

[33]"*Las multas y recargos a malos paradores no sirven para nada.*"
[34]"*En Montevideo, los beneficios para buenos pagadores se los llevan los mismos de siempre*."
[35]"*En Montevideo, vale la pena estar al día con los impuestos.*"
[36]"*Cómo clasificara los impuestos que cobra la Intendencia de Montevideo en general: muy justos, bastante justos, poco justos o nada justos*?"

intermediate outcomes that could be shaped by our natural, field, and survey experimental interventions and that may help explain any broader program impacts we identify. The impact of our treatments on these intermediate outcomes is therefore of interest for shedding light on mechanisms. (Note that we do not intend to do formal mediation analysis, as the assumptions needed for mediation methods to identify causal effects are too demanding in our context (as in many).[37] However, we use variation-in-treatments in our field experiment design to shed light on reasons any effects of winning the tax holiday lottery, as well as the broader impact of the lottery policy).

Tables 7.3 and 7.4 list the data source and outcomes that we use to test each hypothesis, as well as the specific operationalization of each test; we discuss each hypothesis and test in detail next.

5.1 Impact of the tax holiday lottery

Hypothesis 1A: Winning the tax holiday lottery leads to an *increase* in future tax compliance.
Hypothesis 1B: Winning the lottery leads to a *decrease* in future tax compliance.
Hypothesis 1C: Winning the lottery leads to *no change* in future tax compliance.

> **Operationalization of 1A, 1B, and 1C–natural experiment**: Comparison of compliance and missed payments after the tax holiday lottery takes place. We conduct this analysis over the entire period for which data are available after time t at which each respective lottery is held (thus, outcomes are measured at times $t + \ldots$, as indicated in the first row of Table 7.3). We also compare total debt of winners and eligible non-winners as of the end of July 2014 (the only date at which we measure total debt for reasons discussed previously).
>
> **Test 1: K-S test** We construct a plot in which the horizontal axis is time, measured in tax holiday lotteries (which occur every two or four months, depending on the tax). The vertical axis is one of our three measures of tax payment (compliance, missed payments, or total debt). The date at which each taxpayer in the study group won the tax holiday lottery (treatment group) or was eligible to win that lottery (control group) is centered at 0 on the horizontal axis. We then us a K-S test for the equality of distributions. The treatment and control groups should be balanced before 0, due to the randomization provided by the

[37]On mediation, see e.g. Gerber and Green 2012.

lottery. Differences between the distributions after 0 indicates a treatment effect.

Test 2: Diff-in-diff. We calculate the average value of payment at $t+x$ in the treatment and control groups, and subtract the average value at $t-1$ (or $t-0$) for each group. Then, we compare the change in average outcomes in the treatment (winners) and control (eligible non-winners) groups; this is a difference-in-differences analysis. Standard errors are calculated using the conservative variance formula for the difference-in-difference.

Test 3. Persistence effects. We vary x in the calculation above to estimate the persistence of effects, and compare effects for x less than the average value observed in our data set to x greater than the average value.

There are two mechanisms associated with Hypothesis 1A, a positive effect of winning a tax holiday.

Mechanism 1A.1: *Informational*. The municipal government appears to advertise the existence of the lottery quite poorly, and winning the lottery provides information about its existence (as well as a year free of tax payments). Thus, winning the lottery provides taxpayers with information that they have a positive probability of winning in the future if they pay taxes promptly.

Operationalization 1: Field experiment. Comparison of tax compliance and web bill requests of households that receive our lottery/individual reward treatments 1 and 2 (pooled together) and the control group, separately among eligible and ineligible households. A positive effect for eligible households suggests that part of the impact of the lottery on future compliance is due to the information that winning provides. A positive effect for ineligible households suggests larger program impact, as knowledge of the lottery causes bad taxpayers to pay on time (become good taxpayers).

Test 1: Diff.-in-Diff. We compare the change in outcomes for the treatment group (whom we inform of the existence of the lottery) and the placebo control group, using two outcome measures: compliance (0-1) and whether taxpayers viewed their accounts online (web bill request).[38] For the treatment condition, here we pool the three groups that are informed of the lottery: Individual reward 1; Individual reward 2; and Social reward. (Elsewhere,

[38]We will need to have pre-intervention data for the web bill request to do the diff-in-diff.; the municipality has told us it can provide historical data on web requests.

we distinguish effects for these groups). This is intent-to-treat analysis, because we do not take into account whether tax payers already know about the existence of the lottery prior to receiving our informational treatment. We measure the post-intervention outcome at $t+1$—i.e., the next possible tax payment, which occurs in July 2014—as well as $t+4$, the payment that occurs in July 2015. The reason for the latter is that we will next compare effects in the field experiment to effects in the natural experiment.[39] In the natural experiment, the first recorded outcome for lottery winners occurs at $t+4$, since winners have a year free of tax payments and thus cannot opt to comply or not comply with taxes until $t+4$.[40] We do this analysis for both eligible and ineligible taxpayers.

Test 2: Instrumental-variables analysis (IV). Intent-to-treat analysis estimates the effect of providing citizens with information, regardless of whether taxpayers already know about the existence of the lottery. Yet, not all non-winners are ex-ante uninformed. Thus, here we also use instrumental-variables analysis. Dividing the estimated average treatment effect of information (in the difference-in-differences analysis) by the difference in information rates of bad taxpayers in treatment and control groups provides an instrumental-variables estimate of the effect of information on "Compliers"—those who learn about the lottery from our intervention.[41] We estimate the denominator—the proportion of uninformed good taxpayers—using our survey data for good taxpayers in the field experiment's placebo control and Individual Sanction conditions, since these treatments do not inform taxpayers about the lottery."[42] This analysis is important for policy, because it identifies the marginal impact among taxpayers of an informational campaign to advertise the policy

[39]Note that strictly speaking, the effects are only estimated for the same population with the small natural experiment occurring in July 2014, i.e., winners compared to non-winners who are eligible to win as of that date. However, we might think of estimating effects for "eligible taxpayers" in general, including those who were eligible at earlier points in time.

[40]A subtlety here is that $t+1$ represents the first opportunity for taxpayers in the field experiment to pay their taxes, while $t+4$ is the first opportunity in the natural experiment; thus, in the latter, more time passes between intervention and outcome. However, we have confirmed that winners of the tax holiday receive bills every three months with a zero balance during the year that they are exonerated from paying taxes. This makes the time from intervention to outcome more equivalent, if one thinks of receiving the final notice of zero tax due as a part of the treatment.

[41]Recall from section **??** that some of our flyers were mailed to juridical persons (e.g. corporations), rather than physical persons. The denominator of the IV estimator must take this into account, i.e., the proportion of informed or uninformed households is constructed taking into account the entire study group including juridical persons.

[42]As discussed previously, we could also use the control group in the natural experiment, but that introduces some complications we mentioned.

more widely.[43] Again, we do this analysis for both eligible and ineligible taxpayers.

Operationalization 2: We compare the effect on future compliance of information in the field experiment—among eligible households—and the effect of winning the lottery, in the natural experiment. If effects of winning are entirely informational, these two effects should be statistically indistinguishable.

Test 1. We compare the estimated effect of the "existence of lottery" treatments in the field experiment (relative to the placebo control) to the estimated effect of winning the lottery in the natural experiment, using data at $t + 4$ in both cases, and test whether the estimated effects are significantly different.

Operationalization 3: We compare respondents whom we inform through the survey experiment about the existence of the lottery (individual and social benefit) with those not informed about lottery.[44]

Test 1. We compare means on the survey question measuring whether respondents say it is "worth it" to pay taxes.

Mechanism 1A.2: *Attitudinal*. The second mechanism works through taxpayers' faith in the equity and transparency of the tax system, which may boost their willingness to pay taxes (i.e., provide an expressive benefit b as in our decision model). As noted in the introduction, allocating public benefits (such as tax holidays) through lotteries represents a distinctive form of *programmatic politics*—one in which binding, public criteria guide the distribution of benefits (and benefits are not conditional on political support of beneficiaries).[45] This form of distributive politics is particularly egalitarian, because all eligible taxpayers have the same probability of winning benefits.

Beyond tax compliance, allocating public benefits (such as tax holidays) through lotteries represents a form of distributive politics that is quite distinct from the conditionalities of clientelism or patronage politics. For political incumbents, rewarding good taxpayers in an equitable and transparent

[43]Note that IV analysis estimates effects for a particular type: uninformed bad taxpayers who would learn about the lottery from an informational campaign. Yet, such a campaign should be targeted at exactly this population.

[44]Note that in designing the study, we did not originally conceive of the discretion treatment as a pure control group in this sense. But the information about "discretion" is fairly light (though the municipality is said to "choose" winners), so this may work.

[45]For this definition of programmatic politics, see Stokes et al. (2013).

way may also have important electoral consequences, particularly in a context in which amnesties for *bad* taxpayers are common and potentially unpopular. We treat the impacts of the lottery on attitudes towards the transparency and fairness of the tax system as potential mechanisms that may help explain any impact of the lottery on tax compliance. Yet, we are also interested in the attitudinal impact of the tax holiday policy on political perceptions, independent of any consequences for tax payment behavior. In addition, for political incumbents, rewarding good taxpayers in an equitable and transparent way may also have important electoral consequences, particularly in a context in which amnesties for delinquent taxpayers are common.

Note that these are not hypotheses about the impact of winning a *particular* lottery, nor are they hypotheses about being informed that one might win the lottery; rather, they are hypotheses about how the *existence of the policy* shapes attitudes and behavior. Thus, while we conceptualize these attitudinal effects as *mechanisms* that might explain the impact of the lottery policy on tax compliance behavior, effects of the policy on political attitudes and support for incumbents is interesting regardless of whether this in turn induces greater tax compliance. We therefore characterize the following hypothesis about the impact of the tax holiday policy on perceptions of the state and the character of public policy: the tax holiday lottery increases perceived transparency and fairness of the tax system as well as political support for incumbents responsible for the policy.

Operationalization: We assess whether being informed about the lottery and/or winning the lottery boosts 1. trust in the municipality; 2. trust in civil servants;[46] 3. evaluations of the mayor; 4. perceived fairness of municipal taxes in general; 5. perceived fairness of property taxes (field experiment—for the natural experiment, we use the specific tax for which the taxpayer had eligibility); and 6. degree of agreement with amnesties for bad taxpayers. Regarding the 6, note that good taxpayers who are current on their tax payments may resent amnesties for bad taxpayers; a tax holiday lottery that rewards good taxpayers may therefore influence their perceptions of the tax system as well as of political incumbents. It may also cause them to look more favorably on tax amnesties.

Test 1. For both the field and natural experiment, we conduct difference-of-means tests

[46]We first ask about the municipal union, then other municipal civil servants; we use the latter question.

comparing the treatment and control groups.

Test 2. We also test for heterogeneous effects—in the case of the field experiment, comparing eligibles to ineligibles, and in the case of the natural experiment, comparing recent to old winners.

Test 3. We also asses whether the effect of *winning* the lottery on attitudes towards the fairness of the tax system and support for incumbents will be [greater than/equal to] the effect of knowledge of the lottery for eligible taxpayers. Taxpayers who have actually won the lottery have received a concrete benefit (in the form of a year of exoneration from taxes), whereas those we inform about the existence of the lottery have merely received information about the policy. Thus, if the effects on attitudes towards the state and support for political incumbents for both groups are similar (and non-zero), it suggests that the policy mainly works by shaping attitudes about the desirability of this mode of distribution.

Regarding Hypothesis 1B, at least two mechanisms might explain a negative effect of winning the lottery on future compliance:

Mechanism 1B.1: *Income effects*. Winning the lottery gives taxpayers a year free of paying taxes; in principal, this additional income could buttress them against the costs of punishment in case of non-payment for future taxes.

Operationalization: We assess whether the effects of winning the natural experiment–which provides a year free of income–varies according to the cost of payment, here operationalized in terms of the property value.

Test: Heterogeneous effect analysis—difference of means.

Mechanism 1B.2: *Behavioral/"habit" effects*. Taxpayers who stop paying taxes for a year may fall "out of the habit," leading them to miss making tax payments once they are due.

Operationalization: We assess whether the effects of winning the lottery vary by time since winning.

Test: Natural experiment—heterogeneous effects by time since winning.

Operationalization: The behavioral effect should not exist for taxpayers who are set up to pay taxes by automatic debit (placebo outcome) **Test**: Field and natural experiment—assess effect for those who self-report using automatic debits to pay taxes.

Finally, Hypothesis 1C—a null effect of winning the lottery—could follow, for instance, if the lottery has only informational effects on future compliance, but good taxpayers are fully informed about the lottery before winning. In this case, winning the lottery would not inform them about a possible future benefit of staying current on their tax payments. Moreover:

Mechanism 1C.1: *Erroneous beliefs*. Taxpayers who have won the lottery may believe, falsely, that their probability of winning a future lottery is lowered by having won previously. This could either lead to a decrease in tax compliance among winners, relative to eligible non-winners, or to a null effect of winning the lottery. In particular, taxpayers who already knew about the existence of the lottery before winning it might comply more before winning the lottery than after doing so.[47]

Operationalization: Using a survey question, we assess whether taxpayers believe that someone who has won the lottery has less of a chance of winning it in the future.[48] Then, we compare the effects of winning according to beliefs in the dependence of winning on previous outcomes.[49]

Test 1: Heterogeneous effects of winning the lottery, by beliefs about independence of lotteries: compare those who say a winner is less likely to win again, to those who say the chances of winning a second lottery are equal or greater for someone who has already won.

[47]That is, some taxpayers who were previously informed about the lottery may have increased compliance in order to boost the probability of winning; but having won, they may believe erroneously that their chances of winning are reduces, providing a disincentive to continued compliance.

[48]Survey question S.5: If a person wins [the tax holiday] lottery, would you say that the chances of winning again on another occasion are: greater, the same, or less?

[49]Note, however, that beliefs in the dependence of winning are post treatment.

A final possibility under H1.C is that only expressive benefits (e.g. a sense of duty) matter for payment of taxes—and winning the lottery has no effect on these perceived expressive benefits of compliance.

5.2 Rewards vs. punishments.

Hypothesis 2A: (Priming knowledge of) sanctions and punishments for non-compliance *increases* tax compliance.

> **Operationalization**: We compare payments at $t + 1$ (July 2014) of taxpayers who are reminded/informed of fines and punishments for non-payment to the placebo control group.[50]
>
> **Test 1**: Here, we pool the "individual sanction" and "social sanction" treatments, as both treatments remind taxpayers of the existence of fines and punishments for non-payment and we are using administrative outcome data which are available for both groups. (Later, we distinguish between the effects of these treatments). The test statistic is the difference of means, divided by the estimated standard error; as elsewhere, we use the conservative Neyman estimator of the variance (and standard error). Since we have a directional hypothesis (one-tailed test), we reject the null at the nominal 0.05 level if this statistic is greater than 1.65 (though see below for adjustments for multiple comparisons).

We also assess one possible mechanism:

Mechanism 2A.1: *Beliefs about probability of punishment*: the sanction treatments increase beliefs in the probability of punishment. Also, the positive effect of reminders is stronger for those who believe the probability of punishment is greater.

> **Operationalization**: First, we assess whether the sanctions treatment decreases the proportion of taxpayers who say it is "very probable" or "somewhat probable" that the municipal government offers an amnesty to bad taxpayers in the coming year, relative to those who say it is "a little probable, or very improbable." Next, we hypothesize that the effects of the sanction treatment will be weaker for those who say it is very probable or somewhat probable (weak belief in the probability of punishment) to those who say it is a little probable, or very improbable (strong belief in the probability of punishment).
>
> **Test**: First, we use a difference-in-means test to assess whether the sanctions treatment decreases belief in the probability of an amnesty (one-tailed test). Then, we text whether the estimated effects of the sanction treatments are higher for the group that believes the probability of an amnesty is low, using a one-tailed test. Note that beliefs in the probability of an amnesty are plausibly post-treatment for taxpayers in the treatment (sanction) group; so this is not heterogeneous effects analysis but rather a type of mediation analysis. Note that the assumptions needed for to identify this mediation analysis are strong.

[50]Unfortunately, we do not have a survey question that asks individuals about the existence of fines and sanctions, which would allow us to estimate the proportion who are informed about these.

We are especially interested in comparing the size of the effects of manipulating negative and positive incentives to pay taxes, though we are agnostic about the direction of the effect:

Hypothesis 2B: The absolute value of the effect of priming knowledge of sanctions and punishments is [greater/less than] the effect of priming knowledge of the tax holiday lottery.

> **Operationalization**: Here, we are interested in the effects on tax payment at $t + 1$ (July 2014) of the rewards treatments vs. the sanctions treatments. Here, we pool the "individual reward 1," "individual reward 2" and "social reward" as the "rewards treatments" and "individual sanction" and "social sanctions" as the "sanctions treatment."
>
> **Test 1**: We compare the estimated effects in a difference-in-difference analysis. The estimated standard error for the diff-in-diff is the square root of the sum of the estimated variances of the two estimated effects. We use a two-tailed test.

We also hypothesize that the effects of manipulating both positive and negative incentives to pay taxes are heterogeneous, depending on taxpayers' compliance history.

Hypothesis 2C: *Marginal taxpayers I*.The effects of both rewards and sanctions are greater for marginal taxpayers (those with only some history of non-compliance) than for those with no history of late payments or those with an extensive history of late payments.

The rationale here is that the cost of becoming a good taxpayer is greatest for those who are seriously in arrears; they are less likely to be moved by the promise of a one-year tax holiday, conditional on bringing all their accounts up to date.[51] Taxpayers with an extensive history of late payments may also be less responsive to the threat of sanctions.

On the other side of the spectrum, taxpayers with no history of late payments may always pay on time, due to intrinsic motivations (a large expressive benefit b to tax payment); for such taxpayers, paying taxes on time is part of their social identity, and their taxpaying behavior is unlikely to be shaped by the field or natural experimental interventions.

Thus, those taxpayers most likely to be affected are those on the margins: either taxpayers who are currently in arrears, but not very much, or those who are currently eligible for tax holiday lotteries, but who have been in arrears in the past.

> **Operationalization**: We operationalize the idea of marginal taxpayers by identifying "good taxpayers" with some history of past non-payment,[52] and "bad taxpayers" (current ineligibles) who do not owe more than 3 quotas of past tax payments.[53]
>
> **Test 1**: Among good taxpayers, we compare the effect of information about the existence of the lottery, and the effect of information about fines and sanctions, across those good taxpayers who have never owed a debt and those who have at some point in the past been past due on their taxes.[54] Among bad/ineligible taxpayers, we compare the effects for those with 3 or fewer quotas owed to the effects for worse taxpayers. In both cases we use one-tailed tests.

The idea that marginal taxpayers will be most responsive to manipulating perceived benefits and costs suggests a final hypothesis as well:

Hypothesis 2D: *Marginal taxpayers II*. The effects of both rewards and sanctions are greater for taxpayers who face lower costs of coming into compliance.

[51] We will use our sample of taxpayer records to characterize how common these very delinquent taxpayers are.
[52] We will seek to identify these taxpayers using historical payment data on the field experimental study group.
[53] We choose the threshold of three payments because the municipality allows taxpayers who owe no more than three quotas to continue to pay their current tax bill without accumulating interest or charges on the 3 quotas that are past due.
[54] As noted previously, we have screened out of our study group taxpayers who pay the full year's taxes at once and thus do not pay in July. These taxpayers tend overwhelmingly to be good taxpayers. Thus we are already looking at relatively marginal taxpayers in our study group.

> **Operationalization**: Taxpayers with high incomes but low property values (and therefore taxes) face lesser costs in tax payments in general and should be less responsive to manipulation of costs and benefits.[55] On the other side of the spectrum, taxpayers with low incomes and high property values and thus taxes should face greater costs of compliance and should thus also be less responsive to treatments.[56]. Those taxpayers who are on the margin—those intermediate ratios of incomes to taxes owed—should be most responsive to our field and natural experimental treatments. We therefore operationalize the cost of coming into compliance in terms of ratios of income to property values, and we hypothesize that the effects are weaker for taxpayers for whom these ratios are extreme.
>
> **Test 1.** We compare the effects on compliance for taxpayers within one standard deviation of the average value of income-to-property-value ratio[57] with the effects for other taxpayers.

5.3 Individual vs. social incentives.

Our informational interventions will allow us to compare the costs and benefits of a campaign to increase awareness of the tax holiday program, as well as the effects of reminding taxpayers of sanctions for non-payment. Yet, an important question is whether emphasizing the social benefits and costs makes such a campaign more effective. In terms of our theory sketched above, the question is whether manipulating the value of the expressive benefit b can boost the impact on tax compliance. Thus, we have the following hypotheses:

Hypothesis 3A: (*Social benefits*). Emphasizing social rationale for *benefits* provided through the tax holiday lottery increases compliance, relative to an emphasis on individual returns.

> **Operationalization**: We compare the effect of the social benefit treatment to the effect of the individual benefit treatment (or alternately, simply assess whether the mean response for the social benefit treatment is greater/different than the mean response for the individual benefit treatment).

[55] An observational corollary is that these taxpayers should tend to be current on their accounts.
[56] These taxpayers should tend to be in arrears more often
[57] Alternately, those from the 25th to 75th percentile on this measure?

Test 1: Difference-of-means; two-tailed test (since in principle social rational could also *decrease* tax compliance, relative to the individual reward. Here, for the individual reward, we pool individual reward 1 and individual reward 2.

Hypothesis 3B: (*Social sanctions*). Emphasizing the social rationale of *sanctions* for non-payment increases compliance, relative to an emphasis on individual punishment.

Operationalization: We compare the effect of the social sanction treatment to the effect of the individual sanction treatment (or alternately, simply assess whether the mean response for the social sanction treatment is greater/different than the mean response for the individual sanction treatment).

Test 1: Difference-of-means; two-tailed test (since social rationale for sanctions could also *decrease* tax compliance, relative to a focus on individual sanctions.

If both hypotheses 3A and 3B were born out in the data, it would provide evidence that social preferences—operationalized in terms of manipulating the expressive benefit of paying taxes, or social cost of not paying taxes—play an especially important role in shaping compliance behavior. Such a finding could inform the design of informational interventions by the municipality of Montevideo or similar governments.

5.4 Adjustments for multiple comparisons

We will present both nominal p-values and corrected p-values, using a false discovery rate (FDR) correction to control the Type-1 error rate. For both our natural and field experiments, we will control the FDR at level 0.05.[58]

Thus, for a given randomization with m (null) hypotheses and m associated p-values, we order the realized nominal p-values from smallest to largest, $p_{(1)}, \leq p_{(2)} \leq \ldots \leq p_{(m)}$. Let

$$k \text{ be the largest } i \text{ for which } p_{(i)} \leq \frac{i}{m} 0.05.$$

[58]The problem of multiple comparisons also arises with our survey experiments, as we have multiple outcomes for each experiment. We will take a similar approach to adjustment in that case.

Then, we reject all $H_{(i)}$ for $i = 1, 2, \ldots, k$, where $H_{(i)}$ is the null hypothesis corresponding to $p_{(i)}$.[59]

For comparison, we will also present strict Bonferroni corrections, i.e., for each hypothesis $H_{(i)}$, we reject the null at the adjusted 0.05 level if $p_{(i)} \leq \frac{0.05}{m}$. This correction will lead to the most conservative inference for each individual pairwise comparison. Our rejection rule, however, will require controlling the overall false discovery rate.

How large is m under our study design? This differs for the field, natural, and survey experiments. For the natural experiment, we have one randomization into treatment (winner) and control (eligible non-winner) groups. (As discussed, this is really blocked randomization, where the blocking is by type of lottery; however, the blocks are all in expectation the same size for each type of tax). Meanwhile, we will have nominal p-values associated with each of the following comparisons:

- K-S test (three outcomes: compliance, missed payments, and total debt)
- Diff-in-diff (three outcomes)
- Persistence of effects, heterogeneous effects (three outcomes)
- Difference of means (six outcomes: trust in municipality, trust in civil servants, evaluation of mayor, fairness of taxes, fairness of the tax specific to the corresponding lottery,[60] and opinion of lottery)
- Heterogeneous effects, by cost of payment (three outcomes)
- Heterogeneous effects, by time since winning (three outcomes)
- Heterogeneous effects, by beliefs about non-independence of winnings (three outcomes)

The total number of comparisons is 24. We also have the p-value for the comparison of effects in the natural and field experiment. This makes a total of $m = 25$ p-values for the natural experiment.

For the field experiment, we have the following comparisons:

- Diff-in-diff (two outcomes each for $t + 1$ and $t + 4$).[61]

[59]For a description of this procedure, see Yoav Benjamini and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological).* 57 (1): 289-300.

[60]We will do these separately for the winners/eligible losers for each type of tax, i.e. focus on the fairness-of-property-tax outcome onlyfor those who won or were eligible to win a year free of property taxes.

[61]We do not count the $p-values$ associated with the IV analyses as nominal significance should not differ from the intent-to-treat analysis. An initial report may ignore the $t + 4$ outcomes until we have gathered those data a year later.

- Difference of means (six outcomes: trust in municipality, trust in civil servants, evaluation of mayor, fairness of taxes, fairness of property taxes, and and support for amnesties—the analysis is conducted for both eligibles and ineligibles, so twelve tests total);
- Effect of punishments treatment on compliance and belief in probability of punishments (two outcomes), and comparison to effect of reward (one outcome);
- Heterogeneous effects, by payment history (one outcome)
- Effect of social benefits (one outcome), effect of social sanctions (one outcome), comparison of effects (one outcome)

The total number of comparisons is 23. We also include the p-value for the comparison of effects in the natural and field experiment. This makes a total of $m = 24$ p-values for the field experiment.

6 Relevance, Contribution, and Value of Research

Promoting tax compliance is critical in developing countries, where tax monitoring and enforcement is often weak. We believe have a valuable opportunity to use a randomized policy intervention to learn about the effects of positive as well as negative incentives for compliance, as well as the separate impact of priming social versus individual rationale for punishments/benefits. Our informational intervention will allow us to conduct cost-benefit analysis on scaling up promotion of knowledge of the rebate lottery.

7 Data Availability

Data will be posted upon publication of a research report or two years after the start of our informational intervention, whichever is sooner.

Figure 2: Text of informational intervention (Spanish): Placebo control


Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

Por consultas:

FONO TRIBUTOS
1950 3000

Figure 3: Text of informational intervention (Spanish): Lottery/individual reward 1

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

La Intendencia de Montevideo quiere premiar a los buenos pagadores. **Si usted paga en fecha participará automáticamente de un sorteo por la exoneración de un año de Contribución Inmobiliaria.**

Los sorteos se realizan todos los meses pares del año junto con la Lotería Nacional. Los beneficiados serán debidamente informados y se publicarán los resultados en el sitio web de la Intendencia.

¡Usted puede ser el próximo!

Por consultas:

FONO TRIBUTOS
1950 3000

Figure 4: Text of informational intervention (Spanish): Lottery/individual reward 2

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

La Intendencia de Montevideo quiere premiar a los buenos pagadores. **Si usted paga en fecha participará automáticamente de un sorteo por la exoneración de un año de Contribución Inmobiliaria.**

En cada sorteo, 1 de cada 5.000 hogares recibe este beneficio.

Los sorteos se realizan todos los meses pares del año junto con la Lotería Nacional. Los beneficiados serán debidamente informados y se publicarán los resultados en el sitio web de la Intendencia.

¡Usted puede ser el próximo!

Por consultas:

FONO TRIBUTOS
1950 3000

Figure 5: Text of informational intervention (Spanish): Individual sanction

Intendencia
deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

Quienes no paguen en fecha podrían estar sujetos a multas y recargos. La Intendencia de Montevideo podría tomar acciones administrativas y legales para hacer cumplir la normativa en los casos que correspondan.

¡Pague en fecha, evite multas y recargos!

Por consultas:

FONO TRIBUTOS
1950 3000

Figure 6: Text of informational intervention (Spanish): Lottery/Social reward

Intendencia
deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

La Intendencia de Montevideo quiere premiar a los buenos pagadores. **Si usted paga en fecha participará automáticamente de un sorteo por la exoneración de un año de Contribución Inmobiliaria.**

Los sorteos se realizan todos los meses pares del año junto con la Lotería Nacional. Los beneficiados serán debidamente informados y se publicarán los resultados en el sitio web de la Intendencia.

La Intendencia de Montevideo realiza este sorteo para reconocer a los buenos pagadores por su contribución a la construcción de una ciudad más justa y mejor para todos/as.

Por consultas:

FONO TRIBUTOS
1950 3000

Figure 7: Text of informational intervention (Spanish): Social punishment

Intendencia deMontevideo

Estimado/a vecino/a:

Queremos recordarle que en el mes de **julio vence la segunda cuota de la Contribución Inmobiliaria.** Si todavía no recibió su factura, puede obtener un duplicado en nuestro sitio web (www.montevideo.gub.uy).

Quienes no paguen en fecha podrían estar sujetos a multas y recargos. La Intendencia de Montevideo podría tomar acciones administrativas y legales para hacer cumplir la normativa en los casos que corresponda.

Las multas y recargos son una sanción para quienes no pagan sus impuestos y no contribuyen a la construcción de una ciudad más justa y mejor para todos/as.

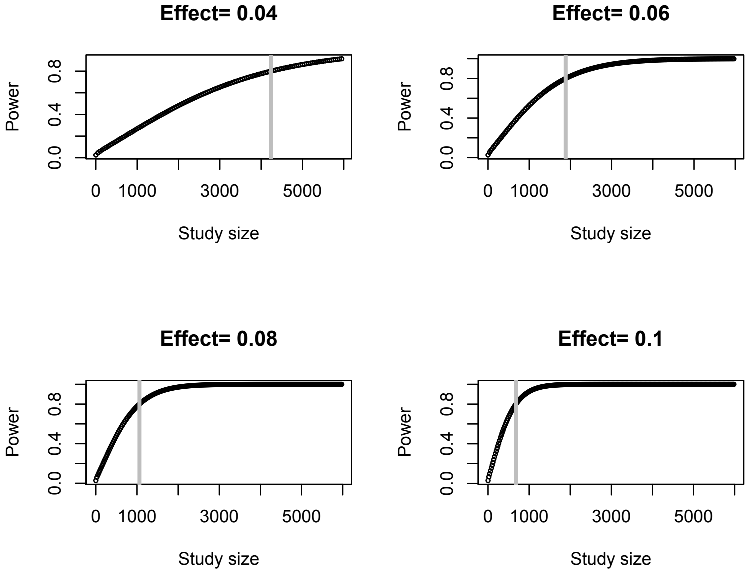
Por consultas:

FONO TRIBUTOS
1950 3000

Figure 8: Informational intervention: Reverse side of flyers with municipal logo

Figure 9

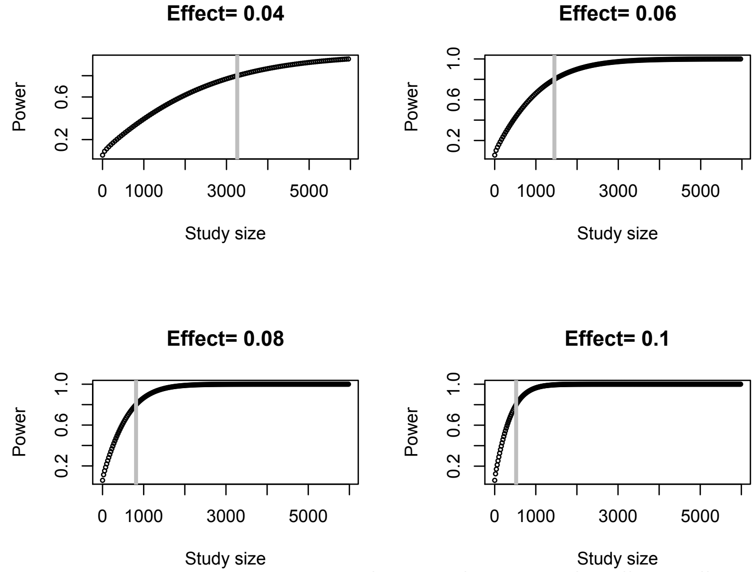
Power Calculations for Different True Effect Sizes
(Two-tailed tests, Binary Outcome)



Plots show statistical power as a function of study size for different effect sizes (binary outcome, e.g tax compliance). Effects are differences of proportions. Vertical line shows the size required for 80% power.

Figure 10

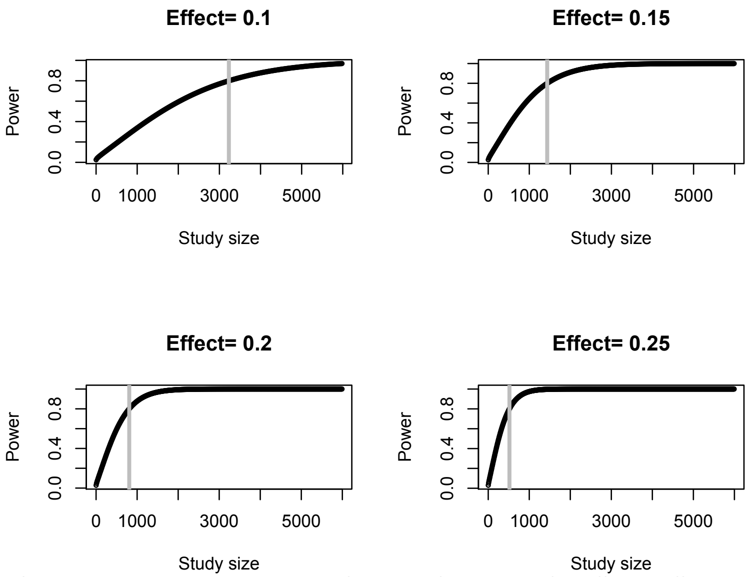
Power Calculations for Different True Effect Sizes
(One-tailed tests, Binary Outcome)



Plots show statistical power as a function of study size for different effect sizes (binary outcome, e.g tax compliance). Effects are differences of proportions. Vertical line shows the size required for 80% power.

Figure 11

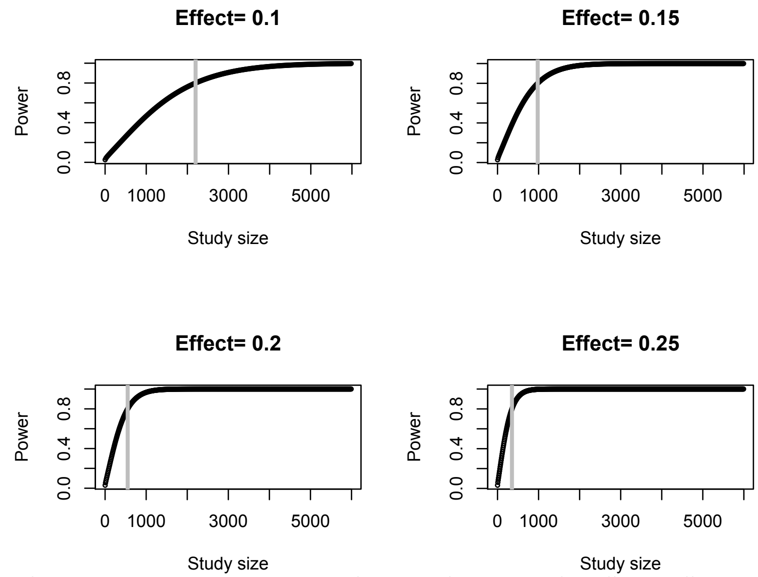
Power Calculations for Different True Effect Sizes
(Two-tailed tests, Graded Outcome)



Circles show statistical power as a function of study size for different effect sizes (graded outcome, e.g attitude scales). Effect sizes are expressed in standard deviations, e.g. 0.1 of one SD. Vertical line shows the size required for 80% power.

Figure 12

Power Calculations for Different True Effect Sizes
(One-tailed tests, Graded Outcome)



Circles show statistical power as a function of study size for different effect sizes (graded outcome, e.g attitude scales). Effect sizes are expressed in standard deviations, e.g. 0.1 of one SD. Vertical line shows the size required for 80% power.

Table 7.3: **Hypotheses, Outcomes, and Tests**

Hypotheses	Data Sources	Outcomes	Comparisons	Tests
Hypotheses 1A, 1B, 1C (Winning lottery)	Natural Exp. (Admin. data)	$t+\ldots$ 1. Compliance (0-1) 2. Missed Payments 3. Total Debt	Winners vs. Non-Winners (Eligibles)	1. K-S test 2. Diff-in-Diff 3. Persistence of effects
Mechanism 1A.1 (*Informational*)	Field Exp. (Admin data) (Eligibles and Ineligibles)	$t+1$ & $t+4$: 1. Compliance (0-1) 2. Web bill request	Existence of Lottery vs. Control	1. Diff.-in-Diff 2. IV
	Field Exp. vs. Nat. Exp. (Admin data)	$t+4$: 1. Compliance (0-1) 2. Web bill request	Effect of Info. vs. Effect of Winning	1. Diff. of Diff.-in-Diffs
	Survey exp. (Survey data)	1. Worth it to pay (Q. S.1.3)	Ex. of Lottery vs. Discretion	1. Diff. of Means
Mechanism 1A.2 (*Attitudinal*)	Field Exp (Survey data) (Eligibles + Ineligibles)	$t+1$ 1. Trust in municipality 2. Trust in civil servants 3. Eval. of Mayor 4. Fairness Taxes 5. Fairness Prop. Taxes 6. Support Amnesties	Existence of Lottery vs. Control	1. Diff. of Means 2 and 3. Diff. of Means (Het. effects)
	Nat. Exp (Survey data)	$t+\ldots$ 1. Trust in municipality 2. Trust in civil servants 3. Eval. of Mayor 4. Fairness Taxes 5. Fairness Spec. Tax 6. Opinion of lottery (Version 1 of survey)	Winners vs. Losers	1.Diff. of Means 2. Diff. of Means Het. effects – recent vs. old winners)
	Survey Exp (Survey data)	t 1. Lotteries are waste of money 3. Eval. of City Hall 4. Benefits go to "same as always"	Lottery treatments vs. non-lottery (discretion) treatment	1.Diff. of Means 2. Diff. of Means Het. effects –
Mechanism 1B.1 (*Income effects*)	Nat. Exp. (Admin. Data)	$t+\ldots$ 1. Compliance 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Heter. effects by cost of payment (property value)

Table 7.4: **Hypotheses, Outcomes, and Tests (Cont.)**

Mechanism 1B.2. (*Habit effects*)	Nat. Exp. (Admin. Data)	$t+\ldots$ 1. Compliance 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Het. effects by time since winning
Mechanism 1C.1. (*Erroneous beliefs*)	Nat. Exp. (Admin/Survey Data)	$t+\ldots$ 1. Compliance (0-1) 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Het. effects by beliefs about non-indep. of winning
Hypothesis 2A (*Punishments*)	Field Exp. (Admin. Data)	$t+1$ 1. Compliance (0-1)	Existence of Fines vs. Control	1. Diff. of Means
	Survey Exp (Survey data)	t 1. Worth it to pay 2. Eval. of City Hall 3. The privileged escape fines	Fines treatment vs. lotteries treatment	1.Diff. of Means
Mechanism 2A.1 (*Prob. punishment*)	Field Exp. (Survey Data)	$t+1$ 1. Belief in Prob. of Fine	Existence of Fines vs. vs. Control	1. Diff. of Means
Hypothesis 2B (*Rewards vs. Punishments*)	Field Exp. (Admin. Data)	$t+1$ 1. Compliance (0-1)	Existence of Fines vs. Control	1. Diff.-in- Diff
Hypothesis 2C (*Marginal taxpayers I*)	Field Exp (Admin. Data)	$t+1$	Benefit vs. Sanction	1. Het. effects by payment history
Hypothesis 2D (*Marginal taxpayers II*)	Field Exp. (Admin./Survey)	$t+1$ 1. Compliance (0-1)	Benefit vs. Sanction	1. Het. effects by payment cost
Hypothesis 3A (*Social benefits*)	Field Exp. (Admin. Data)	$t+1$	Social Ben. vs. Indiv. Ben.	1. Diff. of Means
Hypothesis 3B (*Social sanction*)	Field Exp. (Admin. Data)	$t+1$	Social Sanc. vs. Indiv. Sanc.	1. Diff. of Means
Hypothesis 3A-B (*Social benefits vs. social sanctions*	Field Exp. (Admin. Data)	$t+1$	Effect of vs. Social Sanc. vs. effect of Indiv. Sanc.	1. Diff. of Diff. of Means

In the table, t refers to tax payment periods, of which there are 3-6 per year, depending on the tax. Thus, for winners of the lottery, $t = 0$ is the period in which they won the lottery; $t = 1$ is the following tax payment period; and because they win a year free of tax payments $t + 4$ is the next payment period in which they owe taxes. Property taxes are paid three times per year.

G.2 First amendment to PAP (registered October 19, 2014)

Amended Pre-Analysis Plan. Administrative Data.

Positive vs. Negative Incentives for Compliance: Evaluating a Randomized Tax Holiday

October 19, 2014

Contents

Description of the document and setup of the analysis

This document amends our registered pre-analysis plan and adds the code to be used for the analysis of the administrative data. It will be followed up by an additional document amending and providing mock analysis of the survey data.

This document was produced once we got most of the administrative data, but before conducting any analysis. The results presented here were produced using mock data. We generate our mock data by reshuffling the treatment labels in our datasets without replacement (see below). For a few variables for which we still do not have the data we generate mock variables.

Generating mock data for analysis

Natural Experiment

```r
load("~/Dropbox/Uruguay state capacity/Analysis/data/final_data/naturalex_data.Rda")

# keeping the good taxpayers
naturalex_gtp <- naturalex # dataset has only good taxpayers

treat_by_tax <- ddply(naturalex_gtp, c("tax", "won_lottery"),
                      summarise,
                      N = length(unique(account)))
treat_by_tax
```

Main data

```
##                          tax won_lottery    N
## 1 Contribucion Inmobiliaria           0 1291
## 2 Contribucion Inmobiliaria           1 1275
## 3        Patente de Rodados           0  344
## 4        Patente de Rodados           1  366
```

```
## 5              Saneamiento           0  386
## 6              Saneamiento           1  439
## 7    Tributos Domiciliarios           0  994
## 8    Tributos Domiciliarios           1  963
```

```
# Creating vectors for mock treatment and control respecting the type of tax
set.seed(2067)
vector_CI <- sample(
  c(rep(0,treat_by_tax[1,3]),
    rep(1,treat_by_tax[2,3])), (treat_by_tax[1,3]+treat_by_tax[2,3]), replace=F)
vector_PR <- sample(
  c(rep(0,treat_by_tax[3,3]),
    rep(1,treat_by_tax[4,3])), (treat_by_tax[3,3]+treat_by_tax[4,3]), replace=F)
vector_TS <- sample(
  c(rep(0,treat_by_tax[5,3]),
    rep(1,treat_by_tax[6,3])), (treat_by_tax[5,3]+treat_by_tax[6,3]), replace=F)
vector_TD <- sample(
  c(rep(0,treat_by_tax[7,3]),
    rep(1,treat_by_tax[8,3])), (treat_by_tax[7,3]+treat_by_tax[8,3]), replace=F)

# Binding the treatment vectors with the account number for CI
CI <- cbind(unique(naturalex_gtp$account[naturalex_gtp$tax=="Contribucion Inmobiliaria"]),
            vector_CI)
PR <- cbind(unique(naturalex_gtp$account[naturalex_gtp$tax=="Patente de Rodados"]),
            vector_PR)
TD <- cbind(unique(naturalex_gtp$account[naturalex_gtp$tax=="Tributos Domiciliarios"]),
            vector_TD)
TS <- cbind(unique(naturalex_gtp$account[naturalex_gtp$tax=="Saneamiento"]),
            vector_TS)

mock_assignment <- as.data.frame(rbind(CI, PR, TD, TS))
names(mock_assignment) <- c("account", "won_lottery")

# eliminating real treatment variable from the dataset
check <- sum(naturalex_gtp$won_lottery)
naturalex_gtp <- naturalex_gtp[,-which(names(naturalex_gtp)=="won_lottery")]

# Merging to add mock assignment
naturalex_gtp <- merge(naturalex_gtp, mock_assignment, by="account")
stopifnot(check==sum(naturalex_gtp$won_lottery))

# Take Patente de Rodados out
# We exclude this tax fromt the analysis bacause we need to further work with the
# municipality to understand how they implemented the tax holiday in order to
# set up the analysis of the natural experiment correctly. Once we are able to do
# this, we will reincorporate it to the analysis (and the adjustments for multiple
# comparisons)
naturalex_gtp <- naturalex_gtp[naturalex_gtp$tax!="Patente de Rodados",]

# We use a panel dastaset to analyse the natural experiment as it is allows to build
# the plots and recenter all holidays such that they overlap.
table(naturalex_gtp$tax)
```

```
##
```



```
## Contribucion Inmobiliaria        Patente de Rodados
##                    112904                         0
##               Saneamiento    Tributos Domiciliarios
##                     66000                    172216
```

```
levels(naturalex_gtp$tax) <- c("Property", "Vehicle", "Sewage", "Head")

# Take out years within tax holiday
naturalex_gtp <- naturalex_gtp[is.na(naturalex_gtp$plot_time)==F,]

# cleaning environment
rm(list=(ls()[ls()!="naturalex_gtp"]))
```

Debt data For confidentiality reasons, the municipality could not give us debt data of a given account number. We could, however, list the account numbers in each tax type, taxpayer type and treatment (won_lottery) combination and obtain a list of the debt amounts corresponding to such category (here not linked to an account number).

For pragmatic reasons, we add this data to our main dataset following tax type[1], but the link of the debt data to a given account is ficticious.

```
load("~/Dropbox/Uruguay state capacity/Analysis/data/final_data/naturalex_debt_gtp.Rda")

naturalex_debt_gtp$won_lottery <- sample(naturalex_debt_gtp$won_lottery,
                                     length(naturalex_debt_gtp$won_lottery),
                                     replace=F)
```

Field Experiment

```
load("~/Dropbox/Uruguay state capacity/Analysis/data/final_data/fieldex_data.Rda")

# Generating mock treatment assignment
fieldex$treatment[fieldex$type=="eligible"] <-
  sample(fieldex$treatment[fieldex$type=="eligible"],
         length(fieldex$treatment[fieldex$type=="eligible"]), replace=F)

fieldex$treatment[fieldex$type=="noneligible"] <-
  sample(fieldex$treatment[fieldex$type=="noneligible"],
         length(fieldex$treatment[fieldex$type=="noneligible"]), replace=F)

### Adding compliance indicator
fieldex$compliance_july <- ifelse((fieldex$july_nrbills_owed==0 & fieldex$july_ontime==1),1,0)
```

Our field experiment data includes cases that could not be treated by our July experiment, either because they had paid their bill in advance or because they are retired (retired individuals have differente due dates for their bills).

[1] Note that (1) all taxpayers in the main dataset are goodtaxpayers and (2) for the real analysis we will follow tax-treatment combinations.

```r
# Retired individuals in sample
table(fieldex$retired, fieldex$type)
```

```
##
##     eligible noneligible
##   0    14300       14123
##   1        0         177
```

```r
table(fieldex$retired, fieldex$treatment) # Should be balanced by treatment
```

```
##
##        0    1    2    3    4    5
##   0 5666 2883 2881 5670 5668 5655
##   1   34   17   19   30   32   45
```

```r
# Paid in advance in sample
table(fieldex$paid_in_advance, fieldex$type)
```

```
##
##     eligible noneligible
##   0     9461       12483
##   1     4839        1817
```

```r
table(fieldex$paid_in_advance, fieldex$treatment) # Should be balanced by treatment
```

```
##
##        0    1    2    3    4    5
##   0 4395 2255 2235 4353 4360 4346
##   1 1305  645  665 1347 1340 1354
```

We exclude both this types of cases from our analysis.

```r
fieldex <- fieldex[fieldex$retired==0 & fieldex$paid_in_advance==0,]
# We might use them for placebo tests later.
```

This is the new distribution of eligible and ineligible individuals across treatment conditions.

```
##
##                   0    1    2    3    4    5
##   eligible    1901  968  965 1871 1881 1875
##   noneligible 2460 1271 1251 2452 2447 2426
```

Mock data for the variables we do not have yet

```r
# Field experiment
fieldex$current_debt_DiD  <- rnorm(nrow(fieldex),0,100000)
```


Impact of the tax holiday lottery

Bills and tax holidays

CI: Property tax PR: Vehicle tax TS : Sewage TD: Head

New information about how the lottery and the tax holiday work in practice suggest we should review some of the details of our analysis plan. This section summarizes how the tax holiday system works and how it speaks to our research design.

Note that the first part of the process is common to all taxes: the lottery takes place and the IT department filters accounts[2] for which the last four digits match the lottery value. Because the municipality has no way of distinguishing "personas fisicas" from "personas juridicas", all these accounts are notified that they have won and are encouraged to go to the municipality with proof that they are "personas fisicas" to finish the process that effectively concedes the tax holiday.

All winners (regardless of the lottery), go to the same office ("gestion de contribuyentes") to prove they are "personas juridicas". Note that (a) some winners might never do this, in which case they do not get the tax holiday[3] and (b) winners might not do this as soon as they get the note telling them they have won. Although by law they are supposed to contact the municipality in the 90 day window after they receive the note in order to be eligible for they prize, municipal employees have told us they "would never deny anyone the prize", regardless of when it is they go to "cash it". This point is important, as it suggests a first possible source of delay between the time an account wins the lottery and when the tax holiday is effective.

The process then continues in the office in charge of the tax, which has the final power to decide what bills the tax holiday will effectively apply to. One office ("Impuestos inmobiliarios") controls both the CI and TD. PR and TS are under the jurisdiction of two other offices ("Ingresos Vehiculares" and "Administración de Saneamiento") respectively. Despite by law the tax holiday should already be in place for the bill that immediately follows the lottery, each office tends to adjust the policy to their own admistrative/bureaucratic procedure which, together with difference in the number of annual bills and due dates, introduces some between tax variation.

CI consists of 3 bills a year, in March, July and November one every four months. Since lotteries take place every two months, there are two lotteries per CI bill. The office in charge of this tax implemented the policy following closely its design, and thus most tax holidays start for the bill immediately after the lottery (see table below). Because of the combined timing of bills and lotteries, lotteries always take place between one bill and the next. If $t = 0$ describes the first payment under the tax holiday and $t = 1$ the first payment before the holiday, for CI the lottery always takes place between these two periods. Thus, a good taxpayer by the time of the lottery has complied with the tax for $t = (-3, -2, -1)$. One possible concern with CI is that, because it is an anual tax of a fixed amount, taxpayers can choose to pay the annual rate, paying by march the bills for all the year. However, one can only pay in advance the bills for July and August, but only payments for the current year can be payed in this way. This can also add some delay for the tax holiday to take place, since winners that win after paying the year in advance will not "cash" their prize until the March payment of the following year.

TD consists of 6 bills a year, in February, April, June, August, October and December. Notice that these are also the months when the lotteries take place. These tax holidays were implemented by the same office processing the CI, and its implementation also follows the policy guideline closely: tax holidays start with the first bill that follows the lottery. Unlike CI, however, here the lottery does not take place between payments, they overlap (see table below). Following the notation for CI, here the lottery takes place at $t = -1$. This introduces one specificity to the analysis of the TD holiday lottery. Because they are concurrent, the IT office does not know yet if the bill for that month is correctly paid. Thus, a good taxpayer for TS has, at $t = -1$–the time of the lottery–complied with the tax for $t = (-7, -6, -5, -4, -3, -2)$. By the time the holiday kicks in, it is possible that a winner is no longer a good taxpayer–if she hasn't paid the bill corresponding to $t = -1$. A

[2] We confirmed with the municipality account numbers are unique (ie. they are not repeated across lotteries).

[3] The municipality estimates that around 25% of the people that are registered with real names (and thus likely to be personas fisicas) never engage in the process to get the prize.

natural question here would be if the taxpayer is informed that she has won *before* the due date of the $t = -1$ bill–which overlaps with the lottery. Whereas this is not impossible, it is highly unlikely as the lotteries take place during the last week of the month. TD cannot be paid in advance.

For **TS** there is also 6 bills a year, in January, March, May, July, September and November. The implementation of the tax holiday is close to the policy design, although in practice, the tax holiday does not start with the bill that immediatly follows the lottery but the next. One can think of this case as if the lottery took place between $t = -2$ and $t = -1$. Despite the fact that the lottery takes place *before* the $t = -1$ bill, there is not enough time for the tax holiday to kick in. If someone wins the December lottery, for example, they are judged as good or bad taxpayers looking a the relevant previous six payments, starting by the December bill of the previous year. But by the end of Decembre (when the lottery happes) the municipality is already processing the January bills. Thus, the tax holiday does not kick in until the following bill, in this case March. As for TD, by the time the holiday kicks in, it is possible that a winner is no longer a good taxpayer–if she hasn't paid the bill corresponding to $t = -1$. It is also the case that winner eligibility is judged by the $t = (-7, -6, -5, -4, -3, -2)$ payments. Yet unlike TD, for TS it *is* likely that taxpayers know they have won by the time of their $t = -1$ payment, even if they do not get the tax holiday until $t = 0$. TS cannot be paid in advance.

Finally, **PR** works somewhat different from the other taxes. There are 3 payments—January, May and September—which taxpayers can pay in advance, just as for CI. Despite our close engagement with the municipality, we have not yet been able to understand the concrete process by which the lottery was implemented. For that reason we have excluded this tax from our analysis, but we will add it as soon as we understand the process by which the tax holiday was allocated.

```
|    | FEB | APR | JUN | AUG | OCT | DEC |
|----|-----|-----|-----|-----|-----|-----|
| CI | MAR | JUL | JUL | NOV | NOV | MAR |
| PR | JAN | JAN | JAN | JAN | JAN | JAN |
| TD | APR | JUN | AUG | OCT | DEC | FEB |
| TS | MAY | JUL | SEP | NOV | JAN | MAR |
```

Functions

```
t.test.se <- function(y,x){ # t test with SEs
  diff <- mean(y[x==1],na.rm=T) - mean(y[x==0],na.rm=T)
  sd1a.N <- var(y[x==1],na.rm=T)/length(na.omit(y[x==1]))
  sd1b.N <- var(y[x==0],na.rm=T)/length(na.omit(y[x==0]))
  se.diff <- sqrt(sd1a.N + sd1b.N)
  t <- diff/se.diff
  return(c(diff,se.diff))
}

# Function to test the difference of the differences
comp.eff <- function(dat1_y, dat1_x, dat2_y, dat2_x){

  tt1 <- t.test.se(dat1_y, dat1_x)
  tt2 <- t.test.se(dat2_y, dat2_x)

  dms <- as.data.frame(rbind(tt1, tt2))
  rownames(dms) <- c("DM 1", "DM 2")
  colnames(dms) <- c("DM", "SE")
  print(dms)
```



```
  diff <- tt1[1]-tt2[1]
  se.diff <- sqrt((tt1[2]^2)+(tt2[2]^2))

  res <- c(diff,se.diff)
  names(res) <- c("Diff in effects", "SE")

  return(res)
}
```

Plan for Data Analysis

Hypothesis 1A.

Winning the tax holiday lottery leads to an *increase* in future tax compliance.

Hypothesis 1B.

Winning the lottery leads to a *decrease* in future tax compliance.

Hypothesis 1C.

Winning the lottery leads to *no change* in future tax compliance.

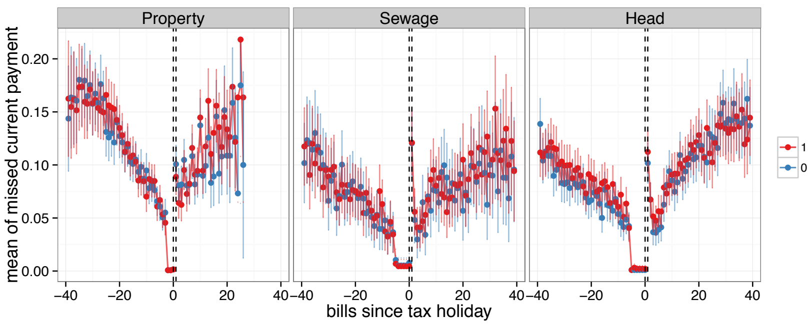
Graphical Analysis We construct a plot in which the horizontal axis is time, measured in tax holiday lotteries (which occur every two or four months, depending on the tax). The vertical axis is one of our three measures of tax payment (compliance, missed payment, or number of payments owed). Note the x axis is not continuous. The left side shows the payment history until an account won or would have won the lottery, centered at 0. The right side of the axis show payments after the tax holiday, starting at 1. The time between 0 and 1 might vary for different taxpayers.[4] Parting the x axis allows for better comparison.

```
# summarizing data to plot
plot_data <- ddply(naturalex_gtp, c("tax", "plot_time", "won_lottery"),
            summarise,
            N  = length(unique(account)),
            nr_missed_mean = mean(nr_missed_payments, na.rm=T),
            se_nr_missed_mean = sd(nr_missed_payments, na.rm=T)/sqrt(N),
            nr_missed_upper= nr_missed_mean + qnorm(.975)*(se_nr_missed_mean),
            nr_missed_lower= nr_missed_mean - qnorm(.975)*(se_nr_missed_mean),
            missed_payment_mean = mean(missed_payment, na.rm=T),
            se_missed_payment_mean = sd(missed_payment, na.rm=T)/sqrt(N),
            missed_payment_upper= missed_payment_mean + qnorm(.975)*(se_missed_payment_mean),
            missed_payment_lower= missed_payment_mean - qnorm(.975)*(se_missed_payment_mean),
            compliance_mean = mean(compliance, na.rm=T),
            se_compliance_mean = sd(compliance, na.rm=T)/sqrt(N),
            compliance_upper= compliance_mean + qnorm(.975)*se_compliance_mean,
            compliance_lower= compliance_mean - qnorm(.975)*se_compliance_mean
            )
```

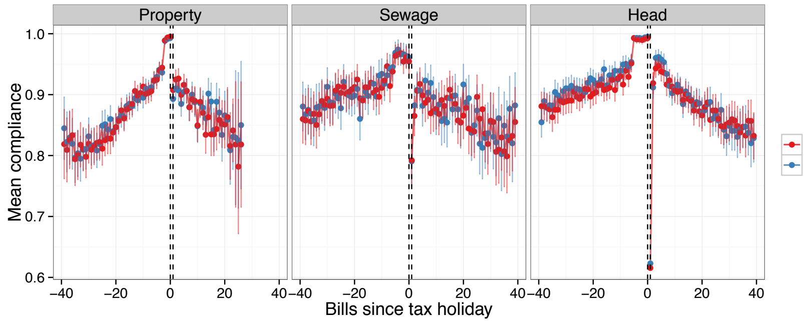
[4] The final plots will use bill date so that the x axis can be interpreted as time since winning (an the taxes are rescaled such that there are 6 payments of TD and TS for every 3 of CI and PR).

```
plot_data <- na.omit(plot_data[plot_data$plot_time>-40 &
                        plot_data$plot_time<40,])
```

```
### Plots by tax
plot_data$won_lottery <- as.factor(plot_data$won_lottery)
plot_data$won_lottery <- relevel(plot_data$won_lottery, ref="1")
plot_data$after <- as.factor(plot_data$plot_time>0)

# missed payment
p <- ggplot(plot_data, aes(x=plot_time,y=missed_payment_mean,
                           group=c(after), color=won_lottery))
p + geom_line(size=.6, alpha=.7) + facet_wrap(~tax) +
  geom_point(size=2.4)  +
  ylim(min(plot_data$missed_payment_mean), max(plot_data$missed_payment_mean)) +
    xlab("bills since tax holiday") +
  ylab("mean of missed current payment") +
  geom_errorbar(aes(ymin=missed_payment_lower,
                    ymax=missed_payment_upper),
              # colour="blue",
               width=.3, alpha=.5) +
  scale_color_brewer(palette="Set1") +
  #scale_colour_manual(
#  values = c("0" = "navyblue","1" = "red")) +
  #ggtitle("Mean Debt (Property Tax, 2000-2014)") +
  theme_bw() +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1.25)),
        axis.title.x = element_text(size = rel(1.25)),
        strip.text.x = element_text(size = rel(1.5)),
        legend.title=element_blank()) +
  geom_vline(aes(xintercept = 1), linetype="dashed") +
  geom_vline(aes(xintercept = 0), linetype="dashed")
```

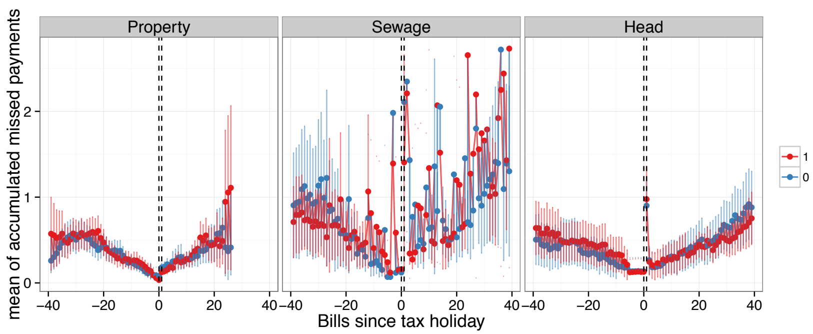



```
# compliance
p <- ggplot(plot_data, aes(x=plot_time,y=compliance_mean,
                           group=c(after), color=won_lottery))
p + geom_line(size=.6, alpha=.7) + facet_wrap(~tax) +
  geom_point(size=2.4)  +
  ylim(min(plot_data$compliance_mean), max(plot_data$compliance_mean)) +
    xlab("Bills since tax holiday") +
  ylab("Mean compliance") +
  geom_errorbar(aes(ymin=compliance_lower,
                    ymax=compliance_upper),
              # colour="blue",
               width=.3, alpha=.5) +
  scale_color_brewer(palette="Set1") +
  #scale_colour_manual(
#  values = c("0" = "navyblue","1" = "red")) +
  #ggtitle("Mean Debt (Property Tax, 2000-2014)") +
  theme_bw() +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1.25)),
        axis.title.x = element_text(size = rel(1.25)),
        strip.text.x = element_text(size = rel(1.5)),
        legend.title=element_blank()) +
  geom_vline(aes(xintercept = 1), linetype="dashed") +
  geom_vline(aes(xintercept = 0), linetype="dashed")
```




```
# nr missed payments
p <- ggplot(plot_data, aes(x=plot_time,y=nr_missed_mean,
                           group=c(after), color=won_lottery))
p + geom_line(size=.6, alpha=.7) + facet_wrap(~tax) +
  geom_point(size=2.4)  +
  ylim(min(plot_data$nr_missed_mean), max(plot_data$nr_missed_mean)) +
    xlab("Bills since tax holiday") +
  ylab("mean of accumulated missed payments") +
  geom_errorbar(aes(ymin=nr_missed_lower,
                    ymax=nr_missed_upper),
              # colour="blue",
```

```
                width=.3, alpha=.5) +
  scale_color_brewer(palette="Set1") +
  #scale_colour_manual(
#  values = c("0" = "navyblue","1" = "red")) +
  #ggtitle("Mean Debt (Property Tax, 2000-2014)") +
  theme_bw() +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1.25)),
        axis.title.x = element_text(size = rel(1.25)),
        strip.text.x = element_text(size = rel(1.5)),
        legend.title=element_blank()) +
  geom_vline(aes(xintercept = 1), linetype="dashed") +
  geom_vline(aes(xintercept = 0), linetype="dashed")
```

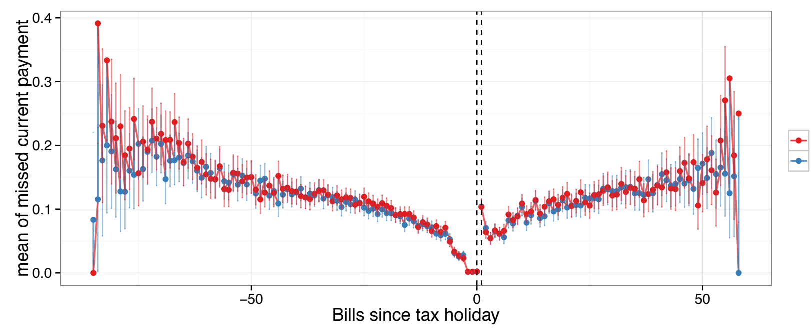



```
#### All taxes in same plot
plot_data2 <- ddply(naturalex_gtp, c("plot_time", "won_lottery"),
                   summarise,
            N  = length(unique(account)),
            nr_missed_mean = mean(nr_missed_payments, na.rm=T),
            se_nr_missed_mean = sd(nr_missed_payments, na.rm=T)/sqrt(N),
            nr_missed_upper= nr_missed_mean + qnorm(.975)*(se_nr_missed_mean),
            nr_missed_lower= nr_missed_mean - qnorm(.975)*(se_nr_missed_mean),
            missed_payment_mean = mean(missed_payment, na.rm=T),
            se_missed_payment_mean = sd(missed_payment, na.rm=T)/sqrt(N),
            missed_payment_upper= missed_payment_mean + qnorm(.975)*(se_missed_payment_mean),
            missed_payment_lower= missed_payment_mean - qnorm(.975)*(se_missed_payment_mean),
            compliance_mean = mean(compliance, na.rm=T),
            se_compliance_mean = sd(compliance, na.rm=T)/sqrt(N),
            compliance_upper= compliance_mean + qnorm(.975)*se_compliance_mean,
            compliance_lower= compliance_mean - qnorm(.975)*se_compliance_mean
            )
```

```
plot_data2$won_lottery <- as.factor(plot_data2$won_lottery)
plot_data2$won_lottery <- relevel(plot_data2$won_lottery, ref="1")
plot_data2$after <- as.factor(plot_data2$plot_time>0)
```

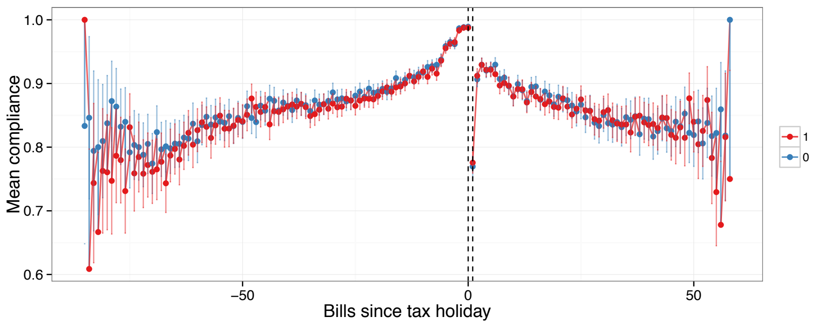


```
# missed payment
p <- ggplot(plot_data2, aes(x=plot_time,y=missed_payment_mean,
                           group=c(after), color=won_lottery))
p + geom_line(size=.7, alpha=.7) +
  geom_point(size=2.4)  +
  ylim(min(plot_data2$missed_payment_mean), max(plot_data2$missed_payment_mean)) +
    xlab("Bills since tax holiday") +
  ylab("mean of missed current payment") +
  geom_errorbar(aes(ymin=missed_payment_lower,
                    ymax=missed_payment_upper),
              # colour="blue",
               width=.3, alpha=.5) +
  scale_color_brewer(palette="Set1") +
  #scale_colour_manual(
#  values = c("0" = "navyblue","1" = "red")) +
  #ggtitle("Mean Debt (Property Tax, 2000-2014)") +
  theme_bw() +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1.25)),
        axis.title.x = element_text(size = rel(1.25)),
        strip.text.x = element_text(size = rel(1.5)),
        legend.title=element_blank()) +
  geom_vline(aes(xintercept = 1), linetype="dashed") +
  geom_vline(aes(xintercept = 0), linetype="dashed")
```




```
# compliance
p <- ggplot(plot_data2, aes(x=plot_time,y=compliance_mean,
                           group=c(after), color=won_lottery))
p + geom_line(size=.6, alpha=.7) +
  geom_point(size=2.4)  +
  ylim(min(plot_data2$compliance_mean), max(plot_data2$compliance_mean)) +
    xlab("Bills since tax holiday") +
  ylab("Mean compliance") +
  geom_errorbar(aes(ymin=compliance_lower,
                    ymax=compliance_upper),
```

```
                # colour="blue",
                 width=.3, alpha=.5) +
  scale_color_brewer(palette="Set1") +
  #scale_colour_manual(
#  values = c("0" = "navyblue","1" = "red")) +
  #ggtitle("Mean Debt (Property Tax, 2000-2014)") +
  theme_bw() +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1.25)),
        axis.title.x = element_text(size = rel(1.25)),
        strip.text.x = element_text(size = rel(1.5)),
        legend.title=element_blank()) +
  geom_vline(aes(xintercept = 1), linetype="dashed") +
  geom_vline(aes(xintercept = 0), linetype="dashed")
```




```
# nr missed payments
p <- ggplot(plot_data2, aes(x=plot_time,y=nr_missed_mean,
                             group=c(after), color=won_lottery))
p + geom_line(size=.7, alpha=1) +
  geom_point(size=2.4)  +
  ylim(min(plot_data2$nr_missed_mean), max(plot_data2$nr_missed_mean)) +
    xlab("Bills since tax holiday") +
  ylab("mean of accumulated missed payments") +
  geom_errorbar(aes(ymin=nr_missed_lower,
                    ymax=nr_missed_upper),
                # colour="blue",
                 width=.3, alpha=.5) +
  scale_color_brewer(palette="Set1") +
  #scale_colour_manual(
#  values = c("0" = "navyblue","1" = "red")) +
  #ggtitle("Mean Debt (Property Tax, 2000-2014)") +
  theme_bw() +
  theme(plot.title = element_text(size = rel(1.75)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1.25)),
        axis.title.x = element_text(size = rel(1.25)),
```



```
        strip.text.x = element_text(size = rel(1.5)),
        legend.title=element_blank()) +
  geom_vline(aes(xintercept = 1), linetype="dashed") +
  geom_vline(aes(xintercept = 0), linetype="dashed")
```

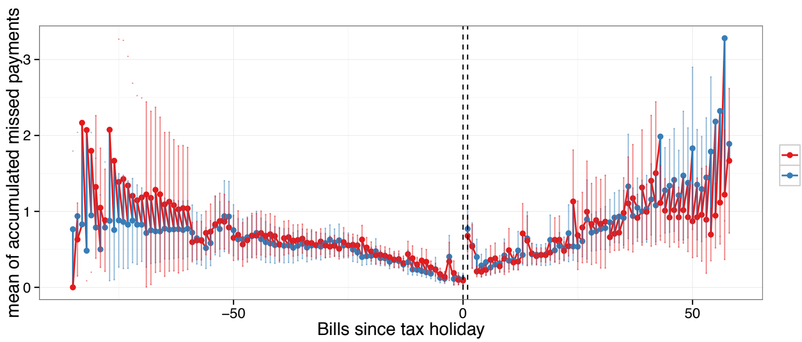


TABLE 1. NATURAL EXPERIMENT. Effects of the tax holiday (difference in differences analysis). Effects of the tax holiday. Comparing winners to non-winners, difference in difference analysis (comparison A=mean of the year before winning vs. mean of the year after the tax holiday; comparison B= mean of three years before winning vs. mean of three years after tax holiday). Tests using compliance as an outcome are conditional on finding effects for either missed payments or number of payments owed for the relevant period. This is because compliance is a stricter test, and if we find effects for neither missed payments of number of payments owed, there will be no effects by construction for compliance.

```
# 1 year diff in diff setup

dd_data <- ddply(naturalex_gtp, c("account","tax", "won_lottery"),
                 summarise,
            compliance_mean_6_DiD_1 = mean(compliance[plot_time>0 & plot_time<=6],
                                       na.rm=T)-
              mean(compliance[plot_time<0 & plot_time>=(-6)],
                                       na.rm=T),
            compliance_mean_3_DiD_1 = mean(compliance[plot_time>0 & plot_time<=3],
                                       na.rm=T)-
              mean(compliance[plot_time<0 & plot_time>=(-3)],
                                       na.rm=T),
            missed_payment_mean_6_DiD_1 = mean(missed_payment[plot_time>0 & plot_time<=6],
                                       na.rm=T)-
              mean(missed_payment[plot_time<0 & plot_time>=(-6)],
                                       na.rm=T),
            missed_payment_mean_3_DiD_1 = mean(missed_payment[plot_time>0 & plot_time<=3],
                                       na.rm=T)-
              mean(missed_payment[plot_time<0 & plot_time>=(-3)],
                                       na.rm=T),
            nr_missed_payments_mean_6_DiD_1 = mean(nr_missed_payments[plot_time>0 & plot_time<=6],
                                       na.rm=T)-
              mean(nr_missed_payments[plot_time<0 & plot_time>=(-6)],
                                       na.rm=T),
```

```r
            nr_missed_payments_mean_3_DiD_1 = mean(nr_missed_payments[plot_time>0 & plot_time<=3],
                                      na.rm=T)-
              mean(nr_missed_payments[plot_time<0 & plot_time>=(-3)],
                                      na.rm=T))

dd_data$compliance_mean_DiD_1yr[
  dd_data$tax=="Property"] <- dd_data$compliance_mean_3_DiD_1[dd_data$tax=="Property"]

dd_data$compliance_mean_DiD_1yr[dd_data$tax=="Head" | dd_data$tax=="Sewage"] <-
  dd_data$compliance_mean_6_DiD_1[dd_data$tax=="Head" | dd_data$tax=="Sewage"]

dd_data$missed_payment_mean_DiD_1yr[
  dd_data$tax=="Property"] <- dd_data$missed_payment_mean_3_DiD_1[dd_data$tax=="Property"]

dd_data$missed_payment_mean_DiD_1yr[dd_data$tax=="Head"| dd_data$tax=="Sewage"] <-
  dd_data$missed_payment_mean_6_DiD_1[dd_data$tax=="Head"| dd_data$tax=="Sewage"]

dd_data$nr_missed_payments_mean_DiD_1yr[
  dd_data$tax=="Property"] <- dd_data$nr_missed_payments_mean_3_DiD_1[dd_data$tax=="Property"]

dd_data$nr_missed_payments_mean_DiD_1yr[dd_data$tax=="Head"| dd_data$tax=="Sewage"] <-
  dd_data$nr_missed_payments_mean_6_DiD_1[dd_data$tax=="Head"| dd_data$tax=="Sewage"]

######## 1 year diff in diff

# all taxes, compliance
with(dd_data,
     t.test(compliance_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_1yr by won_lottery
## t = -0.3488, df = 4921, p-value = 0.7272
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.013585  0.009481
## sample estimates:
## mean in group 0 mean in group 1
##        -0.08039        -0.07834
```

```r
# all taxes, missed payment
with(dd_data,
     t.test(missed_payment_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_1yr by won_lottery
## t = 0.4444, df = 4914, p-value = 0.6568
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```



```
##  -0.007641  0.012120
## sample estimates:
## mean in group 0 mean in group 1
##         0.05699         0.05475
```

```r
# all taxes, nr of missed payments
with(dd_data,
     t.test(nr_missed_payments_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_1yr by won_lottery
## t = 0.8604, df = 4962, p-value = 0.3896
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08389  0.21512
## sample estimates:
## mean in group 0 mean in group 1
##          0.2382          0.1726
```

```r
# Property, compliance
with(dd_data[dd_data$tax=="Property",],
     t.test(compliance_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_1yr by won_lottery
## t = -1.345, df = 2170, p-value = 0.1786
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.033008  0.006146
## sample estimates:
## mean in group 0 mean in group 1
##        -0.06914        -0.05571
```

```r
# Property, missed payment
with(dd_data[dd_data$tax=="Property",],
     t.test(missed_payment_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_1yr by won_lottery
## t = 1.786, df = 2152, p-value = 0.07427
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001588  0.033944
## sample estimates:
## mean in group 0 mean in group 1
##         0.07097         0.05479
```

```r
# Property, nr of missed payments
with(dd_data[dd_data$tax=="Property",],
     t.test(nr_missed_payments_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_1yr by won_lottery
## t = 0.7858, df = 2179, p-value = 0.4321
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.06716  0.15696
## sample estimates:
## mean in group 0 mean in group 1
##         0.07423         0.02932
```

```r
# Sewage, compliance
with(dd_data[dd_data$tax=="Sewage",],
     t.test(compliance_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_1yr by won_lottery
## t = -0.2429, df = 811.4, p-value = 0.8081
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03224  0.02514
## sample estimates:
## mean in group 0 mean in group 1
##        -0.08003        -0.07648
```

```r
# Sewage, missed payment
with(dd_data[dd_data$tax=="Sewage",],
     t.test(missed_payment_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_1yr by won_lottery
## t = -0.2318, df = 803.9, p-value = 0.8167
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01937  0.01528
## sample estimates:
## mean in group 0 mean in group 1
##         0.04575         0.04779
```

```r
# Sewage, nr of missed payments
with(dd_data[dd_data$tax=="Sewage",],
     t.test(nr_missed_payments_mean_DiD_1yr~won_lottery))
```



```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_1yr by won_lottery
## t = 0.9032, df = 778.6, p-value = 0.3667
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4489  1.2142
## sample estimates:
## mean in group 0 mean in group 1
##          0.8079          0.4252
```

```r
# Head, compliance
with(dd_data[dd_data$tax=="Head",],
     t.test(compliance_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_1yr by won_lottery
## t = 1.598, df = 1834, p-value = 0.1103
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.002779  0.027218
## sample estimates:
## mean in group 0 mean in group 1
##        -0.09317        -0.10538
```

```r
# Head, missed payment
with(dd_data[dd_data$tax=="Head",],
     t.test(missed_payment_mean_DiD_1yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_1yr by won_lottery
## t = -1.785, df = 1824, p-value = 0.07436
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.025586  0.001201
## sample estimates:
## mean in group 0 mean in group 1
##         0.04573         0.05793
```

```r
# Head, nr of missed payments
with(dd_data[dd_data$tax=="Head",],
     t.test(nr_missed_payments_mean_DiD_1yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_1yr by won_lottery
```

```
## t = -0.4974, df = 1834, p-value = 0.619
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1112  0.0662
## sample estimates:
## mean in group 0 mean in group 1
##          0.1982          0.2207
```

```
# 3 year diff in diff setup

dd_data <- ddply(naturalex_gtp, c("account","tax", "won_lottery"),
                   summarise,
            compliance_mean_18_DiD_3 = mean(compliance[plot_time>0 & plot_time<=18],
                                            na.rm=T) -
              mean(compliance[plot_time<0 & plot_time>=(-18)],
                   na.rm=T),

            compliance_mean_9_DiD_3 = mean(compliance[plot_time>0 & plot_time<=9],
                                            na.rm=T) -
              mean(compliance[plot_time<0 & plot_time>=(-9)],
                   na.rm=T),

            missed_payment_mean_18_DiD_3 = mean(missed_payment[plot_time>0 & plot_time<=18],
                                                na.rm=T) -
              mean(missed_payment[plot_time<0 & plot_time>=(-18)],
                            na.rm=T),

            missed_payment_mean_9_DiD_3 = mean(missed_payment[plot_time>0 & plot_time<=9],
                                                na.rm=T) -
              mean(missed_payment[plot_time<0 & plot_time>=(-9)],
                   na.rm=T),

            nr_missed_payments_mean_18_DiD_3 = mean(nr_missed_payments[plot_time>0 & plot_time<=18],
                                                    na.rm=T)-
              mean(nr_missed_payments[plot_time<0 & plot_time>=(-18)],
                   na.rm=T),

            nr_missed_payments_mean_9_DiD_3 = mean(nr_missed_payments[plot_time>0 & plot_time<=9],
                            na.rm=T)-
              mean(nr_missed_payments[plot_time<0 & plot_time>=(-9)],
                            na.rm=T))


dd_data$compliance_mean_DiD_3yr[
  dd_data$tax=="Property"] <- dd_data$compliance_mean_9_DiD_3[dd_data$tax=="Property"]

dd_data$compliance_mean_DiD_3yr[dd_data$tax=="Head" | dd_data$tax=="Sewage"] <-
  dd_data$compliance_mean_18_DiD_3[dd_data$tax=="Head" | dd_data$tax=="Sewage"]

dd_data$missed_payment_mean_DiD_3yr[dd_data$tax=="Property"] <-
  dd_data$missed_payment_mean_9_DiD_3[dd_data$tax=="Property"]

dd_data$missed_payment_mean_DiD_3yr[dd_data$tax=="Head"| dd_data$tax=="Sewage"] <-
  dd_data$missed_payment_mean_18_DiD_3[dd_data$tax=="Head"| dd_data$tax=="Sewage"]
```



```
dd_data$nr_missed_payments_mean_DiD_3yr[
  dd_data$tax=="Property"] <- dd_data$nr_missed_payments_mean_9_DiD_3[dd_data$tax=="Property"]

dd_data$nr_missed_payments_mean_DiD_3yr[dd_data$tax=="Head"| dd_data$tax=="Sewage"] <-
  dd_data$nr_missed_payments_mean_18_DiD_3[dd_data$tax=="Head"| dd_data$tax=="Sewage"]

######## 3 year diff in diff

# all taxes, compliance
with(dd_data,
     t.test(compliance_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_3yr by won_lottery
## t = -1.341, df = 4916, p-value = 0.1798
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.018420  0.003453
## sample estimates:
## mean in group 0 mean in group 1
##        -0.04459        -0.03711
```

```
# all taxes, missed payment
with(dd_data,
     t.test(missed_payment_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_3yr by won_lottery
## t = 1.601, df = 4910, p-value = 0.1095
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001792  0.017741
## sample estimates:
## mean in group 0 mean in group 1
##         0.03580         0.02783
```

```
# all taxes, nr of missed payments
with(dd_data,
     t.test(nr_missed_payments_mean_DiD_3yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_3yr by won_lottery
## t = 1.956, df = 4839, p-value = 0.05054
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  -0.0002338  0.1986982
## sample estimates:
## mean in group 0 mean in group 1
##         0.16082         0.06159
```

```r
# Property, compliance
with(dd_data[dd_data$tax=="Property",],
     t.test(compliance_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_3yr by won_lottery
## t = -1.686, df = 2158, p-value = 0.09198
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.034263  0.002586
## sample estimates:
## mean in group 0 mean in group 1
##        -0.04129        -0.02545
```

```r
# Property, missed payment
with(dd_data[dd_data$tax=="Property",],
     t.test(missed_payment_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_3yr by won_lottery
## t = 2.005, df = 2141, p-value = 0.04504
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.000381 0.034079
## sample estimates:
## mean in group 0 mean in group 1
##         0.04199         0.02476
```

```r
# Property, nr of missed payments
with(dd_data[dd_data$tax=="Property",],
     t.test(nr_missed_payments_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_3yr by won_lottery
## t = 1.274, df = 2067, p-value = 0.2027
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03575  0.16840
## sample estimates:
## mean in group 0 mean in group 1
##        0.060352       -0.005971
```



```r
# Sewage, compliance
with(dd_data[dd_data$tax=="Sewage",],
     t.test(compliance_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_3yr by won_lottery
## t = -0.0469, df = 766.8, p-value = 0.9626
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02320  0.02212
## sample estimates:
## mean in group 0 mean in group 1
##        -0.03790        -0.03736
```

```r
# Sewage, missed payment
with(dd_data[dd_data$tax=="Sewage",],
     t.test(missed_payment_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_3yr by won_lottery
## t = 0.2239, df = 772.4, p-value = 0.8229
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01522  0.01914
## sample estimates:
## mean in group 0 mean in group 1
##         0.02827         0.02631
```

```r
# Sewage, nr of missed payments
with(dd_data[dd_data$tax=="Sewage",],
     t.test(nr_missed_payments_mean_DiD_3yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_3yr by won_lottery
## t = 1.322, df = 789.5, p-value = 0.1866
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1496  0.7662
## sample estimates:
## mean in group 0 mean in group 1
##          0.5641          0.2557
```

```r
# Head, compliance
with(dd_data[dd_data$tax=="Head",],
     t.test(compliance_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_3yr by won_lottery
## t = -0.0565, df = 1842, p-value = 0.955
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01648  0.01556
## sample estimates:
## mean in group 0 mean in group 1
##        -0.05094        -0.05048
```

```
# Head, missed payment
with(dd_data[dd_data$tax=="Head",],
     t.test(missed_payment_mean_DiD_3yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_3yr by won_lottery
## t = -0.034, df = 1835, p-value = 0.9729
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.0147  0.0142
## sample estimates:
## mean in group 0 mean in group 1
##         0.03184         0.03209
```

```
# Head, nr of missed payments
with(dd_data[dd_data$tax=="Head",],
     t.test(nr_missed_payments_mean_DiD_3yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_3yr by won_lottery
## t = 1.067, df = 1341, p-value = 0.2862
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.05465  0.18497
## sample estimates:
## mean in group 0 mean in group 1
##         0.11521         0.05005
```

TABLE 2. NATURAL EXPERIMENT. Effect of the tax holiday (T-test) Comparing winners to non-winners, difference of means test for the total debt as of October, 2014. We limit the comparison to pre-2013 winners as the tax holiday might still apply for more recent winners.

```
with(naturalex_debt_gtp[naturalex_debt_gtp$tax!="pr",],
     t.test(debt_amount ~ won_lottery))
```

```
##
```



```
##  Welch Two Sample t-test
##
## data:  debt_amount by won_lottery
## t = 0.9541, df = 3798, p-value = 0.3401
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -313.3  907.2
## sample estimates:
## mean in group 0 mean in group 1
##           709.3           412.4
```

TABLE 3. NATURAL EXPERIMENT. Persistence of the effects of the tax lottery CONDITIONAL TEST). Comparing winners to non-winners, difference in difference analysis looking at the change between the 5 years after the tax holiday and the 5 years before winning the lottery. We conduct this test only if we find effects for missed payments, number of payments owed or total debt for the 3 year window.

```
# 5 year diff in diff setup

dd_data <-
  ddply(naturalex_gtp, c("account","tax", "won_lottery"), summarise,
     compliance_mean_30_DiD_5 = mean(compliance[plot_time>0 & plot_time<=30],
                                       na.rm=T) -
       mean(compliance[plot_time<0 & plot_time>=(-30)], na.rm=T),

      compliance_mean_15_DiD_5 = mean(compliance[plot_time>0 & plot_time<=15],
                                       na.rm=T) -
       mean(compliance[plot_time<0 & plot_time>=(-15)], na.rm=T),

      missed_payment_mean_30_DiD_5 = mean(missed_payment[plot_time>0 & plot_time<=30],
                                           na.rm=T) -
       mean(missed_payment[plot_time<0 & plot_time>=(-30)], na.rm=T),

      missed_payment_mean_15_DiD_5 = mean(missed_payment[plot_time>0 & plot_time<=15],
                                           na.rm=T) -
       mean(missed_payment[plot_time<0 & plot_time>=(-15)], na.rm=T),

        nr_missed_payments_mean_30_DiD_5 = mean(nr_missed_payments[plot_time>0 & plot_time<=30],
                                               na.rm=T)-
       mean(nr_missed_payments[plot_time<0 & plot_time>=(-30)], na.rm=T),

        nr_missed_payments_mean_15_DiD_5 = mean(nr_missed_payments[plot_time>0 & plot_time<=15],
                                   na.rm=T)-
      mean(nr_missed_payments[plot_time<0 & plot_time>=(-15)],
                                   na.rm=T))


dd_data$compliance_mean_DiD_5yr[
  dd_data$tax=="Property"] <- dd_data$compliance_mean_15_DiD_5[dd_data$tax=="Property"]

dd_data$compliance_mean_DiD_5yr[dd_data$tax=="Head" | dd_data$tax=="Sewage"] <-
  dd_data$compliance_mean_30_DiD_5[dd_data$tax=="Head" | dd_data$tax=="Sewage"]

dd_data$missed_payment_mean_DiD_5yr[dd_data$tax=="Property"] <-
  dd_data$missed_payment_mean_15_DiD_5[dd_data$tax=="Property"]
```

```
dd_data$missed_payment_mean_DiD_5yr[dd_data$tax=="Head"| dd_data$tax=="Sewage"] <-
  dd_data$missed_payment_mean_30_DiD_5[dd_data$tax=="Head"| dd_data$tax=="Sewage"]

dd_data$nr_missed_payments_mean_DiD_5yr[
  dd_data$tax=="Property"] <- dd_data$nr_missed_payments_mean_15_DiD_5[dd_data$tax=="Property"]

dd_data$nr_missed_payments_mean_DiD_5yr[dd_data$tax=="Head"| dd_data$tax=="Sewage"] <-
  dd_data$nr_missed_payments_mean_30_DiD_5[dd_data$tax=="Head"| dd_data$tax=="Sewage"]

######## 5 year diff in diff

# all taxes, compliance
with(dd_data,
     t.test(compliance_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_5yr by won_lottery
## t = -0.8931, df = 4906, p-value = 0.3718
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.016503  0.006173
## sample estimates:
## mean in group 0 mean in group 1
##        -0.03423        -0.02907
```

```
# all taxes, missed payment
with(dd_data,
     t.test(missed_payment_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_5yr by won_lottery
## t = 1.117, df = 4897, p-value = 0.2639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.004383  0.015999
## sample estimates:
## mean in group 0 mean in group 1
##         0.02635         0.02054
```

```
# all taxes, nr of missed payments
with(dd_data,
     t.test(nr_missed_payments_mean_DiD_5yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_5yr by won_lottery
```



```
## t = 0.7472, df = 4826, p-value = 0.455
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.07256  0.16194
## sample estimates:
## mean in group 0 mean in group 1
##         0.10292         0.05823
```

```
# Property, compliance
with(dd_data[dd_data$tax=="Property",],
     t.test(compliance_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_5yr by won_lottery
## t = -1.015, df = 2155, p-value = 0.3102
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.028519  0.009067
## sample estimates:
## mean in group 0 mean in group 1
##        -0.02758        -0.01785
```

```
# Property, missed payment
with(dd_data[dd_data$tax=="Property",],
     t.test(missed_payment_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_5yr by won_lottery
## t = 1.185, df = 2136, p-value = 0.2362
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.006846  0.027749
## sample estimates:
## mean in group 0 mean in group 1
##         0.02761         0.01716
```

```
# Property, nr of missed payments
with(dd_data[dd_data$tax=="Property",],
     t.test(nr_missed_payments_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_5yr by won_lottery
## t = 0.4836, df = 2114, p-value = 0.6287
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08083  0.13374
```

```
## sample estimates:
## mean in group 0 mean in group 1
##        0.002816       -0.023641
```

```r
# Sewage, compliance
with(dd_data[dd_data$tax=="Sewage",],
     t.test(compliance_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_5yr by won_lottery
## t = 0.0521, df = 746.5, p-value = 0.9584
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02245  0.02368
## sample estimates:
## mean in group 0 mean in group 1
##        -0.03139        -0.03200
```

```r
# Sewage, missed payment
with(dd_data[dd_data$tax=="Sewage",],
     t.test(missed_payment_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_5yr by won_lottery
## t = -0.0023, df = 778.1, p-value = 0.9982
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01918  0.01913
## sample estimates:
## mean in group 0 mean in group 1
##         0.01742         0.01744
```

```r
# Sewage, nr of missed payments
with(dd_data[dd_data$tax=="Sewage",],
     t.test(nr_missed_payments_mean_DiD_5yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_5yr by won_lottery
## t = 0.3846, df = 793.4, p-value = 0.7006
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4039  0.6008
## sample estimates:
## mean in group 0 mean in group 1
##          0.4681          0.3697
```



```r
# Head, compliance
with(dd_data[dd_data$tax=="Head",],
     t.test(compliance_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_mean_DiD_5yr by won_lottery
## t = -0.2447, df = 1877, p-value = 0.8067
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01931  0.01502
## sample estimates:
## mean in group 0 mean in group 1
##        -0.04283        -0.04069
```

```r
# Head, missed payment
with(dd_data[dd_data$tax=="Head",],
     t.test(missed_payment_mean_DiD_5yr ~ won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_mean_DiD_5yr by won_lottery
## t = 0.3309, df = 1877, p-value = 0.7408
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01265  0.01778
## sample estimates:
## mean in group 0 mean in group 1
##         0.02847         0.02590
```

```r
# Head, nr of missed payments
with(dd_data[dd_data$tax=="Head",],
     t.test(nr_missed_payments_mean_DiD_5yr~won_lottery))
```

```
##
##  Welch Two Sample t-test
##
## data:  nr_missed_payments_mean_DiD_5yr by won_lottery
## t = 0.6976, df = 1555, p-value = 0.4855
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1126  0.2369
## sample estimates:
## mean in group 0 mean in group 1
##         0.07169         0.00953
```

Mechanisms for H1A

Mechanism 1A.1: ***Informational.***

TABLE 4. FIELD EXPERIMENT. Informational mechanism. Good and bad taxpayers, comparison of treatments 1, 2 and 4 (pooled) vs. the placebo control group (treatment 0). First differences use the value of the dependent variable for the pre-treatment period (March 2014).

```
### Creating first differences outcomes for the field experiment

fieldex$missed_payment_DiD  <- fieldex$july_ontime - fieldex$march_ontime  # missed payment
fieldex$web_bill_DiD  <- fieldex$july_web_access - fieldex$march_web_access # web access
fieldex$payments_owed_DiD  <- fieldex$july_nrbills_owed - fieldex$adeudadas_2014_MAR # nr payments owed
names(fieldex)
```

```
##  [1] "account"               "due_date"
##  [3] "bill_adddress"         "property_value_2014"
##  [5] "bill_postalcode"       "has_debt_since"
##  [7] "adeudadas_2014_MAR"    "treatment"
##  [9] "type"                  "localidad"
## [11] "phase"                 "correlativo"
## [13] "address_freq"          "bill_padron"
## [15] "property_padron"       "property_address"
## [17] "retired"               "paid_in_advance"
## [19] "march_web_access"      "march_web_access_date"
## [21] "march_ontime"          "july_ontime"
## [23] "july_nrbills_owed"     "july_web_access"
## [25] "july_web_access_date"  "adeudadas_2009_MAR"
## [27] "adeudadas_2009_JUL"    "adeudadas_2009_NOV"
## [29] "adeudadas_2010_MAR"    "adeudadas_2010_JUL"
## [31] "adeudadas_2010_NOV"    "adeudadas_2011_MAR"
## [33] "adeudadas_2011_JUL"    "adeudadas_2011_NOV"
## [35] "adeudadas_2012_MAR"    "adeudadas_2012_JUL"
## [37] "adeudadas_2012_NOV"    "adeudadas_2013_MAR"
## [39] "adeudadas_2013_JUL"    "adeudadas_2013_NOV"
## [41] "compliance_july"       "current_debt_DiD"
## [43] "missed_payment_DiD"    "web_bill_DiD"
## [45] "payments_owed_DiD"
```

```
# compliance
fieldex$compliance_march <- ifelse(fieldex$march_ontime==1 & fieldex$adeudadas_2014_MAR==0, 1, 0)
fieldex$compliance_july <- ifelse(fieldex$july_ontime==1 & fieldex$july_nrbills_owed==0, 1, 0)
fieldex$compliance_DiD <- fieldex$compliance_july - fieldex$compliance_march

## Pooling treatments 1, 2, and 4 vs 0
fieldex$pooled_124_0 <- NA
fieldex$pooled_124_0[fieldex$treatment==0] <- 0
fieldex$pooled_124_0[fieldex$treatment==1] <- 1
fieldex$pooled_124_0[fieldex$treatment==2] <- 1
fieldex$pooled_124_0[fieldex$treatment==4] <- 1
```

```
# Missed payment
with(fieldex, t.test(missed_payment_DiD  ~ pooled_124_0))
```



```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_DiD by pooled_124_0
## t = 0.0254, df = 8488, p-value = 0.9797
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01506  0.01545
## sample estimates:
## mean in group 0 mean in group 1
##         0.02707         0.02687
```

```
# Web bill request
with(fieldex, t.test(web_bill_DiD ~ pooled_124_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  web_bill_DiD by pooled_124_0
## t = -0.5018, df = 8392, p-value = 0.6158
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.016399  0.009715
## sample estimates:
## mean in group 0 mean in group 1
##         0.07613         0.07947
```

```
# Payments owed
with(fieldex, t.test(payments_owed_DiD ~ pooled_124_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  payments_owed_DiD by pooled_124_0
## t = 1.561, df = 11034, p-value = 0.1185
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.007046  0.062209
## sample estimates:
## mean in group 0 mean in group 1
##          0.1844          0.1568
```

```
# Current debt
with(fieldex, t.test(current_debt_DiD ~ pooled_124_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  current_debt_DiD by pooled_124_0
## t = -1.755, df = 8730, p-value = 0.07925
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  -6901.6   380.8
## sample estimates:
## mean in group 0 mean in group 1
##           -1925            1336
```

```
# Compliance (test conditional on previous outcomes)
with(fieldex, t.test(compliance_DiD ~ pooled_124_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_DiD by pooled_124_0
## t = -0.1547, df = 8660, p-value = 0.8771
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01584  0.01353
## sample estimates:
## mean in group 0 mean in group 1
##         0.02552         0.02668
```

TABLE 6. FIELD EXPERIMENT: Comparison of effects for good and bad taxpayers: difference of the difference of means for the comparison of treatments 1, 2 and 4 (pooled) vs. the placebo control group (treatment 0).

```
# Missed payments
comp.eff(fieldex[fieldex$type=="eligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_124_0,
         fieldex[fieldex$type=="noneligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_124_0)
```

```
##             DM       SE
## DM 1  0.005281 0.006721
## DM 2 -0.004853 0.012627

## Diff in effects              SE
##         0.01013         0.01430
```

```
# Web access
comp.eff(fieldex[fieldex=="eligible",]$web_bill_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_124_0,
         fieldex[fieldex$type=="noneligible",]$web_bill_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_124_0)
```

```
##             DM       SE
## DM 1       NaN       NA
## DM 2 2.651e-05 0.008193

## Diff in effects              SE
##             NaN              NA
```


Mechanism 1A.2. *Attitudinal.*

Mechanisms for H1B

Regarding Hypothesis 1B, at least two mechanisms might explain a negative effect of winning the lottery on future compliance:

Mechanism 1B.1: *Income effects.*

TABLE 7. NATURAL EXPERIMENT. Income effects. Comparison of winners vs. non-winners: heterogeneous effects of winning the lottery by tax bracket.

```
# Coding tax brackets
  a <- 418958
  b <- 1047393
  c <- 2094784
  d <- 41895699

x <- naturalex_gtp$property_value_2014
naturalex_gtp$prop_tax_bracket <- ifelse(x==NA, NA,
                                         naturalex_gtp$prop_tax_bracket)
naturalex_gtp$prop_tax_bracket <- ifelse(x <= a, 1,
                                         naturalex_gtp$prop_tax_bracket)
naturalex_gtp$prop_tax_bracket <- ifelse(((x > a) & (x <= b)), 2,
                                         naturalex_gtp$prop_tax_bracket)
naturalex_gtp$prop_tax_bracket <- ifelse(((x > b) & (x <= c)), 3,
                                         naturalex_gtp$prop_tax_bracket)
naturalex_gtp$prop_tax_bracket <- ifelse(((x > c) & (x <= d)), 4,
                                         naturalex_gtp$prop_tax_bracket)
naturalex_gtp$prop_tax_bracket <- ifelse(x > d, 5,
                                         naturalex_gtp$prop_tax_bracket)

tax_bracket_data <- ddply(naturalex_gtp[is.na(naturalex_gtp$prop_tax_bracket)==F,],
                          c("prop_tax_bracket"),
                 summarise,
                 mean_missed_winners = mean(missed_payment[plot_time==1 & won_lottery==1],
                                            na.rm=T),
                 N_winners = length(na.omit(missed_payment[plot_time==1 & won_lottery==1])),
                 mean_missed_losers = mean(missed_payment[plot_time==1 & won_lottery==0],
                                           na.rm=T),
                 N_losers = length(na.omit(missed_payment[plot_time==1 & won_lottery==0])))

tax_bracket_data
```

```
##   prop_tax_bracket mean_missed_winners N_winners mean_missed_losers
## 1                1              0.1051       923            0.10584
## 2                2              0.1127       772            0.11443
## 3                3              0.0907       441            0.08794
## 4                4              0.1029       311            0.09265
## 5                5              0.0000        22            0.14286
##   N_losers
## 1      907
## 2      804
```

```
## 3      398
## 4      313
## 5       21
```

```r
chisq.test(tax_bracket_data[,c(3,5)])
```

```
##
##  Pearson's Chi-squared test
##
## data:  tax_bracket_data[, c(3, 5)]
## X-squared = 2.886, df = 4, p-value = 0.5772
```

Mechanism 1B.2: ***Behavioral/'habit' effects.***

TABLE 8. NATURAL EXPERIMENT. Habit effects. Winners vs. non-winners: heterogeneous treatment effects by time since winning (heterogeneous effects; 1, 2 and 3 years).

```r
naturalex_gtp$year_since_win[naturalex_gtp$tax=="Property" &
                          naturalex_gtp$plot_time > 3 & naturalex_gtp$plot_time < 7]  <- 2
naturalex_gtp$year_since_win[naturalex_gtp$tax=="Property" &
                          naturalex_gtp$plot_time > 6 & naturalex_gtp$plot_time < 10]  <- 3
naturalex_gtp$year_since_win[naturalex_gtp$tax=="Property" &
                          naturalex_gtp$plot_time > 9 & naturalex_gtp$plot_time < 13]  <- 4

naturalex_gtp$year_since_win[(naturalex_gtp$tax=="Sewage" | naturalex_gtp$tax=="Head") &
                          naturalex_gtp$plot_time > 6 & naturalex_gtp$plot_time < 13]  <- 2
naturalex_gtp$year_since_win[(naturalex_gtp$tax=="Sewage" | naturalex_gtp$tax=="Head") &
                          naturalex_gtp$plot_time > 12 & naturalex_gtp$plot_time < 19] <- 3
naturalex_gtp$year_since_win[(naturalex_gtp$tax=="Sewage" | naturalex_gtp$tax=="Head") &
                          naturalex_gtp$plot_time > 18 & naturalex_gtp$plot_time < 25]  <- 4


comp.eff(naturalex_gtp[naturalex_gtp$year_since_win==2,]$missed_payment,
         naturalex_gtp[naturalex_gtp$year_since_win==2,]$won_lottery,
         naturalex_gtp[naturalex_gtp$year_since_win==3,]$missed_payment,
         naturalex_gtp[naturalex_gtp$year_since_win==3,]$won_lottery)
```

```
##            DM       SE
## DM 1 0.002196 0.003928
## DM 2 0.001752 0.004475

## Diff in effects              SE
##       0.0004442       0.0059547
```

```r
comp.eff(naturalex_gtp[naturalex_gtp$year_since_win==3,]$missed_payment,
         naturalex_gtp[naturalex_gtp$year_since_win==3,]$won_lottery,
         naturalex_gtp[naturalex_gtp$year_since_win==4,]$missed_payment,
         naturalex_gtp[naturalex_gtp$year_since_win==4,]$won_lottery)
```

```
##            DM       SE
## DM 1 0.001752 0.004475
## DM 2 0.007384 0.005126
```



```
## Diff in effects              SE
##       -0.005631        0.006805
```

Mechanisms for H1C

Mechanism 1C.1: ***Erroneous beliefs.***

Rewards vs. punishments. (Positive vs. negative incentives)

Hypothesis 2A.

TABLE 9. FIELD EXPERIMENT. Priming knowledge of punishment. Good and bad taxpayers, comparison of treatments 3 and 5 (pooled) vs. the placebo control group (treatment 0).

```r
# Creating pooled variable
fieldex$pooled_35_0 <- NA
fieldex$pooled_35_0[fieldex$treatment==0] <- 0
fieldex$pooled_35_0[fieldex$treatment==3] <- 1
fieldex$pooled_35_0[fieldex$treatment==5] <- 1


# Missed payment
with(fieldex, t.test(missed_payment_DiD ~ pooled_35_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_DiD by pooled_35_0
## t = 0.6746, df = 8472, p-value = 0.5
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01001  0.02051
## sample estimates:
## mean in group 0 mean in group 1
##         0.02707         0.02182
```

```r
# Web bill request
with(fieldex, t.test(web_bill_DiD ~ pooled_35_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  web_bill_DiD by pooled_35_0
## t = -0.2676, df = 8509, p-value = 0.789
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01492  0.01134
## sample estimates:
## mean in group 0 mean in group 1
##         0.07613         0.07792
```

```r
# Payments owed
with(fieldex, t.test(payments_owed_DiD ~ pooled_35_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  payments_owed_DiD by pooled_35_0
## t = -0.3257, df = 11757, p-value = 0.7446
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.04237  0.03029
## sample estimates:
## mean in group 0 mean in group 1
##          0.1844          0.1904
```

```r
# Current debt
with(fieldex, t.test(current_debt_DiD ~ pooled_35_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  current_debt_DiD by pooled_35_0
## t = -0.3088, df = 8674, p-value = 0.7575
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4208  3062
## sample estimates:
## mean in group 0 mean in group 1
##           -1925           -1352
```

```r
# Compliance (test conditional on results of previous tests)
with(fieldex, t.test(compliance_DiD ~ pooled_35_0))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_DiD by pooled_35_0
## t = 0.6624, df = 8569, p-value = 0.5077
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.009692  0.019585
## sample estimates:
## mean in group 0 mean in group 1
##         0.02552         0.02058
```

TABLE 10. FIELD EXPERIMENT: Comparison of effects for good and bad taxpayers: difference of the difference in means for the comparison of treatments 3 and 5 (priming knowledge of sanctions, pooled) vs. the placebo control group (treatment 0).

```r
# Missed payments
comp.eff(fieldex[fieldex$type=="eligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_35_0,
         fieldex[fieldex$type=="noneligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_35_0)
```

```
##             DM       SE
## DM 1  0.001792 0.006792
## DM 2 -0.011222 0.012603

## Diff in effects              SE
##         0.01301         0.01432
```

```r
# Web access
comp.eff(fieldex[fieldex$type=="eligible",]$web_bill_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_35_0,
         fieldex[fieldex$type=="noneligible",]$web_bill_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_35_0)
```

```
##             DM      SE
## DM 1 0.0004106 0.01096
## DM 2 0.0029640 0.00830

## Diff in effects              SE
##       -0.002553        0.013751
```

We also assess one possible mechanism:

Mechanism 2A.1: ***Beliefs about probability of punishment***

We are especially interested in comparing the size of the effects of manipulating negative and positive incentives to pay taxes, though we are agnostic about the direction of the effect:

Hypothesis 2B.

TABLE 11. FIELD EXPERIMENT. Positive vs negative incentives. Good and bad taxpayers, comparison of treatments 1, 2 and 4 (positive incentives, pooled) vs 3 and 5 (negative incentives, pooled). Test using compliance conditional on significant effects for missed payment, number of payments owed or total debt.

```r
fieldex$pooled_124_35 <- NA
fieldex$pooled_124_35[fieldex$treatment==1] <- 1
fieldex$pooled_124_35[fieldex$treatment==2] <- 1
fieldex$pooled_124_35[fieldex$treatment==4] <- 1
fieldex$pooled_124_35[fieldex$treatment==3] <- 0
fieldex$pooled_124_35[fieldex$treatment==5] <- 0


# Missed payment
with(fieldex, t.test(missed_payment_DiD ~ pooled_124_35))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_DiD by pooled_124_35
## t = -0.8012, df = 17108, p-value = 0.423
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.017413  0.007308
## sample estimates:
## mean in group 0 mean in group 1
##         0.02182         0.02687
```

```
# Web bill request
with(fieldex, t.test(web_bill_DiD ~ pooled_124_35))
```

```
##
##  Welch Two Sample t-test
##
## data:  web_bill_DiD by pooled_124_35
## t = -0.2909, df = 17392, p-value = 0.7711
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01199  0.00889
## sample estimates:
## mean in group 0 mean in group 1
##         0.07792         0.07947
```

```
# Payments owed
 with(fieldex, t.test(payments_owed_DiD ~ pooled_124_35))
```

```
##
##  Welch Two Sample t-test
##
## data:  payments_owed_DiD by pooled_124_35
## t = 1.879, df = 17206, p-value = 0.06025
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001449  0.068687
## sample estimates:
## mean in group 0 mean in group 1
##          0.1904          0.1568
```

```
# Current debt
with(fieldex, t.test(current_debt_DiD ~ pooled_124_35))
```

```
##
##  Welch Two Sample t-test
##
## data:  current_debt_DiD by pooled_124_35
## t = -1.776, df = 17405, p-value = 0.07573
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```



```
##  -5653.9   278.4
## sample estimates:
## mean in group 0 mean in group 1
##           -1352            1336
```

```
# Compliance debt
with(fieldex, t.test(compliance_DiD ~ pooled_124_35))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_DiD by pooled_124_35
## t = -1.007, df = 17370, p-value = 0.3141
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.017992  0.005781
## sample estimates:
## mean in group 0 mean in group 1
##         0.02058         0.02668
```

TABLE 12. FIELD EXPERIMENT. Comparison of effects of positive vs negative incentives for good and bad taxpayers: difference of the difference in means for the comparison of treatments 1, 2 and 4 (pooled) and 3 and 5 (pooled).

```
# Missed payments
comp.eff(fieldex[fieldex$type=="eligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_124_35,
         fieldex[fieldex$type=="noneligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_124_35)
```

```
##            DM       SE
## DM 1 0.003489 0.005451
## DM 2 0.006368 0.010235
```

```
## Diff in effects              SE
##        -0.00288         0.01160
```

```
# Web access
comp.eff(fieldex[fieldex$type=="eligible",]$web_bill_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_124_35,
         fieldex[fieldex$type=="noneligible",]$web_bill_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_124_35)
```

```
##             DM       SE
## DM 1  0.007406 0.008713
## DM 2 -0.002938 0.006602
```

```
## Diff in effects              SE
##         0.01034         0.01093
```

Hypothesis 2C: *Marginal taxpayers I.*

TABLE 13. FIELD EXPERIMENT. Marginal taxpayers. Good taxpayers. Heterogeneous effects, taxpayers at risk. Comparison of treatment effect of 1, 2 and 4 (pooled) vs control (A-Information about the tax lottery), on one test and 3 and 5 (pooled) vs control on another (B-Information about sanctions).}

```
# Identifying good taxpayers with history of debt
# Ever owed a bill since March 2009?
names(fieldex[,c(26:40,7)])
```

```
##  [1] "adeudadas_2009_MAR" "adeudadas_2009_JUL" "adeudadas_2009_NOV"
##  [4] "adeudadas_2010_MAR" "adeudadas_2010_JUL" "adeudadas_2010_NOV"
##  [7] "adeudadas_2011_MAR" "adeudadas_2011_JUL" "adeudadas_2011_NOV"
## [10] "adeudadas_2012_MAR" "adeudadas_2012_JUL" "adeudadas_2012_NOV"
## [13] "adeudadas_2013_MAR" "adeudadas_2013_JUL" "adeudadas_2013_NOV"
## [16] "adeudadas_2014_MAR"
```

```
sum_bills_owed <- apply(fieldex[,c(26:40,7)], 1, sum)

fieldex$goodtp_at_risk <- ifelse(sum_bills_owed>0, 1, 0)
table(fieldex$goodtp_at_risk[fieldex$type=="eligible"])
```

```
##
##    0    1
## 7164 2297
```

```
### A. Information on the lottery - Heterogeneous effects for taxpayers at risk

# Missed payments
comp.eff(fieldex$missed_payment_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$pooled_124_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$missed_payment_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0],
        fieldex$pooled_124_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0])
```

```
##            DM      SE
## DM 1 0.013129 0.01708
## DM 2 0.002724 0.00696

## Diff in effects              SE
##         0.01041         0.01845
```

```
# Web access
comp.eff(fieldex$web_bill_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$pooled_124_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$web_bill_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0],
        fieldex$pooled_124_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0])
```

```
##            DM      SE
## DM 1 0.028336 0.02132
## DM 2 0.001317 0.01277
```



```
## Diff in effects              SE
##         0.02702         0.02485
```

```
### B. Information on sanctions - Heterogeneous effects for taxpayers at risk

# Missed payments
comp.eff(fieldex$missed_payment_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$pooled_35_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$missed_payment_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0],
        fieldex$pooled_35_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0])
```

```
##             DM       SE
## DM 1  0.012920 0.017062
## DM 2 -0.001441 0.007071

## Diff in effects              SE
##         0.01436         0.01847
```

```
# Web access
comp.eff(fieldex$web_bill_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$pooled_35_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==1],
        fieldex$web_bill_DiD[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0],
        fieldex$pooled_35_0[fieldex$type=="eligible" & fieldex$goodtp_at_risk==0])
```

```
##             DM      SE
## DM 1  0.008871 0.02113
## DM 2 -0.002059 0.01279

## Diff in effects              SE
##         0.01093         0.02470
```

TABLE 14. FIELD EXPERIMENT. Marginal taxpayers. Bad taxpayers. Heterogeneous effects, salvageable taxpayers. Comparison of treatment effect of 1, 2 and 4 (pooled) vs control (A-Information about the tax lottery), on one test and 3 and 5 (pooled) vs control on another (B-Information about sanctions). Test using compliance conditional on significant effects for missed payment, number of payments owed or total debt.

```
# Identifying bad taxpayers not too in debt

# Ever owed a bill since March 2009?
names(fieldex)[7]
```

```
## [1] "adeudadas_2014_MAR"
```

```
fieldex$salvageable_btp <- ifelse(fieldex$adeudadas_2014_MAR>3, 0, 1)

### A. Information on the lottery - Heterogeneous effects for taxpayers at risk

# Missed payments
comp.eff(fieldex$missed_payment_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
        fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
        fieldex$missed_payment_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
        fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##             DM      SE
## DM 1 -0.003946 0.01620
## DM 2 -0.006449 0.01261

## Diff in effects              SE
##        0.002503        0.020529
```

```r
# Web access
comp.eff(fieldex$web_bill_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$web_bill_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##             DM       SE
## DM 1  0.002772 0.010583
## DM 2 -0.007161 0.008238

## Diff in effects              SE
##        0.009933        0.013411
```

```r
# Nr of payments owed
comp.eff(fieldex$payments_owed_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$payments_owed_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##             DM      SE
## DM 1 -0.009227 0.01682
## DM 2 -0.175959 0.10960

## Diff in effects              SE
##          0.1667          0.1109
```

```r
# Compliance (conditional)
comp.eff(fieldex$compliance_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$compliance_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_124_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##             DM       SE
## DM 1 -0.002700 0.015822
## DM 2 -0.001144 0.005121

## Diff in effects              SE
##       -0.001557        0.016630
```

```r
### B. Information on sanctions - Heterogeneous effects for taxpayers at risk

# Missed payments
comp.eff(fieldex$missed_payment_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$missed_payment_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```



```
##             DM      SE
## DM 1 -0.011346 0.01621
## DM 2 -0.005506 0.01291

## Diff in effects              SE
##        -0.00584         0.02072
```

```r
# Web access
comp.eff(fieldex$web_bill_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$web_bill_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##            DM       SE
## DM 1  0.01219 0.010779
## DM 2 -0.01996 0.008121

## Diff in effects              SE
##         0.03216         0.01350
```

```r
# Nr of payments owed
comp.eff(fieldex$payments_owed_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$payments_owed_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##            DM      SE
## DM 1  0.03338 0.02670
## DM 2 -0.07526 0.09817

## Diff in effects              SE
##          0.1086          0.1017
```

```r
# Compliance (conditional)
comp.eff(fieldex$compliance_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==1],
         fieldex$compliance_DiD[fieldex$type=="noneligible" & fieldex$salvageable_btp==0],
         fieldex$pooled_35_0[fieldex$type=="noneligible" & fieldex$salvageable_btp==0])
```

```
##              DM       SE
## DM 1 -0.0127398 0.015810
## DM 2  0.0009157 0.005252

## Diff in effects              SE
##        -0.01366         0.01666
```

Hypothesis 2D: *Marginal taxpayers II.*

Individual vs. social incentives.

Hypothesis 3A: *Social benefits.*

TABLE 15. FIELD EXPERIMENT. Good and bad taxpayers. Social vs individual rewards. Comparison of treatments 1, 2 (pooled) vs 4. Test using compliance conditional on significant effects for missed payment, number of payments owed or total debt.

```
fieldex$pooled_12_4 <- NA
fieldex$pooled_12_4[fieldex$treatment==1] <- 1
fieldex$pooled_12_4[fieldex$treatment==2] <- 1
fieldex$pooled_12_4[fieldex$treatment==4] <- 0

# Missed payment
with(fieldex, t.test(missed_payment_DiD ~ pooled_12_4))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_DiD by pooled_12_4
## t = -0.4156, df = 8623, p-value = 0.6777
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02119  0.01378
## sample estimates:
## mean in group 0 mean in group 1
##         0.02499         0.02870
```

```
# Web bill request
with(fieldex, t.test(web_bill_DiD ~ pooled_12_4))
```

```
##
##  Welch Two Sample t-test
##
## data:  web_bill_DiD by pooled_12_4
## t = -0.6687, df = 8780, p-value = 0.5037
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.019616  0.009637
## sample estimates:
## mean in group 0 mean in group 1
##         0.07694         0.08193
```

```
# Payments owed
with(fieldex, t.test(payments_owed_DiD ~ pooled_12_4))
```

```
##
##  Welch Two Sample t-test
##
```



```
## data:  payments_owed_DiD by pooled_12_4
## t = -1.499, df = 7244, p-value = 0.134
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08363  0.01116
## sample estimates:
## mean in group 0 mean in group 1
##          0.1384          0.1746
```

```
# Current debt
with(fieldex, t.test(current_debt_DiD ~ pooled_12_4))
```

```
##
##  Welch Two Sample t-test
##
## data:  current_debt_DiD by pooled_12_4
## t = -0.7607, df = 8781, p-value = 0.4468
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5836  2573
## sample estimates:
## mean in group 0 mean in group 1
##           508.1          2139.6
```

```
# Compliance (conditional)
with(fieldex, t.test(compliance_DiD ~ pooled_12_4))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_DiD by pooled_12_4
## t = -0.4933, df = 8763, p-value = 0.6218
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02114  0.01264
## sample estimates:
## mean in group 0 mean in group 1
##         0.02453         0.02878
```

TABLE 16. FIELD EXPERIMENT. Social (4) vs individual rewards (1 and 2, pooled), comparison of effect between good and bad taxpayers.

```
# Missed payments
comp.eff(fieldex[fieldex$type=="eligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_12_4,
         fieldex[fieldex$type=="noneligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_12_4)
```

```
##              DM      SE
## DM 1 -0.005270 0.00758
## DM 2  0.009944 0.01452
```

```
## Diff in effects              SE
##        -0.01521         0.01638
```

```r
# Web access
comp.eff(fieldex[fieldex$type=="eligible",]$web_bill_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_12_4,
         fieldex[fieldex$type=="noneligible",]$web_bill_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_12_4)
```

```
##            DM       SE
## DM 1 0.001810 0.012327
## DM 2 0.007491 0.009147
```

```
## Diff in effects              SE
##       -0.005681        0.015350
```

Hypothesis 3B: *Social sanctions.*

TABLE 17. FIELD EXPERIMENT. Good and bad taxpayers. Social vs individual sanctions. Comparison of treatments 3 vs 5. Test using compliance conditional on significant effects for missed payment, number of payments owed or total debt.

```r
fieldex$pooled_3_5 <- NA
fieldex$pooled_3_5[fieldex$treatment==3] <- 0
fieldex$pooled_3_5[fieldex$treatment==5] <- 1


# Missed payment
with(fieldex, t.test(missed_payment_DiD ~ pooled_3_5))
```

```
##
##  Welch Two Sample t-test
##
## data:  missed_payment_DiD by pooled_3_5
## t = 0.8602, df = 8476, p-value = 0.3897
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.00981  0.02515
## sample estimates:
## mean in group 0 mean in group 1
##         0.02565         0.01798
```

```r
# Web bill request
with(fieldex, t.test(web_bill_DiD ~ pooled_3_5))
```

```
##
##  Welch Two Sample t-test
##
## data:  web_bill_DiD by pooled_3_5
## t = -0.4795, df = 8609, p-value = 0.6316
## alternative hypothesis: true difference in means is not equal to 0
```



```
## 95 percent confidence interval:
##  -0.01854  0.01125
## sample estimates:
## mean in group 0 mean in group 1
##         0.07610         0.07975
```

```r
# Payments owed
with(fieldex, t.test(payments_owed_DiD ~ pooled_3_5))
```

```
##
##  Welch Two Sample t-test
##
## data:  payments_owed_DiD by pooled_3_5
## t = -0.2992, df = 8092, p-value = 0.7648
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.05988  0.04402
## sample estimates:
## mean in group 0 mean in group 1
##          0.1864          0.1944
```

```r
# Current debt
with(fieldex, t.test(current_debt_DiD ~ pooled_3_5))
```

```
##
##  Welch Two Sample t-test
##
## data:  current_debt_DiD by pooled_3_5
## t = -0.5044, df = 8619, p-value = 0.614
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5261  3108
## sample estimates:
## mean in group 0 mean in group 1
##         -1889.1          -812.4
```

```r
# Compliance
with(fieldex, t.test(compliance_DiD ~ pooled_3_5))
```

```
##
##  Welch Two Sample t-test
##
## data:  compliance_DiD by pooled_3_5
## t = 0.503, df = 8600, p-value = 0.615
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01244  0.02102
## sample estimates:
## mean in group 0 mean in group 1
##         0.02272         0.01842
```

TABLE 18. FIELD EXPERIMENT. Social vs individual sanctions. Comparison of effects for good and bad taxpayers. Comparison of treatments 3 vs 5. Test using compliance conditional on significant effects for missed payment, number of payments owed or total debt.

```r
# Missed payments
comp.eff(fieldex[fieldex$type=="eligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_3_5,
         fieldex[fieldex$type=="noneligible",]$missed_payment_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_3_5)
```

```
##             DM       SE
## DM 1 -0.004118 0.007833
## DM 2 -0.009563 0.014434
```

```
## Diff in effects              SE
##        0.005445        0.016422
```

```r
# Web access
comp.eff(fieldex[fieldex$type=="eligible",]$web_bill_DiD,
         fieldex[fieldex$type=="eligible",]$pooled_3_5,
         fieldex[fieldex$type=="noneligible",]$web_bill_DiD,
         fieldex[fieldex$type=="noneligible",]$pooled_3_5)
```

```
##            DM       SE
## DM 1 0.005127 0.012306
## DM 2 0.002296 0.009526
```

```
## Diff in effects              SE
##        0.002831        0.015562
```

G.3 Second amendment to PAP (registered December 13, 2014)

Second Amendment to Pre-Analysis Plan:

Analysis of Survey Data

Positive vs. Negative Incentives for Compliance: Evaluating a Randomized Tax Holiday

December 13, 2014

Contents

Description of the document

This document is a second amendment to our original pre-analysis plan, dated July 23, 2014, which is registered at the Experiments in Governance and Politics registration page as well as the study registry of the American Economic Association.[1] We registered our first amendment, dated October 19, 2014, in advance of analyzing administrative (tax payment) data. This second amendment concerns analysis of household survey data, which we have received from our survey firm since filing our first amendment. We register this document in advance of conducting experimental tests using the survey data (though after analyzing descriptive quantities such as unconditional means and standard deviations). We encountered substantial logistical difficulties in implementing aspects of our household survey. Fortunately, these difficulties do not affect our analysis of survey experiments, though they do affect inferences about the effects of our field and natural experiments using the household survey (rather than administrative/tax payment) data.

This document describes our amended plan for analyzing the survey data in light of these considerations. First, we describe the challenges we experienced with respect to survey implementation. Next, we discuss modifications to the analysis plan in light of these challenges. Finally, we present code for a mock analysis of the survey data.

Challenges to survey implementation

We hired a reputable Uruguayan survey firm, CIFRA, to implement our household surveys. We chose CIFRA from among four competing bids due to the reputation of the firm for producing quality work, as well as the lesser cost and shorter promised timeline for the fieldwork relative to other firms.

Our agreement with CIFRA called for face-to-face interviews of 6,000 taxpayers, selected from the total study group of 33,429 households randomly selected from the population of payers of the property tax (Contribución Inmobiliaria) in Montevideo. Fieldwork began July 16, 2014. CIFRA promised to implement our survey in a period of roughly four-five weeks, so that the fieldwork would be completed by the end of August or beginning of September. The timing of the surveys was critical, as we planned to assess the effects on attitudes and behaviors of citizens who received our randomized informational treatments (that is, flyers with different messages about tax payments), which were mailed in phases beginning June 27, 2014 and ending July 16, 2014.

The taxpayer accounts we obtained from the municipality (i.e., our study group, which is in turn a random sample of taxpayers in Montevideo) contained addresses associated to each account. The survey firm intended to locate physical taxpayers using these addresses. There were several considerations regarding the accounts obtained

[1]See registered design 84 [initial date: 20140723] at the EGAP registration page.

from the municipality, which we describe on pages 21-23 of our original pre-analysis plan. First, they contained both physical persons as well as firms ("legal persons" or *personas jurídicas*), since the municipality has no way of identifying ex-ante who is a physical person.[2] Second, the mailing address associated with an account may or may not correspond to the physical address of the property on which taxes are paid.[3] For this reason, it was important for us to work with CIFRA to screen our data and identify households consisting of physical taxpayers.

To accomplish our objective, and in light of several logistical difficulties we encountered, we provided data to CIFRA in two phases:

1. First, as described in our original pre-analysis plan, we randomly sampled 8,000 taxpayer accounts from the study group for our informational field experiment and provided these records to CIFRA, with instructions to seek 6,000 completed surveys from among these. Our procedure required CIFRA to visit each household a minimum of three times, one of which visits had to take place during the weekend. In case of direct refusal or inability of the firm to locate a taxpayer, the firm was to replace that case with another from the original list we provided.[4]

   As of July 25, nine days after initiating fieldwork, the firm had only visited 300 addresses (and only completed 47 interviews). The management of the firm informed us that they had great difficulty locating households, given the primitive mapping technologies they used, and they could not readily organize the fieldwork so as to minimize their costs and expedite the fieldwork. This slow rhythm of work obviously appeared to compromise the firm's ability to complete the surveys in the agreed period.

2. We therefore reorganized the data by geocoding all of the addresses using QGIS together with further municipal data on the location of properties. To do this, we obtained from the municipality the number associated with the physical address of each property on which tax is paid for each account (the *padrón*) and matched this to the official map that assigns the number/*padrón* to a single lot in the city of Montevideo. We then eliminated (for purposes of the household survey) those accounts for which we could not find a valid geolocated address; this reduced the available accounts to 20,866 in the field experiment (from 28,600) and to 4,947 (from 5,129) in the natural experiment (considering only properties subject to property tax or Contribución Inmobiliaria).

[2]Indeed, this is why lottery winners must present themselves to the municipality to establish that they are physical persons and thus eligible for the tax holiday.

[3]In many cases, this did not matter because our informational treatments were mailed to the address associated with the account, and we tracked tax compliance (using administrative data) and attitudes and behaviors (using survey data) using the same address and account data. Thus, we could readily measure the effect of our flyers on a person living at X address who pays taxes on property Y, whether or not X=Y. However, in some cases a renter living in property X might receive the mailer, but the payment of taxes could be made by an absentee landlord (in other cases, the landlord might receive the bill for property X at a different address, and this would be captured in the account data). In other cases, an administrative or real estate agent might receive and pay the bills.

[4]With assistance from us, the firm screened this list of 8,000 to eliminate duplicate entries (e.g. houses that receive more than one bill because they own more than one property) and screen out invalid addresses, leaving a list of 7,107 accounts from which they should select cases for interviews.

Next, we grouped these valid addresses into 931 census tracts. Finally, we randomly sampled 697 census tracts (weighted by the number of households in our study group located in each tract) and sent this first batch of tracts to CIFRA; this batch of census tracts included 25,998 individual addresses. Our hope was that this would dramatically increase the speed with which CIFRA could locate taxpayers in our sample.

Timeline of survey data collection

Unfortunately, our efforts did not have the intended effect. Table 1 shows the progress of the fieldwork conducted by CIFRA. Over the course of eighteen weeks, the firm averaged around 132 successful interviews per week. As of December 11, 2014, CIFRA had completed 2,349 interviews.

Table 0.1: Cumulative progress of survey fieldwork by CIFRA

REPORT DATE	**INTERVIEWS COMPLETED**	HOUSEHOLDS VISITED	Reason for failure			Number of Enumerators
			Refusal	Not physical	Other*	
July 25, 2014	47	300	94	84	75	29
August 5, 2014	240	1,473	449	221	563	53
August 13, 2014	441	2,734	845	350	1,098	49
August 19, 2014	609	3,932	1,088	694	1,541	48
August 26, 2014	821	4,924	1,330	754	2,019	45
September 2, 2014	1,008	6,978	1,876	1,010	3,084	40
September 9, 2014	1,124	8,070	2,283	1,112	3,551	38
September 16, 2014	1,238	9,492	2,627	1,414	4,213	36
September 23, 2014	1,405	11,005	2,998	1,647	4,959	40
September 30, 2014	1,536	12,325	3,426	1,716	5,647	43
October 7, 2014	1,664	13,387	3,784	1,746	6,193	30
October 14, 2014	1,723	14,614	4,024	1,843	7,024	25
October 21, 2014	1,812	15,531	4,326	1,957	7,436	36
October 28, 2014	1,912	16,168	4,429	2,047	7,748	32
November 4, 2014	1,980	16,939	4,719	2,129	8,111	35
November 11, 2014	2,109	17,811	4,938	2,292	8,472	30
November 18, 2014	2,194	18,446	5,120	2,343	8,789	25
November 25, 2014	2,283	18,886	5,202	2,375	9,026	28

* "Others" includes the following reasons for failure: address does not exist; residence is empty; nobody home (after three visits); deceased or repeated address in data set.

This slow rhythm of data collection can potentially compromise inferences from the survey data in two key ways.

1. First—especially for interviews that happened after the intended end of data collection at the end of August—it is perhaps implausible to think that flyers received several weeks or months prior to the interview would substantially impact citizens' responses to the eight questions we registered as main outcomes for our analysis of survey data, namely, (1) trust in municipal government; (2) trust in civil servants; (3) evaluation of the mayor; (4) fairness of municipal taxes in general; (5) fairness of the property tax; (6) attitudes towards tax amnesties;

(7) agreement that sometimes taxes are not worth paying; and (8) party vote intention (or change in vote intention from previous election).[5] Moreover, our survey instrument frequently included temporal references such as "in the last month, have you received a flyer from the municipal government?" which we constructed on the understanding that the survey would be completed in four-five weeks after the mailing of flyers; these references do not make sense for respondents interviewed in September and thereafter.

2. Second, and even more consequentially, power analyses suggested a requisite overall sample size of 6,000 to detect plausible effects of our informational treatments with 80% power, when stratifying between good taxpayers (N=3,000) and bad taxpayers (N=3,000) and estimating effects of individual treatments with respect to the control group. The sample size as of December 11, 2014 is 2,453, that is, less than half of this targeted sample size.

For both reasons, our power to detect effects of our informational treatments using survey data is substantially lower than anticipated by our pre-analysis plan. We do not think it is plausible that we will detect effects of our flyers on the attitudinal outcomes registered in our original plan.

Fortunately, we included in our survey instrument experimental primes that substantially replicate the informational treatments in the field experiment; and we asked several outcome questions that allow us to probe the effects of these primes on political attitudes. Thus, the survey can serve its intended function of measuring the effects of information about punishment for non-payment, or benefits for tax payment (e.g. the tax holiday lottery). We failed to describe the survey experiments in sufficient detail in our original pre-analysis plan and thus do so here.

Revisions to analysis plan

In light of the complications that arose during survey data collection, and our reliance on the survey experiment to draw inferences about the effects of informational treatments, we make the following amendments to our plan for analysis of survey data:

1. We consider the survey measures registered in our field and natural experiments (see page 30 of our original pre-analysis plan) to be secondary outcomes. Our primary outcomes for the field and natural experiments are the administrative measures of tax compliance, missed payments, and total debt, as well as (for the field experiment) our measure of Web bill access (see page 29 of our original plan). We will adjust statistical tests for multiple comparisons with respect to those primary outcomes measured through administrative data.

[5] See page 30 of our original pre-analysis plan.

2. We will focus the analysis of survey data on the survey experiments (see previous section and page 30-31 of our original plan). Note that these questions are not subject to the same issues regarding delays in the surveys, because randomization of treatments is achieved within the survey itself. Analysis of these questions does not give us purchase on the effects of our informational flyers but does serve some of the same objectives, in terms of giving us insight into how priming sanctions/punishments vs. rewards may influence attitudes.

3. Finally, we will use the survey data to gain insight into observational quantities such as the proportion of taxpayers in Montevideo who know about the existence of the tax holiday lottery, or the existence of tax amnesties as well as fines and punishment for non-payment. These analysis should be taken with some grain of salt as appropriate, given the substantial non-response documented in Table 1 (i.e., we may only project to the population of taxpayers in Montevideo under some strong assumptions). But these data will still be useful for assessing reasons why the natural or field experimental treatments did not have a larger effect on tax payment behavior (as measured through the field experiment).

Description of survey experiment

Our survey instrument included the following informational treatments; the order was randomized and, as we describe below, individuals were randomly assigned to receive certain of these prompts and not others.

- INDIVIDUAL PUNISHMENT: "The municipal government of Montevideo applies fines and charges to those who do not pay their taxes on time. These punishments can be very costly for whoever does not pay their taxes."[6]
- SOCIAL PUNISHMENT: "The municipal government of Montevideo applies fines and charges to those who do not pay their taxes on time. Fines and charges are a punishment for those who do not contribute to the construction of a city that is more just and better for all."[7]
- INDIVIDUAL BENEFIT: "The municipal government of Montevideo raffles, in conjunction with the National Lottery, the exoneration of municipal taxes for one year. This policy individually rewards those who are up to date on their tax payments."[8]
- SOCIAL BENEFIT: "The municipal government of Montevideo raffles, in conjunction with the National Lottery, the exoneration of municipal taxes for one year among people who are up to date on their tax payments. The municipal government conducts this lottery to recognize good taxpayers for their contribution to the construction of a city that is more just and better for all.[9]
- "DISCRETIONAL" BENEFIT: "At times the municipal government of Montevideo chooses people who are up to date on their municipal tax payments, to grant them the exoneration of municipal taxes for one year."[10]

[6] In Spanish, "La Intendencia de Montevideo aplica multas y recargos a quienes no pagan sus impuestos en fecha. Estas sanciones pueden ser muy costosas para quien no paga sus impuestos."

[7] La Intendencia de Montevideo aplica multas y recargos a quienes no pagan sus impuestos en fecha. Las multas y recargos son una sanci?n para quienes no contribuyen a la construcci?n de una ciudad m?s justa y mejor para todos y todas.

[8] La Intendencia de Montevideo sortea junto a la Loter?a Nacional la exoneraci?n de los tributos municipales por un a?o. Esta pol?tica premia individualmente a quienes est?n al d?a con sus impuestos.

Hypotheses: survey experiment

We will test several hypotheses about the effects of these informational treatments on several of the attitudinal outcome measures registered in our original pre-analysis plan (see p. 30-31). Respondents were asked for their degree of agreement on a 0-10 scale with the following statements:

1. "People only pay their taxes on time when there are substantial fines and charges" (survey question M.1.1);[11]
2. "In Montevideo, punishments don't apply to the privileged" (question M.1.4)[12]; and
3. "Fines and charges for bad taxpayers are pointless" (question M.1.5).[13]
4. "Policies that reward good taxpayers are a waste of money" (question S.1.1)[14];
5. "In Montevideo, benefits for good taxpayers go to the same people as always (question S.1.4)."[15]
6. "In general, the municipal government does a good job" (questions M.1.3 and S.1.2)[16];
7. "In Montevideo, it is worth it to be up to date on ones taxes" (question M.1.2 and S.1.3)"[17]

   as well as their response to a question with four ordered response categories:

8. "How would you classify the taxes that the municipal government charges, in general: very just, fairly just, a little just, or not just at all?" (questions M.1.6 and S.1.5)[18]

We reiterate and slightly modify hypotheses outlined in our original pre-analysis plan (see e.g. Table 7.3). In particular, we are interested in the testing the alternative of a treatment effect against the following null hypotheses:

[9] La Intendencia de Montevideo sortea junto a la Loter?a Nacional la exoneraci?n de los impuestos por un a?o entre las personas que est?n al d?a. La Intendencia realiza este sorteo para reconocer a los buenos pagadores por su contribuci?n a la construcci?n de una ciudad m?s justa y mejor para todos/as.

[10] A veces la Intendencia de Montevideo elige personas que est?n al d?a con el pago de impuestos municipales, para otorgarles un a?o de exoneraci?n de pago de estos impuestos.

[11] In Spanish, "La gente paga sus impuestos en fecha solo cuando las multas y los recargos son importantes."

[12] In Spanish, "En Montevideo, a los privilegiados, no se les aplican los castigos."

[13] In Spanish, "Las multas y recargos a malos pagadores no sirven para nada."

[14] In Spanish, "Las políticas que premian a los buenos pagadores son un derroche de dinero."

[15] In Spanish, "En Montevideo, los beneficios para buenos pagadores se los llevan los mismos de siempre."

[16] In Spanish, "En general, la Intendencia hace un buen trabajo."

[17] In Spanish, "En Montevideo, vale la pena estar al día con los impuestos."

[18] In Spanish, "Cómo clasificaría los impuestos que cobra la Intendencia de Montevideo en general: muy justos, bastante justos, poco justos o nada justos?"

1. PUNISHMENTS VERSUS BENEFITS: There is no difference in the response of taxpayers to the benefit treatments vs. the punishment treatments, for the following outcomes[19]:

   - “In general, the municipal government does a good job" (question S.1.2 and question M.1.3)[20];

     Here, we expect benefits to have a larger positive effect than punishments (where the effect may be negative), and thus we expect benefits - punishments to be positive; we will conduct a one-tailed test consistent with this alternative hypothesis.

   - “In Montevideo, it is worth it to be up to date on ones taxes" (question M.1.2 and question S.1.3);

     Here, we have no strong directional hypothesis: either punishments or benefits may have the larger effect on this variable. We will conduct a two-tailed test.

   - “How would you classify the taxes that the municipal government charges, in general: very just, fairly just, a little just, or not just at all?" (questions M.1.6 and S.1.5); and

     Here, we again expect benefits to have the larger positive effect (i.e. greater perceptions that taxes are just), leading to a one-tailed test.

2. DISCRETIONARY VS. NON-DISCRETIONARY/RANDOM BENEFITS: There is no difference in the response of taxpayers to the discretionary benefit vs. individual/social benefits conditions in which we mention the lottery, for the following outcomes[21]:

   - “In Montevideo, it is worth it to be up to date on ones taxes" (question S.1.3);

     Here, we suspect the non-discretionary benefits (knowledge of lottery) will have the larger positive effect, as we only state that rewards are conditional on being up to date (being a "good taxpayer") in the non-discretionary/random benefits conditions. We will conduct a one-tailed test.

   - “Policies that reward good taxpayers are a waste of money" (question S.1.1);

     We conduct a two-tailed test, because taxpayers might dislike discretion, but they might also dislike lotteries if they perceive them to be ineffective.

   - “In general, the municipal government does a good job" (question S.1.2);

     We conduct a two-tailed test, because taxpayers might like "universal particularism" (Monestier et al. 2014) but they might also like the lottery policy.

   - “In Montevideo, benefits for good taxpayers always go to the same people" (question S.1.4); and

[19] This corresponds to Hypothesis 2A in the original plan (see Table 7.3), and we add one outcome not registered there

[20] Question S.1.2. is the same as M.1.3; however, the former follows the prompt about benefits whereas the latter follows the prompts about punishments. A similar comment applies to other outcomes for the PUNISHMENT VERSUS BENEFITS hypotheses.

[21] This corresponds to Mechanisms IA.1 and IA.2 in the original plan (see Table 7.3).

     We conduct a one-tailed test under the alternative hypothesis that the discretionary benefits treatment increases agreement with this statement more than the lottery treatment.

   - “How would you classify the taxes that the municipal government charges, in general: very just, fairly just, a little just, or not just at all?" (question S.1.5).

     We conduct a one-tailed test under the alternative hypothesis that the non-discretionary/lottery treatment increases perceptions that the system is just.

3. SOCIAL VERSUS INDIVIDUAL BENEFITS: There is no difference in the response of taxpayers to the social benefit and the individual benefit primes, for the following outcomes[22]:

   - “Policies that reward good taxpayers are a waste of money" (question S.1.1); and

     We conduct a two-tailed test, because who the heck knows? (Certainly not us. That’s why we are doing the research!).

   - “In Montevideo, it is worth it to be up to date on ones taxes" (question S.1.3).

     The treatment effects depend on whether people are rational individual maximizers or social beings. We are agnostic. Two-tailed test.

Adjustments for multiple comparisons

Our original pre-analysis plan did not specify corrections for the survey experimental analysis in detail, so we do so here. As per our original plan, we will present both nominal p-values and corrected p-values, using a false discovery rate (FDR) correction to control the Type-1 error rate. For the survey experiments, as for our natural and field experiments, we will control the FDR at level 0.05.

For a given randomization with m (null) hypotheses and m associated p-values, we order the realized nominal p-values from smallest to largest, $p_{(1)}, \leq p_{(2)} \leq \ldots \leq p_{(m)}$. Let

$$k \text{ be the largest } i \text{ for which } p_{(i)} \leq \tfrac{i}{m} 0.05.$$

Then, we reject all $H_{(i)}$ for $i = 1, 2, \ldots, k$, where $H_{(i)}$ is the null hypothesis corresponding to $p_{(i)}$.[23]

For comparison, we will also present strict Bonferroni corrections, i.e., for each hypothesis $H_{(i)}$, we reject the null at the adjusted 0.05 level if $p_{(i)} \leq \frac{0.05}{m}$. This correction will lead to the most conservative inference for each

[22] Though we specified a comparison of social vs. individual benefits for the field experiment in our original plan, we omitted this comparison of survey experimental treatments; we add that here.

[23] For a description of this procedure, see Yoav Benjamini and Yosef Hochberg. 1995. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological).* 57 (1): 289-300.

individual pairwise comparison. Our rejection rule, however, will require controlling the overall false discovery rate.

How large is m under our study design? This differs for the survey experiment, we have three independent or nearly independent set of tests corresponding to different randomizations. For each set of tests, there are two treatment conditions. The number of outcome variables/test statistics (and associated p-values) in each set is as follows:

- PUNISHMENT VS. BENEFITS: $m = 3$
- DISCRETIONARY BENEFITS VS. NON-DISCRETIONARY BENEFITS/LOTTERY: $m = 5$
- SOCIAL VS. INDIVIDUAL BENEFITS: $m = 2$.

Testing for order effects

Each survey respondent was exposed to two of the informational treatments (though in different orders, as explained next). For purposes of our main analysis, we pool treatments, e.g. for H1 we compare all exposures to the Social Benefit prime to all exposures to the Individual Benefit prime; and for H2 we pool the social and individual benefit treatments and compare them to the pooled social and individual punishment conditions.

To allow us to test for order effects, we varied the placement of the treatments across four versions of the questionnaires, as shown in Table 2. The "Discretional Benefit" treatment only appears in one questionnaire (D) and appears before the other informational prime in that questionnaire; thus, it cannot be subject to order effects. To test whether the other three primes are subject to order effects, we will assess whether responses to the individual benefit, individual punishment, and social punishment conditions are statistically identical when these treatments appear as the first informational treatment vs. as the second informational treatment in the questionnaire. Note that each of these three treatments appears once in the first position on a questionnaire version and once in the second position. Thus, even if there are order effects, they should not bias conclusions about the effects of these treatments compared to each other.

Table 0.2: Survey experimental treatments: distribution across four versions of questionnaire

	Version of Questionnaire			
	A	B	C	D
First treatment	Individual Punishment	Social Punishment	Individual Benefit	Discretional Benefit
Second treatment	Social Benefit	Individual Benefit	Individual Punishment	Social Punishment

Revision to Table 7.3 in our original pre-analysis plan

In light of the modifications discussed in this document, we amend the set of hypotheses, outcomes and tests described in Table 7.3 of our original pre-analysis plan, mainly in order to exclude tests in our field experiment using survey data. Here, strikethroughs indicate tests that we will not longer undertake. In a few cases, we add outcomes for the survey experiment we had forgotten to register in the original plan. We also update this table to reflect changes to the analysis of the administrative/tax payment data registered in our first amendment to the pre-analysis plan (which we filed before analyzing the administrative data).

Code for mock analysis of survey data

In what follows, we use a dataset containing responses from 2,349 survey respondents, provided to us by CIFRA on December 11, 2013. We will replace this with the final dataset containing 2,400 respondents when we receive it from the survey firm.

Here, after cleaning some variables where we detected coding errors (e.g. as indicated by extreme/impossible outliers), we shuffle treatment labels, that is, we assign each respondent at random to a version of the questionnaire (and thus of survey experimental prompts) that does not necessarily reflect the real version/prompts that the respondent received. This allows us to run code for our survey experimental tests without seeing the real results. We created the commented code below in R and generated a TeX file using R markdown.

```r
rm(list=ls())
library(foreign)
library(ggplot2)
```

Loading data

```r
setwd("~/Dropbox/Uruguay state capacity/C722MDEO 11-12-2014")

cA <- read.csv("C722MDEOA.csv")
cB <- read.csv("C722MDEOB.csv")
cC <- read.csv("C722MDEOC.csv")
cD <- read.csv("C722MDEOD.csv")


data <- rbind(cA,cB,cC,cD)
```

Correcting outliers and coded vars

```r
data$B1[data$B1==7070] <- NA
data$B1[data$B1==87] <- 0
data$B2[data$B2==87] <- 0
data$S4[data$S4==7] <- "Otro"
data$D11[data$D11==7] <- "Otro"
data$D11p2[data$D11p2==5] <- "Otro"
data$E3[data$E3==99] <- "No voto"
data$E4p2[data$E4p2==99] <- "No voto"
data$E5[data$E5==99] <- "No vota en Montevideo"
data$E6p2[data$E6p2==99] <- "No simpatiza con ningun partido"
data$D9p3[data$D9p3==0] <- NA # confirmar con CIFRA
data$G3[data$G3==8] <- NA

# Replacing 88, 89, 99 as NA
for (i in 3:ncol(data)){
  data[,i][data[,i]==88] <- NA
  data[,i][data[,i]==89] <- NA
  data[,i][data[,i]==99] <- NA
}
```

Shuffling treatment labels

```r
data$treatment <- sample(data$tipo, length(data$tipo), replace=F)
```

When we are ready to analize the real data we will replace the line above with the following line:

```r
# data$treatment <- data$tipo
```


t-test function

```r
# t test with SEs
# df for two sample t test with unequal variances
t.test.se <- function(y,x, two.tailed=TRUE){

  # Calculating difference in means
  mean1 <- mean(y[x==1], na.rm=T)
  mean0 <- mean(y[x==0], na.rm=T)
  diff <- mean1 - mean0

  # Calculating SE of the difference
  N1 <- length(na.omit(y[x==1]))
  N0 <- length(na.omit(y[x==0]))
  var1 <- var(y[x==1],na.rm=T)
  var0 <- var(y[x==0],na.rm=T)
  varN1 <- var1/N1
  varN0 <- var0/N0
  se.diff <- sqrt(varN1 + varN0)

  # T-statistic
  t <- diff/se.diff

  # Degrees of freedom
  df.num <- ((varN1 + varN0)^2)
  df.den <- (varN1^2)/(N1-1) + (varN0^2)/(N0-1)
  df <- df.num/df.den

  # P-value
  if(two.tailed==TRUE){
        if (t>=0) { p <- pt(t, df, lower.tail=F) +
                      pt(-t, df, lower.tail=T) }
        if (t<0) {p <- pt(t, df, lower.tail=T) +
                    pt(-t, df, lower.tail=F)}
    }

  if(two.tailed==FALSE){
        if (t>=0) {p <- pt(t, df, lower.tail=F)}
        if (t<0) {p <- pt(t, df, lower.tail=T)}
     }

  # Preparing output
  res <- c(mean1, mean0, diff, se.diff,
          t, (N1+N0), df, p)
  names(res) <- c("Mean 1", "Mean 0", "Difference",
                  "SE Diff","t-stat", "N", "df", "p-value")

  return(c(res))
}
```

Analysis

Social benefits vs individual benefits.

```
# Pooling the individual benefits treatments (questionaire versions 2 and 3)
# and recoding as treatment=1.
data$social_individual <- ifelse((data$treatment==2|data$treatment==3),1,NA)
# Recoding the social benefits treatment as 0.
data$social_individual <- ifelse((data$treatment==1),0,data$social_individual)

# Outcome: "Policies that reward good taxpayers are a waste of money"
# totally disagree (0) - totally agree (10)
ben1 <- with(data,
     t.test.se(S1p1, social_individual, two.tailed=TRUE))

# Outcome: "It is worth it to be up to date with ones taxes"
# totally disagree (0) - totally agree (10)
ben2 <- with(data,
     t.test.se(S1p3, social_individual, two.tailed=TRUE))

social_individual <- rbind(ben1, ben2)
rownames(social_individual) <- c("Rewards are waste of money",
                                "Worth it to be up to date")
```

```
social_individual <- social_individual[order(social_individual[,8], decreasing=F),]

# Ordering p-values in decreasing order
ordered.ps <- social_individual[, 8]

# Building reference vector to compare to ordered p-values
FDR_reference <- .05*(1:length(ordered.ps)/length(ordered.ps))

# Comparing p-values to referece vector
FDR <- as.data.frame(cbind(ordered.ps,FDR_reference,
                          ordered.ps<=FDR_reference))

if (sum(FDR[,3])>0){
  fdr <- which(FDR[,1]==max(FDR[,1][FDR[,3]==1]))
  FDR[,4] <- c(rep("reject null", fdr), rep("do not reject", nrow(FDR)-fdr))}
if (sum(FDR[,3])==0){
  FDR[,4] <- rep("do not reject", nrow(FDR))}

Bonferroni_reference <- rep(.05/nrow(FDR), nrow(FDR))

Bonferroni_reject <- ifelse(ordered.ps<=Bonferroni_reference,
                           "reject null", "do not reject")

social_individual <- cbind(social_individual,
                          FDR[,c(2,4)],
                          Bonferroni_reference,
```



```
                           Bonferroni_reject)

names(social_individual)[10] <- "FDR_reject"

social_individual
```

Adjusting p-values

```
##                            Mean 1 Mean 0 Difference SE Diff  t-stat    N
## Rewards are waste of money  3.130  3.002    0.12840  0.1849  0.6946 1660
## Worth it to be up to date   7.412  7.509   -0.09638  0.1537 -0.6269 1682
##                              df p-value FDR_reference    FDR_reject
## Rewards are waste of money 1175  0.4875         0.025 do not reject
## Worth it to be up to date  1199  0.5308         0.050 do not reject
##                            Bonferroni_reference Bonferroni_reject
## Rewards are waste of money                0.025     do not reject
## Worth it to be up to date                 0.025     do not reject
```

Discretionary vs lottery allocation of benefits

```
# Generating discretionary benefits dummy where surveys with the discretionary version
# of the survey are 1.
data$treat_discretion <- ifelse((data$treatment==4),1,0)

# Outcome: "In Montevideo, rewards for good taxpayers go to the same people as always"
# totally disagree (0) - totally agree (10)
discretion1 <- with(data,
     t.test.se(S1p4, treat_discretion, two.tailed=FALSE))
discretion1
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##  5.310e+00  5.293e+00  1.665e-02  2.211e-01  7.534e-02  1.542e+03
##         df    p-value
##  7.068e+02  4.700e-01
```

```
# Outcome: "Policies that reward good taxpayers are a waste of money"
# totally disagree (0) - totally agree (10)
discretion2 <- with(data,
     t.test.se(S1p1, treat_discretion, two.tailed=TRUE))
discretion2
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##  3.108e+00  3.086e+00  2.187e-02  1.789e-01  1.222e-01  2.234e+03
##         df    p-value
##  9.674e+02  9.027e-01
```

```
# Outcome: "It is worth it to be up to date with ones taxes"
# totally disagree (0) - totally agree (10)
discretion3 <- with(data,
     t.test.se(S1p3, treat_discretion, two.tailed=FALSE))
 discretion3
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##  7.736e+00  7.445e+00  2.910e-01  1.341e-01  2.169e+00  2.266e+03
##         df    p-value
##  1.123e+03  1.513e-02
```

```
# Outcome: "In general, the municipal government does a good job"
# totally disagree (0) - totally agree (10)
discretion4 <- with(data,
     t.test.se(S1p2, treat_discretion, two.tailed=TRUE))
discretion4
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##     4.7785     4.5649     0.2136     0.1369     1.5602  2313.0000
##         df    p-value
##  1102.6385     0.1190
```

```
# Outcome: "How would you classify the taxes that the municipal government charges in general"
# very just (1) - not just at all (4)
discretion5 <-with(data,
     t.test.se(S1p5, treat_discretion, two.tailed=FALSE))
discretion5
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##    2.66497    2.74222   -0.07725    0.03420   -2.25910 2291.00000
##         df    p-value
## 1073.84987    0.01204
```

```
discretion <- rbind(discretion1, discretion2, discretion3,
                    discretion4, discretion5)
rownames(discretion) <- c("Rewards go to the same people as always",
                          "Rewards are waste of money",
                          "Worth it to be up to date",
                          "Mun.gov. does a good job",
                          "Mun. taxes are just")
```

```
discretion <- discretion[order(discretion[,8], decreasing=F),]

# Ordering p-values in decreasing order
ordered.ps <- discretion[, 8]

# Building reference vector to compare to ordered p-values
FDR_reference <- .05*(1:length(ordered.ps)/length(ordered.ps))

# Comparing p-values to referece vector
FDR <- as.data.frame(cbind(ordered.ps,FDR_reference,
                           ordered.ps<=FDR_reference))

if (sum(FDR[,3])>0){
  fdr <- which(FDR[,1]==max(FDR[,1][FDR[,3]==1]))
  FDR[,4] <- c(rep("reject null", fdr), rep("do not reject", nrow(FDR)-fdr))}
```



```
if (sum(FDR[,3])==0){
  FDR[,4] <- rep("do not reject", nrow(FDR))}

Bonferroni_reference <- rep(.05/nrow(FDR), nrow(FDR))

Bonferroni_reject <- ifelse(ordered.ps<=Bonferroni_reference,
                            "reject null", "do not reject")

discretion <- cbind(discretion,
                           FDR[,c(2,4)],
                           Bonferroni_reference,
                           Bonferroni_reject)

names(discretion)[10] <- "FDR_reject"

discretion
```

Adjusting p-values

```
##                                         Mean 1 Mean 0 Difference SE Diff
## Mun. taxes are just                      2.665  2.742   -0.07725  0.0342
## Worth it to be up to date                7.736  7.445    0.29100  0.1341
## Mun.gov. does a good job                 4.779  4.565    0.21358  0.1369
## Rewards go to the same people as always  5.310  5.293    0.01665  0.2211
## Rewards are waste of money               3.108  3.086    0.02187  0.1789
##                                          t-stat    N     df p-value
## Mun. taxes are just                     -2.25910 2291 1073.8 0.01204
## Worth it to be up to date                2.16933 2266 1123.2 0.01513
## Mun.gov. does a good job                 1.56022 2313 1102.6 0.11899
## Rewards go to the same people as always  0.07534 1542  706.8 0.46998
## Rewards are waste of money               0.12223 2234  967.4 0.90274
##                                         FDR_reference    FDR_reject
## Mun. taxes are just                              0.01   reject null
## Worth it to be up to date                        0.02   reject null
## Mun.gov. does a good job                         0.03 do not reject
## Rewards go to the same people as always          0.04 do not reject
## Rewards are waste of money                       0.05 do not reject
##                                         Bonferroni_reference
## Mun. taxes are just                                     0.01
## Worth it to be up to date                               0.01
## Mun.gov. does a good job                                0.01
## Rewards go to the same people as always                 0.01
## Rewards are waste of money                              0.01
##                                         Bonferroni_reject
## Mun. taxes are just                         do not reject
## Worth it to be up to date                   do not reject
## Mun.gov. does a good job                    do not reject
## Rewards go to the same people as always     do not reject
## Rewards are waste of money                  do not reject
```

Conditional on finding significant p-values ($p<.05$) for either of these tests, we will test for order effects by comparing the mean of the "individual benefits" treatment when it appears first to its mean when it appears second to see if they are statistically different.

Fines and charges vs. benefits of tax holidays

```r
# Creating dataframe with all treatments and outcomes.

# Keeping outcomes we want for benefits
ben <- data[(data$treatment!=4),c(59,60,62,111)]
names(ben)
```

```
## [1] "S1p2"      "S1p3"      "S1p5"      "treatment"
```

```r
# Benefits pooled (S1p2) (A, B and C)
ben$benefits_punishments <- 1

# versus fines and charges pooled (M1p3)  (A, B , C and D)
fin <- data[,c(50,51,54,111)]
names(fin)
```

```
## [1] "M1p2"      "M1p3"      "M1p6"      "treatment"
```

```r
# Pooling punishments
fin$benefits_punishments <- 0

# For three questions the outcomes are the same in the punishments and benefits conditions
# but have different survey question numbers. Here we rename the variables so
# that we can bind the datasets into one.
names(ben)
```

```
## [1] "S1p2"                 "S1p3"                 "S1p5"
## [4] "treatment"            "benefits_punishments"
```

```r
names(ben)[1:3] <- c("M1p3", "M1p2", "M1p6")

pooled <- rbind(ben,fin)

# Checking that the dimensions of the pooled dataset are correct.
stopifnot(nrow(data)*2-nrow(cD)==nrow(pooled))
```

```r
# Outcome: "In general, the municipal government does a good job"
# totally disagree (0) - totally agree (10)
benefits_punishments1 <- with(pooled,
     t.test.se(M1p3, benefits_punishments, two.tailed=FALSE))
benefits_punishments1
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##    4.56494    4.76322   -0.19828    0.09662   -2.05223 4006.00000
##         df    p-value
## 3694.54263    0.02011
```



```r
# Outcome: "In Montevideo, it is worth it to be up to date on ones taxes"
# totally disagree (0) - totally agree (10)
benefits_punishments2 <- with(pooled,
     t.test.se(M1p2, benefits_punishments, two.tailed=TRUE))
benefits_punishments2
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##  7.445e+00  7.853e+00 -4.082e-01  9.419e-02 -4.334e+00  3.955e+03
##         df    p-value
##  3.459e+03  1.508e-05
```

```r
# Outcome: "How would you classify the taxes that the municipal government charges?"
# very just (1) - not just at all (4)
benefits_punishments3 <- with(pooled,
     t.test.se(M1p6, benefits_punishments, two.tailed=FALSE))
benefits_punishments3
```

```
##     Mean 1     Mean 0 Difference    SE Diff     t-stat          N
##  2.742e+00  2.705e+00  3.729e-02  2.377e-02  1.569e+00  3.994e+03
##         df    p-value
##  3.651e+03  5.841e-02
```

```r
benefits_punishments <- rbind(benefits_punishments1,
                              benefits_punishments2,
                              benefits_punishments3)
rownames(benefits_punishments) <- c("Mun. gov. does a good job",
                                    "Worth it to be up to date",
                                    "Mun. taxes are just")
```

```r
benefits_punishments <- benefits_punishments[order(benefits_punishments[,8], decreasing=F),]

# Ordering p-values in decreasing order
ordered.ps <- benefits_punishments[, 8]

# Building reference vector to compare to ordered p-values
FDR_reference <- .05*(1:length(ordered.ps)/length(ordered.ps))

# Comparing p-values to referece vector
FDR <- as.data.frame(cbind(ordered.ps,FDR_reference,
                           ordered.ps<=FDR_reference))

if (sum(FDR[,3])>0){
  fdr <- which(FDR[,1]==max(FDR[,1][FDR[,3]==1]))
  FDR[,4] <- c(rep("reject null", fdr), rep("do not reject", nrow(FDR)-fdr))}
if (sum(FDR[,3])==0){
  FDR[,4] <- rep("do not reject", nrow(FDR))}

Bonferroni_reference <- rep(.05/nrow(FDR), nrow(FDR))
```

```r
Bonferroni_reject <- ifelse(ordered.ps<=Bonferroni_reference,
                            "reject null", "do not reject")

benefits_punishments <- cbind(benefits_punishments,
                              FDR[,c(2,4)],
                              Bonferroni_reference,
                              Bonferroni_reject)

names(benefits_punishments)[10] <- "FDR_reject"

benefits_punishments
```

Adjusting p-values

```
##                           Mean 1 Mean 0 Difference SE Diff t-stat    N
## Worth it to be up to date  7.445  7.853   -0.40819 0.09419 -4.334 3955
## Mun. gov. does a good job  4.565  4.763   -0.19828 0.09662 -2.052 4006
## Mun. taxes are just        2.742  2.705    0.03729 0.02377  1.569 3994
##                             df   p-value FDR_reference    FDR_reject
## Worth it to be up to date 3459 1.508e-05       0.01667   reject null
## Mun. gov. does a good job 3695 2.011e-02       0.03333   reject null
## Mun. taxes are just       3651 5.841e-02       0.05000 do not reject
##                           Bonferroni_reference Bonferroni_reject
## Worth it to be up to date              0.01667       reject null
## Mun. gov. does a good job              0.01667     do not reject
## Mun. taxes are just                    0.01667     do not reject
```

Conditional on finding significant p-values (p<.05) for either of these tests, we will test for order effects by comparing the mean of the "individual benefits", "individual punishment" and "social punishment" treatments when they appear first to their mean when they appear second to see if they are statistically different.

Table 0.3: **Revision to Table 7.3 in pre-analysis plan: Hypotheses, Outcomes, and Tests**

Hypotheses	Data Sources	Outcomes	Comparisons	Tests
Hypotheses 1A, 1B, 1C (Winning lottery)	Natural Exp. (Admin. data)	$t+\ldots$ 1. Compliance (0-1) 2. Missed Payments 3. Total Debt	Winners vs. Non-Winners (Eligibles)	1. ~~K-S test~~ 2. Diff-in-Diff 3. Persistence of effects
Mechanism 1A.1 (*Informational*)	Field Exp. (Admin data) (Eligibles and Ineligibles)	$t+1$ & $t+4$: 1. Compliance (0-1) 2. Web bill request	Existence of Lottery vs. Control	1. Diff.-in-Diff 2. IV
	Field Exp. vs. Nat. Exp. (Admin data)	$t+4$: 1. Compliance (0-1) 2. Web bill request	Effect of Info. vs. Effect of Winning	1. Diff. of Diff.-in-Diffs
	Survey exp. (Survey data)	1. Worth it to pay	Ex. of Lottery vs. Discretion	1. Diff. of Means
Mechanism 1A.2 (*Attitudinal*)	~~Field Exp~~ ~~(Survey data)~~ ~~(Eligibles + Ineligibles)~~	~~$t+1$~~ ~~1. Trust in municipality~~ ~~2. Trust in civil servants~~ ~~3. Eval. of Mayor~~ ~~4. Fairness Taxes~~ ~~5. Fairness Prop. Taxes~~ ~~6. Support Amnesties~~	~~Existence of Lottery vs. Control~~	~~1. Diff. of Means~~ ~~2 and 3. Diff. of Means~~ ~~(Het. effects)~~
	~~Nat. Exp~~ ~~(Survey data)~~	~~$t+\ldots$~~ ~~1. Trust in municipality~~ ~~2. Trust in civil servants~~ ~~3. Eval. of Mayor~~ ~~4. Fairness Taxes~~ ~~5. Fairness Spec. Tax~~ ~~6. Opinion of lottery~~ ~~(Version 1 of survey)~~	~~Winners vs. Losers~~	~~1.Diff. of Means~~ ~~2. Diff. of Means~~ ~~Het. effects~~ – ~~recent vs. old winners)~~
	Survey Exp (Survey data)	t 1. Lotteries are waste of money 2. Eval. of City Hall 3. Benefits go to "same as always" 4. Taxes are fair	Lottery treatments vs. non-lottery (discretion) treatment	1.Diff. of Means ~~2. Diff. of Means~~ ~~Het. effects~~
Mechanism 1B.1 (*Income effects*)	Nat. Exp. (Admin. Data)	$t+\ldots$ 1. Compliance 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Heter. effects by cost of payment (property value)

Table 0.4: **Revision to Table 7.3 in pre-analysis plan: Hypotheses, Outcomes, and Tests (Cont.)**

Mechanism 1B.2. (*Habit effects*)	Nat. Exp. (Admin. Data)	$t+\dots$ 1. Compliance 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Het. effects by time since winning
Mechanism 1C.1. (*Erroneous beliefs*)	Nat. Exp. (Admin/Survey Data)	$t+\dots$ 1. Compliance (0-1) 2. Missed Payments 3. Total Debt	Winners vs. Losers	~~1. Het. effects by beliefs about non-indep. of winning~~
Hypothesis 2A (*Punishments*)	Field Exp. (Admin. Data) Survey Exp (Survey data)	$t+1$ 1. Compliance (0-1) t 1. Worth it to pay 2. Eval. of City Hall 4. Taxes are fair	Existence of Fines vs. Control Fines treatment vs. lotteries treatment	1. Diff. of Means 1.Diff. of Means
Mechanism 2A.1 (*Prob. punishment*)	Field Exp. (Survey Data)	$t+1$ 1. Belief in Prob. of Fine	Existence of Fines vs. vs. Control	1. Diff. of Means
Hypothesis 2B (*Rewards vs. Punishments*)	Field Exp. (Admin. Data)	$t+1$ 1. Compliance (0-1)	Existence of Fines vs. Control	1. Diff.-in-Diff
Hypothesis 2C (*Marginal taxpayers I*)	Field Exp (Admin. Data)	$t+1$	Benefit vs. Sanction	1. Het. effects by payment history
Hypothesis 2D (*Marginal taxpayers II*)	Field Exp. (Admin./Survey)	$t+1$ 1. Compliance (0-1)	Benefit vs. Sanction	1. Het. effects by payment cost
Hypothesis 3A (*Social benefits*)	Field Exp. (Admin. Data)	$t+1$	Social Ben. vs. Indiv. Ben.	1. Diff. of Means
Hypothesis 3B (*Social sanction*)	Field Exp. (Admin. Data)	$t+1$	Social Sanc. vs. Indiv. Sanc.	1. Diff. of Means
Hypothesis 3A-B (*Social benefits vs. social sanctions*)	Field Exp. (Admin. Data)	$t+1$	Effect of vs. Social Sanc. vs. effect of Indiv. Sanc.	1. Diff. of Diff. of Means
Hypothesis 3A-B (*Ind. benefits vs. social bens*)	Survey Exp. (Survey data)	t Rewards waste of money Benefits to same as always	Social Sanc. vs. effect of Indiv. Sanc.	Diff. of Means

In the table, t refers to tax payment periods, of which there are 3-6 per year, depending on the tax. Thus, for winners of the lottery, $t=0$ is the period in which they won the lottery; $t=1$ is the following tax payment period; and because they win a year free of tax payments $t+4$ is the next payment period in which they owe taxes. Property taxes are paid three times per year.

Table 0.5: **Revision to Table 7.3 in pre-analysis plan: Hypotheses, Outcomes, and Tests**

Hypotheses	Data Sources	Outcomes	Comparisons	Tests
Hypotheses 1A, 1B, 1C (Winning lottery)	Natural Exp. (Admin. data)	$t+\dots$ 1. Compliance (0-1) 2. Missed Payments 3. Total Debt	Winners vs. Non-Winners (Eligibles)	1. ~~K-S test~~ 2. Diff-in-Diff 3. Persistence of effects
Mechanism 1A.1 (*Informational*)	Field Exp. (Admin data) (Eligibles and Ineligibles) Field Exp. vs. Nat. Exp. (Admin data) Survey exp. (Survey data)	$t+1$ & $t+4$: 1. Compliance (0-1) 2. Web bill request $t+4$: 1. Compliance (0-1) 2. Web bill request 1. Worth it to pay	Existence of Lottery vs. Control Effect of Info. vs. Effect of Winning Ex. of Lottery vs. Discretion	1. Diff.-in-Diff 2. IV 1. Diff. of Diff.-in-Diffs 1. Diff. of Means
Mechanism 1A.2 (*Attitudinal*)	~~Field Exp (Survey data) (Eligibles + Ineligibles)~~ ~~Nat. Exp (Survey data)~~ Survey Exp (Survey data)	~~$t+1$~~ ~~1. Trust in municipality~~ ~~2. Trust in civil servants~~ ~~3. Eval. of Mayor~~ ~~4. Fairness Taxes~~ ~~5. Fairness Prop. Taxes~~ ~~6. Support Amnesties~~ ~~$t+\dots$~~ ~~1. Trust in municipality~~ ~~2. Trust in civil servants~~ ~~3. Eval. of Mayor~~ ~~4. Fairness Taxes~~ ~~5. Fairness Spec. Tax~~ ~~6. Opinion of lottery (Version 1 of survey)~~ t 1. Lotteries are waste of money 2. Eval. of City Hall 3. Benefits go to "same as always" 4. Taxes are fair	~~Existence of Lottery vs. Control~~ ~~Winners vs. Losers~~ Lottery treatments vs. non-lottery (discretion) treatment	~~1. Diff. of Means~~ ~~2 and 3. Diff. of Means (Het. effects)~~ ~~1.Diff. of Means~~ ~~2. Diff. of Means Het. effects – recent vs. old winners)~~ 1.Diff. of Means ~~2. Diff. of Means Het. effects~~
Mechanism 1B.1 (*Income effects*)	Nat. Exp. (Admin. Data)	$t+\dots$ 1. Compliance 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Heter. effects by cost of payment (property value)

Table 0.6: **Revision to Table 7.3 in pre-analysis plan: Hypotheses, Outcomes, and Tests (Cont.)**

Mechanism 1B.2. (*Habit effects*)	Nat. Exp. (Admin. Data)	$t+\ldots$ 1. Compliance 2. Missed Payments 3. Total Debt	Winners vs. Losers	1. Het. effects by time since winning
Mechanism 1C.1. (*Erroneous beliefs*)	Nat. Exp. (Admin/Survey Data)	$t+\ldots$ 1. Compliance (0-1) 2. Missed Payments 3. Total Debt	Winners vs. Losers	~~1. Het. effects by beliefs about non-indep. of winning~~
Hypothesis 2A (*Punishments*)	Field Exp. (Admin. Data) Survey Exp (Survey data)	$t+1$ 1. Compliance (0-1) t 1. Worth it to pay 2. Eval. of City Hall 4. Taxes are fair	Existence of Fines vs. Control Fines treatment vs. lotteries treatment	1. Diff. of Means 1.Diff. of Means
Mechanism 2A.1 (*Prob. punishment*)	Field Exp. (Survey Data)	$t+1$ 1. Belief in Prob. of Fine	Existence of Fines vs. vs. Control	1. Diff. of Means
Hypothesis 2B (*Rewards vs. Punishments*)	Field Exp. (Admin. Data)	$t+1$ 1. Compliance (0-1)	Existence of Fines vs. Control	1. Diff.-in-Diff
Hypothesis 2C (*Marginal taxpayers I*)	Field Exp (Admin. Data)	$t+1$	Benefit vs. Sanction	1. Het. effects by payment history
Hypothesis 2D (*Marginal taxpayers II*)	Field Exp. (Admin./Survey)	$t+1$ 1. Compliance (0-1)	Benefit vs. Sanction	1. Het. effects by payment cost
Hypothesis 3A (*Social benefits*)	Field Exp. (Admin. Data)	$t+1$	Social Ben. vs. Indiv. Ben.	1. Diff. of Means
Hypothesis 3B (*Social sanction*)	Field Exp. (Admin. Data)	$t+1$	Social Sanc. vs. Indiv. Sanc.	1. Diff. of Means
Hypothesis 3A-B (*Social benefits vs. social sanctions*	Field Exp. (Admin. Data)	$t+1$	Effect of vs. Social Sanc. vs. effect of Indiv. Sanc.	1. Diff. of Diff. of Means
Hypothesis 3A-B (*Ind. benefits vs. social bens*)	Survey Exp. (Survey data)	t Rewards waste of money Benefits to same as always	Social Sanc. vs. effect of Indiv. Sanc.	Diff. of Means

In the table, t refers to tax payment periods, of which there are 3-6 per year, depending on the tax. Thus, for winners of the lottery, $t = 0$ is the period in which they won the lottery; $t = 1$ is the following tax payment period; and because they win a year free of tax payments $t + 4$ is the next payment period in which they owe taxes. Property taxes are paid three times per year.

G.4 PAP code applied to real data

In the document in this section, we replicate all the registered analyses for which we provided code in the first amendment to our PAP (registered October 19, 2014). In that amendment, we used nearly (but not fully) complete data, reshuffled treatment labels, and reported the mock analysis. Here we reproduce all those analyses using real outcome data. For readibility and ease, we have modernized some of the code using functions from packages not available at the time we filed the first amendment. However, all analyses follow those laid out in the registered amendment to our PAP.

PAP Analysis for Disrupting Compliance: The Impact of a Randomized Tax Holiday in Uruguay

2023-05-06

```r
##################################################################################
# PAP 1 & 2 code execution
# Last revised April 2023
##################################################################################

# Basic setup --
rm(list=ls())
set.seed(1234)
options(scipen=99999, digits=3)


###############################################
message("required libraries and setwd")
```

```
## required libraries and setwd
```

```r
# Load/install packages --
if (!require("pacman")) install.packages("pacman")
```

```
## Loading required package: pacman
```

```r
pacman::p_load(
  plyr,
  ggplot2,
  reshape2,
  zoo,
  sandwich,
  AER,
  xtable,
  stats,
  tidyr,
  dplyr,
  weights,
  estimatr
)

conflicted::conflicts_prefer(dplyr::filter)
```

```
## [conflicted] Will prefer dplyr::filter over any other package.
```



```r
if (grepl ("/Users/gtunon", getwd ()) == TRUE){
  home <- "~/Dropbox/Working_papers/Uruguay_state_capacity/JOP_replication"
} else {
  # home <- " "
}
setwd(home)

source("code/t_test.R")

##############################################################
# PAP 1 - NATURAL AND FIELD EXPERIMENTS
load("data/panel_taxtime.Rda")
load("data/naturalex_debt_gtp.Rda")
load("data/fieldex_data.Rda")
##############################################################

# switch tax names to english
taxes_panel$tax <- taxes_panel$TRIBUTO
taxes_panel$tax <- as.factor(taxes_panel$tax)
levels(taxes_panel$tax) <-  c("Property", "Vehicle", "Sewage", "Head")

# check when the holiday takes place for each tax
holiday <- taxes_panel %>% dplyr::filter(cuota_exonerada==1) %>% group_by(tax) %>%
  dplyr::summarise(
    holiday_start = min(t),
    holiday_end = max(t)
  )

holiday
```

```
## # A tibble: 4 x 3
##   tax      holiday_start holiday_end
##   <fct>            <int>       <int>
## 1 Property             1           4
## 2 Vehicle             -2           3
## 3 Sewage               1           7
## 4 Head                 2           8
```

```r
# For the treatment group, we replace with NAs the observations under the holiday
taxes_panel <- taxes_panel %>% inner_join(holiday)
```

```
## Joining with `by = join_by(tax)`
```

```r
taxes_panel$en_fecha[taxes_panel$TREATMENT==1 &
  taxes_panel$t>=taxes_panel$holiday_start &
                      taxes_panel$t<=taxes_panel$holiday_end] <- NA
taxes_panel$cuotas_adeudadas[taxes_panel$TREATMENT==1 &
                      taxes_panel$t>=taxes_panel$holiday_start &
                      taxes_panel$t<=taxes_panel$holiday_end] <- NA

taxes_panel$missed_payment <- as.numeric(taxes_panel$en_fecha==0)
```

```r
taxes_panel$compliance <- as.numeric(taxes_panel$cuotas_adeudadas==0)

taxes_panel <- taxes_panel[!(taxes_panel$tax=="Vehicle" &
                              taxes_panel$t %in% c(23,24)),]


##############################################

## Impact of the tax holiday lottery

plot_data <- taxes_panel %>% filter(ES_BP==1) %>%
  group_by(tax, t, TREATMENT) %>% dplyr::summarize(
  N = n(),

  # missed current payment
  missed_payment_mean = mean(missed_payment, na.rm=T),
  se_missed_payment_mean = sd(missed_payment, na.rm=T)/sqrt(N),
  missed_payment_upper= missed_payment_mean + qnorm(.975)*(se_missed_payment_mean),
  missed_payment_lower= missed_payment_mean - qnorm(.975)*(se_missed_payment_mean),

  compliance_mean = mean(compliance, na.rm=T),
  se_compliance_mean = sd(compliance, na.rm=T)/sqrt(N),
  compliance_upper= compliance_mean + qnorm(.975)*se_compliance_mean,
  compliance_lower= compliance_mean - qnorm(.975)*se_compliance_mean,

  nr_missed_mean = mean(cuotas_adeudadas, na.rm=T),
  se_nr_missed_mean = sd(cuotas_adeudadas, na.rm=T)/sqrt(N),
  nr_missed_upper= nr_missed_mean + qnorm(.975)*(se_nr_missed_mean),
  nr_missed_lower= nr_missed_mean - qnorm(.975)*(se_nr_missed_mean)

  )
```

```
## `summarise()` has grouped output by 'tax', 't'. You can override using the
## `.groups` argument.
```

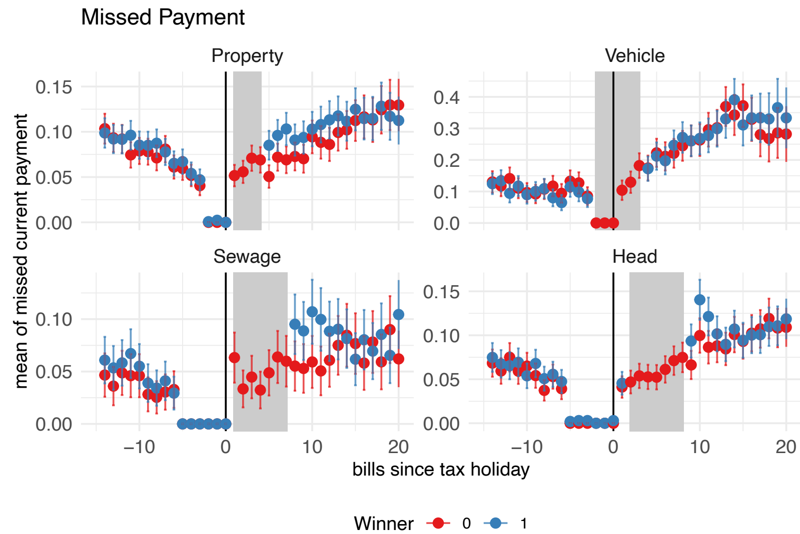
```r
plot_data <- plot_data %>% inner_join(holiday)
```

```
## Joining with `by = join_by(tax)`
```

```r
plot_data$Winner <- as.factor(plot_data$TREATMENT)

# missed payment
plot_data %>% filter(t>-15 & t<21) %>%
ggplot(aes(x=t, y=missed_payment_mean, color=Winner)) +
  facet_wrap(tax ~ . , scales = "free_y") +
  geom_rect(aes(xmin=holiday_start, xmax=holiday_end, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80", alpha=.1) +
  geom_vline(aes(xintercept = 0)) +
  geom_point(size=2.4) +
  #ylim(0, .5) +
  xlim(-15,20) +
  xlab("bills since tax holiday") +
```



```r
  ylab("mean of missed current payment") +
  geom_errorbar(aes(ymin=missed_payment_lower,
                    ymax=missed_payment_upper), # colour="blue",
                width=.3, alpha=.75) +
  scale_color_brewer(palette="Set1") +
  ggtitle("Missed Payment") +
  theme_minimal() +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1.2)),
        legend.position = "bottom")
```



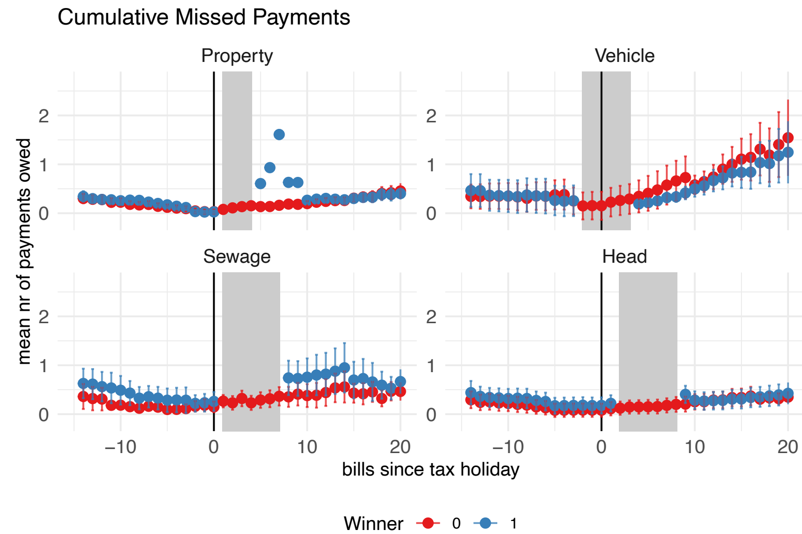

```r
# compliance
plot_data %>% filter(t>-15 & t<21) %>%
ggplot(aes(x=t, y=compliance_mean, color=Winner)) +
  facet_wrap(tax ~ . , scales = "free_y") +
  geom_rect(aes(xmin=holiday_start, xmax=holiday_end, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80", alpha=.1) +
  geom_vline(aes(xintercept = 0)) +
  geom_point(size=2.4) +
  #ylim(0, .5) +
  xlim(-15,20) +
  xlab("bills since tax holiday") +
```

```r
  ylab("mean of compliance") +
  geom_errorbar(aes(ymin=compliance_lower,
                    ymax=compliance_upper), # colour="blue",
                width=.3, alpha=.75) +
  scale_color_brewer(palette="Set1") +
  ggtitle("Compliance") +
  theme_minimal() +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1.2)),
        legend.position = "bottom")
```


```r
# nr missed payments
plot_data %>% filter(t>-15 & t<21) %>%
ggplot(aes(x=t, y=nr_missed_mean, color=Winner)) +
  facet_wrap(tax ~ . , scales = "free_y") +
  geom_rect(aes(xmin=holiday_start, xmax=holiday_end, ymin=-Inf, ymax=Inf),
            fill="gray80", color="gray80", alpha=.1) +
  geom_vline(aes(xintercept = 0)) +
  geom_point(size=2.4) +
  ylim(-.2, 2.75) +
  xlim(-15,20) +
  xlab("bills since tax holiday") +
```



```r
  ylab("mean nr of payments owed") +
  geom_errorbar(aes(ymin=nr_missed_lower,
                    ymax=nr_missed_upper), # colour="blue",
                width=.3, alpha=.75) +
  scale_color_brewer(palette="Set1") +
  ggtitle("Cumulative Missed Payments") +
  theme_minimal() +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1.2)),
        legend.position = "bottom")
```




```r
## All taxes together - for this we first need to standardize the timeline
taxes_panel$t_st <- taxes_panel$t

# drop data within holiday window for all eligible taxpayers
taxes_panel <- taxes_panel %>% filter(!(ES_BP==1 & t>0 & t<=holiday_end))
# for eligibles, fix full timeframe so that first payment period after holiday is always 3
taxes_panel$t_st[taxes_panel$ES_BP==1 & taxes_panel$t > 0] <-
  taxes_panel$t - taxes_panel$holiday_end + 3

# drop data within holiday window (but outside the standard 3 payment period holiday)
# for all ineligible taxpayers
```

```
taxes_panel <- taxes_panel %>% filter(!(ES_BP==0 & t>=4 & t<= holiday_end))

holiday
```

```
## # A tibble: 4 x 3
##   tax      holiday_start holiday_end
##   <fct>            <int>       <int>
## 1 Property             1           4
## 2 Vehicle             -2           3
## 3 Sewage               1           7
## 4 Head                 2           8
```

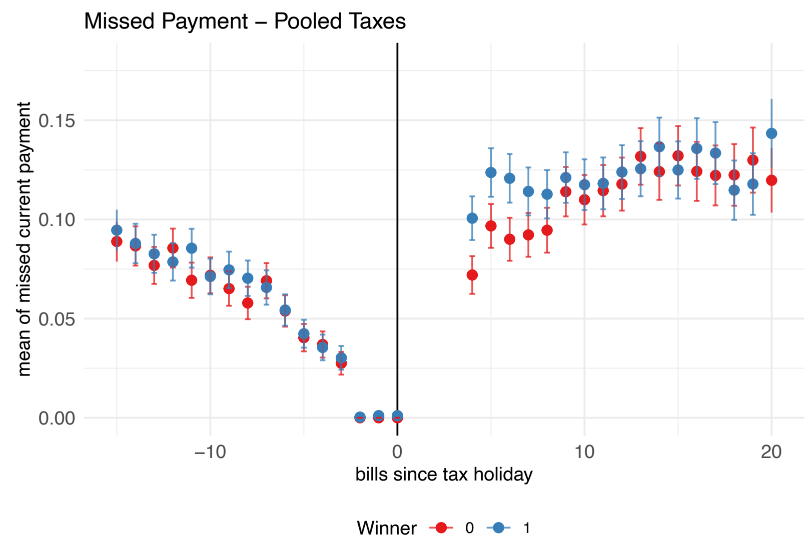
```
taxes_panel$st <- taxes_panel$holiday_end - 3

taxes_panel$t_st <- ifelse((taxes_panel$ES_BP==1 & taxes_panel$t > 0) |
                             (taxes_panel$ES_BP==0 & taxes_panel$t > 4),
                           taxes_panel$t - taxes_panel$st, taxes_panel$t)


plot_data <- taxes_panel %>% filter(ES_BP==1) %>%
  group_by(t_st, TREATMENT) %>% dplyr::summarize(
    N = n(),

    # missed current payment
    missed_payment_mean = mean(missed_payment, na.rm=T),
    se_missed_payment_mean = sd(missed_payment, na.rm=T)/sqrt(N),
    missed_payment_upper= missed_payment_mean + qnorm(.975)*(se_missed_payment_mean),
    missed_payment_lower= missed_payment_mean - qnorm(.975)*(se_missed_payment_mean),

    compliance_mean = mean(compliance, na.rm=T),
    se_compliance_mean = sd(compliance, na.rm=T)/sqrt(N),
    compliance_upper= compliance_mean + qnorm(.975)*se_compliance_mean,
    compliance_lower= compliance_mean - qnorm(.975)*se_compliance_mean,

    nr_missed_mean = mean(cuotas_adeudadas, na.rm=T),
    se_nr_missed_mean = sd(cuotas_adeudadas, na.rm=T)/sqrt(N),
    nr_missed_upper= nr_missed_mean + qnorm(.975)*(se_nr_missed_mean),
    nr_missed_lower= nr_missed_mean - qnorm(.975)*(se_nr_missed_mean)

  )
```

```
## `summarise()` has grouped output by 't_st'. You can override using the
## `.groups` argument.
```

```
plot_data$Winner <- as.factor(plot_data$TREATMENT)

# missed payment
ggplot(plot_data, aes(x=t_st, y=missed_payment_mean, color=Winner)) +
  geom_vline(aes(xintercept = 0)) +
  geom_point(size=2.4) +
  ylim(0, .18) +
  xlim(-15,20) +
```

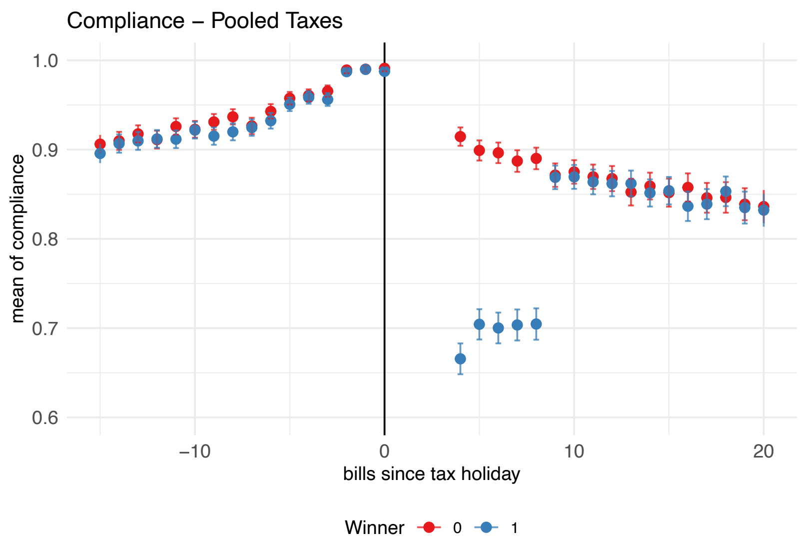


```
  xlab("bills since tax holiday") +
  ylab("mean of missed current payment") +
  geom_errorbar(aes(ymin=missed_payment_lower,
                    ymax=missed_payment_upper), # colour="blue",
                width=.3, alpha=.75) +
  scale_color_brewer(palette="Set1") +
  ggtitle("Missed Payment - Pooled Taxes") +
  theme_minimal() +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1.2)),
        legend.position = "bottom")
```




```
# compliance
ggplot(plot_data, aes(x=t_st, y=compliance_mean, color=Winner)) +
  geom_vline(aes(xintercept = 0)) +
  geom_point(size=2.4) +
  ylim(0.6, 1) +
  xlim(-15,20) +
  xlab("bills since tax holiday") +
  ylab("mean of compliance") +
  geom_errorbar(aes(ymin=compliance_lower,
                    ymax=compliance_upper), # colour="blue",
```

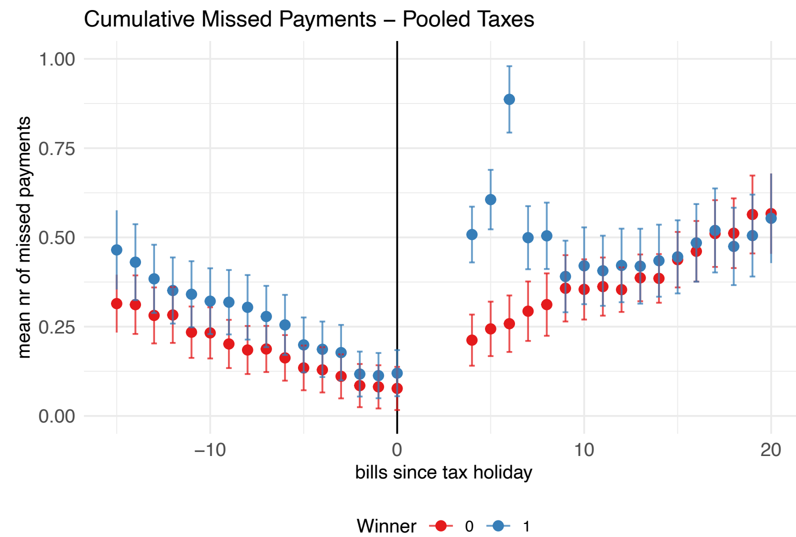
```
                 width=.3, alpha=.75) +
  scale_color_brewer(palette="Set1") +
  ggtitle("Compliance - Pooled Taxes") +
  theme_minimal() +
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1.2)),
        legend.position = "bottom")
```




```
# nr missed payments
ggplot(plot_data, aes(x=t_st, y=nr_missed_mean, color=Winner)) +
  geom_vline(aes(xintercept = 0)) +
  geom_point(size=2.4) +
  ylim(0, 1) +
  xlim(-15,20) +
  xlab("bills since tax holiday") +
  ylab("mean nr of missed payments") +
  geom_errorbar(aes(ymin=nr_missed_lower,
                    ymax=nr_missed_upper), # colour="blue",
                 width=.3, alpha=.75) +
  scale_color_brewer(palette="Set1") +
  ggtitle("Cumulative Missed Payments - Pooled Taxes") +
  theme_minimal() +
```



```
  theme(plot.title = element_text(size = rel(1.2)),
        axis.text = element_text(size = rel(1)),
        axis.title.y = element_text(size = rel(1)),
        axis.title.x = element_text(size = rel(1)),
        strip.text.x = element_text(size = rel(1.2)),
        legend.position = "bottom")
```




```
#################################################################
## TABLE 1 - NATURAL EXPERIMENT. Effects of the tax holiday
## (difference in differences analysis).
#################################################################

# rescaling the time variable to account for the taxes that have twice as
# many payments per year
taxes_panel$t_st_2 <- ifelse(taxes_panel$tax=="Sewage" | taxes_panel$tax=="Head",
                             taxes_panel$t_st/2, taxes_panel$t_st)


# 1 year diff in diff setup
dd_data1 <-  rbind.data.frame(taxes_panel %>% filter(ES_BP==1) %>%
                                group_by(CUENTA, tax, TREATMENT) %>% dplyr::summarise(
                                  compliance_mean_DiD_1yr =
                                    mean(compliance[t_st>3 & t_st<=6], na.rm=T) -
                                    mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                  missed_payment_mean_DiD_1yr =
```

```r
                                  mean(missed_payment[t_st>3 & t_st<=6], na.rm=T)-
                                  mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                nr_missed_payments_mean_DiD_1yr =
                                  mean(cuotas_adeudadas[t_st> 3 & t_st<=6], na.rm=T)-
                                  mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T),
                                compliance_mean_DiD_1yr.yr2 =
                                  mean(compliance[t_st>6 & t_st<=9], na.rm=T) -
                                  mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                missed_payment_mean_DiD_1yr.yr2 =
                                  mean(missed_payment[t_st>6 & t_st<=9], na.rm=T)-
                                  mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                nr_missed_payments_mean_DiD_1yr.yr2 =
                                  mean(cuotas_adeudadas[t_st> 6 & t_st<=9], na.rm=T)-
                                  mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T),
                                compliance_mean_DiD_1yr.yr3 =
                                  mean(compliance[t_st>9 & t_st<=12], na.rm=T) -
                                  mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                missed_payment_mean_DiD_1yr.yr3 =
                                  mean(missed_payment[t_st>9 & t_st<=12], na.rm=T)-
                                  mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                nr_missed_payments_mean_DiD_1yr.yr3 =
                                  mean(cuotas_adeudadas[t_st>9 & t_st<=12], na.rm=T)-
                                  mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T)
                              )
)
```

```
## `summarise()` has grouped output by 'CUENTA', 'tax'. You can override using the
## `.groups` argument.
```

```r
DiD_1yr_compliance <- difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1)
DiD_1yr_compliance
```

```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.199    0.00935   -21.3
##
## TREATMENT 0.000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper   DF
## TREATMENT   -0.218   -0.181 4611
```

```r
DiD_1yr_missed_payment <- difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1)
DiD_1yr_missed_payment
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0202     0.0065    3.11  0.00191  0.00745    0.033 5307
```

```r
DiD_1yr_nrmissed_payments <- difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1)
DiD_1yr_nrmissed_payments
```



```
## Design:  Standard
##           Estimate Std. Error t value                Pr(>|t|) CI Lower CI Upper
## TREATMENT    0.349     0.0389    8.98 0.0000000000000000000391    0.273    0.425
##             DF
## TREATMENT 4410
```

```r
difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT    -0.46     0.0155   -29.6
##
## TREATMENT 0.000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper   DF
## TREATMENT    -0.49   -0.429 1699
```

```r
difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0273    0.00866    3.16  0.00163   0.0103   0.0443 2147
```

```r
difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT    0.866     0.0419    20.6
##                                                                                            Pr(>|t|
## TREATMENT 0.0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000037
##           CI Lower CI Upper   DF
## TREATMENT    0.784    0.948 2336
```

```r
difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT  -0.0333     0.0205   -1.63    0.104  -0.0735  0.00688 748
```

```r
difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   0.0398     0.0143    2.78  0.00553   0.0117   0.0678 712
```

```r
difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT    0.249      0.166     1.5    0.134  -0.0769    0.574 676
```

```r
difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  -0.0363     0.0116   -3.14  0.00171   -0.059  -0.0136 1821
```

```r
difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0323     0.0111    2.92  0.00352   0.0106    0.054 1739
```

```r
difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  -0.0354     0.0399  -0.885    0.376   -0.114    0.043 1596
```

```r
difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   0.0459      0.024    1.91   0.0559 -0.00116   0.0931 702
```

```r
difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT  -0.0556     0.0273   -2.03   0.0425   -0.109 -0.00189 668
```

```r
difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1[dd_data1$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   -0.256      0.166   -1.54    0.124   -0.583   0.0707 411
```



```r
# 3 year diff in diff setup
dd_data <-  rbind.data.frame(taxes_panel %>% filter(ES_BP==1) %>%
                                 group_by(CUENTA, tax, TREATMENT) %>% dplyr::summarise(

                                   compliance_mean_DiD_3yr =
                                     mean(compliance[t_st>3 & t_st<=12], na.rm=T)-
                                     mean(compliance[t_st<0 & t_st>=(-9)], na.rm=T),
                                   missed_payment_mean_DiD_3yr =
                                     mean(missed_payment[t_st>3 & t_st<=12], na.rm=T)-
                                     mean(missed_payment[t_st<0 & t_st>=(-9)], na.rm=T),
                                   nr_missed_payments_mean_DiD_3yr =
                                     mean(cuotas_adeudadas[t_st> 3 & t_st<=12], na.rm=T)-
                                     mean(cuotas_adeudadas[t_st<0 & t_st>=(-9)], na.rm=T)
                                 )
)
```

```
## `summarise()` has grouped output by 'CUENTA', 'tax'. You can override using the
## `.groups` argument.
```

```r
difference_in_means(compliance_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data)
```

```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.116    0.00751   -15.5
##                                                                    Pr(>|t|) CI Lower
## TREATMENT 0.000000000000000000000000000000000000000000000000000000584   -0.131
##           CI Upper   DF
## TREATMENT   -0.102 5322
```

```r
difference_in_means(missed_payment_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data)
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0145    0.00576    2.52   0.0118  0.00321   0.0258 5451
```

```r
difference_in_means(nr_missed_payments_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data)
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT    0.136     0.0448    3.04  0.00234   0.0485    0.224 4666
```

```r
difference_in_means(compliance_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.281     0.0125   -22.5
```

```
##
## TREATMENT 0.0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper   DF
## TREATMENT   -0.306   -0.257 2029
```

```
difference_in_means(missed_payment_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0224     0.0083     2.7  0.00702  0.00612   0.0387 2208
```

```
difference_in_means(nr_missed_payments_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value                 Pr(>|t|) CI Lower
## TREATMENT    0.424     0.0461     9.2 0.0000000000000000000801    0.334
##           CI Upper   DF
## TREATMENT    0.515 2336
```

```
difference_in_means(compliance_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT  -0.0103     0.0169   -0.61    0.542  -0.0435   0.0229 745
```

```
difference_in_means(missed_payment_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   0.0234     0.0118    1.99   0.0468 0.000328   0.0466 732
```

```
difference_in_means(nr_missed_payments_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT    0.152        0.2   0.757    0.449   -0.242    0.545 598
```

```
difference_in_means(compliance_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  -0.0038     0.0104  -0.364    0.716  -0.0243   0.0167 1829
```



```
difference_in_means(missed_payment_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  0.00956    0.00931    1.03    0.305 -0.00871   0.0278 1748
```

```
difference_in_means(nr_missed_payments_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   -0.114     0.0528   -2.16   0.0307   -0.218  -0.0106 1770
```

```
difference_in_means(compliance_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   0.0165     0.0212   0.776    0.438  -0.0252   0.0581 719
```

```
difference_in_means(missed_payment_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT  -0.0127     0.0232  -0.548    0.584  -0.0582   0.0328 700
```

```
difference_in_means(nr_missed_payments_mean_DiD_3yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   -0.208      0.183   -1.14    0.255   -0.567    0.151 534
```

```
##############################################################
## TABLE 2 - NATURAL EXPERIMENT. Effect of the tax holiday (T-test)
## Comparing winners to non-winners, difference of means test
# for the total debt as of October, 2014.
##############################################################

difference_in_means(debt_amount ~ won_lottery,
                    data = naturalex_debt_gtp)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## won_lottery      184        297   0.619    0.536     -398      765 4948
```

```r
###############################################################
## TABLE 3 - NATURAL EXPERIMENT. Effects of the tax holiday
## (difference in differences analysis), 5 year window
###############################################################


# 5 year diff in diff setup
dd_data <-  rbind.data.frame(taxes_panel %>% filter(ES_BP==1) %>%
                               group_by(CUENTA, tax, TREATMENT) %>% dplyr::summarise(
                                 compliance_mean_DiD_5yr =
                                   mean(compliance[t_st>3 & t_st<=18], na.rm=T)-
                                   mean(compliance[t_st<0 & t_st>=(-15)], na.rm=T),
                                 missed_payment_mean_DiD_5yr =
                                   mean(missed_payment[t_st>3 & t_st<=18], na.rm=T)-
                                   mean(missed_payment[t_st<0 & t_st>=(-15)], na.rm=T),
                                 nr_missed_payments_mean_DiD_5yr =
                                   mean(cuotas_adeudadas[t_st> 3 & t_st<=18], na.rm=T)-
                                   mean(cuotas_adeudadas[t_st<0 & t_st>=(-15)], na.rm=T)
                               )
)
```

```
## `summarise()` has grouped output by 'CUENTA', 'tax'. You can override using the
## `.groups` argument.
```

```r
difference_in_means(compliance_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data)
```

```
## Design:  Standard
##           Estimate Std. Error t value                              Pr(>|t|)
## TREATMENT   -0.089    0.00724   -12.3 0.0000000000000000000000000000000000272
##           CI Lower CI Upper   DF
## TREATMENT   -0.103  -0.0748 5446
```

```r
difference_in_means(missed_payment_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data)
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT    0.012    0.00581    2.06    0.039 0.000606   0.0234 5453
```

```r
difference_in_means(nr_missed_payments_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data)
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0795     0.0482    1.65   0.0991   -0.015    0.174 5268
```

```r
difference_in_means(compliance_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Property",])
```



```
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.216     0.0122   -17.7
##                                                                      Pr(>|t|)
## TREATMENT 0.0000000000000000000000000000000000000000000000000000000000000000343
##           CI Lower CI Upper   DF
## TREATMENT    -0.24   -0.192 2116
```

```r
difference_in_means(missed_payment_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0195     0.0085     2.3   0.0215  0.00288   0.0362 2241
```

```r
difference_in_means(nr_missed_payments_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Property",])
```

```
## Design:  Standard
##           Estimate Std. Error t value       Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT    0.321     0.0514    6.25 0.000000000484     0.22    0.422 2345
```

```r
difference_in_means(compliance_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT 0.000237     0.0149  0.0159    0.987   -0.029   0.0295 763
```

```r
difference_in_means(missed_payment_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   0.0134     0.0109    1.23     0.22 -0.00804   0.0349 731
```

```r
difference_in_means(nr_missed_payments_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Sewage",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   0.0837      0.176   0.476    0.635   -0.262     0.43 593
```

```r
difference_in_means(compliance_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT -0.00388     0.0106  -0.365    0.715  -0.0247   0.0169 1829
```

```r
difference_in_means(missed_payment_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  0.00911    0.00922   0.988    0.323 -0.00897   0.0272 1750
```

```r
difference_in_means(nr_missed_payments_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Head",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   -0.111     0.0778   -1.42    0.156   -0.263   0.0421 1822
```

```r
difference_in_means(compliance_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT  0.00897     0.0215   0.417    0.677  -0.0333   0.0512 720
```

```r
difference_in_means(missed_payment_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT  -0.0116     0.0234  -0.498    0.619  -0.0575   0.0342 700
```

```r
difference_in_means(nr_missed_payments_mean_DiD_5yr ~ TREATMENT,
                    data = dd_data[dd_data$tax=="Vehicle",])
```

```
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## TREATMENT   -0.255      0.195    -1.3    0.193   -0.638    0.129 610
```

```r
##############################################################
## TABLE 4 - FIELD EXPERIMENT. Informational mechanism.
# Good and bad taxpayers, comparison of treatments 1, 2 and
# 4 (pooled) vs. the placebo control group (treatment 0).
# First differences use the value of the dependent variable
# for the pre-treatment period (March 2014).
##############################################################

### Creating first differences outcomes for the field experiment
fieldex$missed_payment_DiD <- fieldex$JUL_2014_ontime - fieldex$MAR_2014_ontime # missed payment
fieldex$web_bill_DiD <- fieldex$july_web_access - fieldex$march_web_access # web access
fieldex$payments_owed_DiD <- fieldex$july_nrbills_owed - fieldex$adeudadas_2014_MAR # nr bills owed

# compliance
fieldex$compliance_march <- ifelse(fieldex$march_ontime==1 & fieldex$adeudadas_2014_MAR==0, 1, 0)
```



```r
fieldex$compliance_july <- ifelse(fieldex$JUL_2014_ontime==1 & fieldex$july_nrbills_owed==0, 1, 0)
fieldex$compliance_DiD <- fieldex$compliance_july - fieldex$compliance_march

# results
difference_in_means(missed_payment_DiD ~ pooled_124_0, data = fieldex)
```

```
## Design:  Standard
##                Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0 -0.000745    0.00868 -0.0858    0.932  -0.0178   0.0163 7227
```

```r
difference_in_means(payments_owed_DiD ~ pooled_124_0, data = fieldex)
```

```
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0   0.0131      0.023   0.568     0.57   -0.032   0.0582 7787
```

```r
difference_in_means(compliance_DiD ~ pooled_124_0, data = fieldex)
```

```
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0 0.0000208    0.00837 0.00249    0.998  -0.0164   0.0164 7298
```

```r
difference_in_means(web_bill_DiD ~ pooled_124_0, data = fieldex)
```

```
## Design:  Standard
##              Estimate Std. Error t value   Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0  -0.0458    0.00798   -5.74 0.00000001  -0.0614  -0.0301 6601
```

```r
##############################################################
## TABLE 6 - FIELD EXPERIMENT: Comparison of effects for good
# and bad taxpayers: difference of the difference of means for
# the comparison of treatments 1, 2 and 4 (pooled) vs. the
# placebo control group (treatment 0).
##############################################################

# Function to test the difference of the differences
comp.eff <- function(dm1, dm2){

  print("Difference in Means 1")
  print(dm1)

  print("Difference in Means 2")
  print(dm2)

  print("#####. Difference in Effects")

  diff <- dm1$coefficients - dm2$coefficients
  se.diff <- sqrt((dm1$std.error^2)+(dm2$std.error^2))
  t.val.diff <- diff/se.diff
  df <- dm1$nobs + dm2$nobs -2
```

```
  # Calculate the p-value
  p_val <- 2 * (1 - pt(abs(t.val.diff), df=df))

  res <- c(diff,se.diff,t.val.diff, p_val)
names(res) <- c("Diff in effects", "SE", "t", "p-value")
print(res)
return(res)
}

comp.eff(difference_in_means(missed_payment_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "good taxpayer")),
         difference_in_means(missed_payment_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0 0.000155     0.0076  0.0204    0.984  -0.0147   0.0151 3036
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0 -0.00208     0.0138  -0.151     0.88  -0.0291   0.0249 4138
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00223         0.01573         0.14189         0.88717

## Diff in effects              SE               t         p-value
##         0.00223         0.01573         0.14189         0.88717
```

```
comp.eff(difference_in_means(web_bill_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "good taxpayer")),
         difference_in_means(web_bill_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value   Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0  -0.0623     0.0128   -4.86 0.00000127  -0.0874  -0.0371 2683
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0  -0.0336     0.0101   -3.33 0.000889  -0.0533  -0.0138 3955
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0287          0.0163         -1.7594          0.0785

## Diff in effects              SE               t         p-value
##         -0.0287          0.0163         -1.7594          0.0785
```

```
##############################################################
## TABLE 7-  NATURAL EXPERIMENT. Income effects. Comparison
```



```
# of winners vs. non-winners: heterogeneous effects of winning
# the lottery by tax bracket.
##############################################################

# Coding tax brackets
a <- 418958
b <- 1047393
c <- 2094784
d <- 41895699

taxes_panel$tax_bracket <- "D"
taxes_panel$tax_bracket[taxes_panel$VALOR_CATASTRALACTUAL<d] <- "C"
taxes_panel$tax_bracket[taxes_panel$VALOR_CATASTRALACTUAL<c] <- "B"
taxes_panel$tax_bracket[taxes_panel$VALOR_CATASTRALACTUAL<b] <- "A"
taxes_panel$tax_bracket[is.na(taxes_panel$VALOR_CATASTRALACTUAL)] <- NA

by_tax_bracket <- taxes_panel %>% dplyr::filter(!is.na(tax_bracket) & t_st==4) %>%
  group_by(tax_bracket) %>%
  dplyr::summarize(
    N_treat = sum(TREATMENT==1),
    N_control = sum(TREATMENT==0),
    missed_treat = mean(missed_payment[TREATMENT==1], na.rm=T),
    missed_control = mean(missed_payment[TREATMENT==0], na.rm=T)
  )

by_tax_bracket
```

```
## # A tibble: 4 x 5
##   tax_bracket N_treat N_control missed_treat missed_control
##   <chr>         <int>     <int>        <dbl>          <dbl>
## 1 A              4531      4563        0.334          0.323
## 2 B               885       895        0.234          0.194
## 3 C               710       672        0.247          0.214
## 4 D                39        45        0.212          0.163
```

```
chisq.test(by_tax_bracket[,c(4,5)])
```

```
##
##  Pearson's Chi-squared test
##
## data:  by_tax_bracket[, c(4, 5)]
## X-squared = 0.004, df = 3, p-value = 1
```

```
# HTEs by Valor Catastral

# 1 year diff in diff setup
dd_data_vc <-  rbind.data.frame(taxes_panel %>% filter(ES_BP==1 & tax=="Property") %>%
                               group_by(CUENTA, TREATMENT, VALOR_CAT2004) %>% dplyr::summarise(
                                 compliance_mean_DiD_1yr =
                                   mean(compliance[t_st>3 & t_st<=6], na.rm=T) -
                                   mean(compliance[t_st<0 & t_st>=(-3)], na.rm=T),
                                 missed_payment_mean_DiD_1yr =
```

```r
                                   mean(missed_payment[t_st>3 & t_st<=6], na.rm=T)-
                                   mean(missed_payment[t_st<0 & t_st>=(-3)], na.rm=T),
                                 nr_missed_payments_mean_DiD_1yr =
                                   mean(cuotas_adeudadas[t_st> 3 & t_st<=6], na.rm=T)-
                                   mean(cuotas_adeudadas[t_st<0 & t_st>=(-3)], na.rm=T)
                               )
)
```

```
## `summarise()` has grouped output by 'CUENTA', 'TREATMENT'. You can override
## using the `.groups` argument.
```

```r
dd_data_vc$high_propvalue <- ifelse(dd_data_vc$VALOR_CAT2004 >
                                 median(dd_data_vc$VALOR_CAT2004, na.rm=T),
                               1, 0)

income_missed <- comp.eff(difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 1)),
         difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0347      0.011    3.15  0.00166   0.0131   0.0563 1032
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0188     0.0134    1.41     0.16 -0.00743   0.0451 1069
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.0159          0.0173          0.9157          0.3599
```

```r
income_missed
```

```
## Diff in effects              SE               t         p-value
##          0.0159          0.0173          0.9157          0.3599
```

```r
income_nrmissed <- comp.eff(difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 1)),
         difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT     0.94      0.048    19.6
##                                                                                Pr(>|t|)
## TREATMENT 0.00000000000000000000000000000000000000000000000000000000000000000000644
##           CI Lower CI Upper  DF
## TREATMENT    0.846     1.03 725
## [1] "Difference in Means 2"
```



```
## Design:  Standard
##           Estimate Std. Error t value                             Pr(>|t|)
## TREATMENT    0.812     0.0667    12.2 0.0000000000000000000000000000000442
##           CI Lower CI Upper   DF
## TREATMENT    0.681    0.943 1133
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.1287          0.0822          1.5661          0.1175
```

```r
income_nrmissed
```

```
## Diff in effects              SE               t         p-value
##          0.1287          0.0822          1.5661          0.1175
```

```r
income_compliance <- comp.eff(difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 1)),
         difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                             data = filter(dd_data_vc, high_propvalue == 0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.505     0.0219   -23.1
##                                                                                              Pr(
## TREATMENT 0.000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper  DF
## TREATMENT   -0.548   -0.462 821
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.418      0.022     -19
##                                                                              Pr(>|t|)
## TREATMENT 0.000000000000000000000000000000000000000000000000000000000000000000184
##           CI Lower CI Upper  DF
## TREATMENT   -0.462   -0.375 856
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        -0.08671         0.03103        -2.79414         0.00525
```

```r
income_compliance
```

```
## Diff in effects              SE               t         p-value
##        -0.08671         0.03103        -2.79414         0.00525
```

```r
###################################################################
## TABLE 8. NATURAL EXPERIMENT. Habit effects. Winners vs.
# non-winners: heterogeneous treatment effects by time since
# winning (heterogeneous effects; 1, 2 and 3 years).
###################################################################

# Yr 1 vs 2
comp.eff(difference_in_means(missed_payment_mean_DiD_1yr ~ TREATMENT,
```

```r
                    data = dd_data1),
difference_in_means(missed_payment_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data1))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0202     0.0065    3.11  0.00191  0.00745    0.033 5307
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|)  CI Lower CI Upper   DF
## TREATMENT   0.0125    0.00688    1.82   0.0688 -0.000964    0.026 4985
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00767         0.00947         0.80971         0.41812

## Diff in effects              SE               t         p-value
##         0.00767         0.00947         0.80971         0.41812
```

```r
comp.eff(difference_in_means(nr_missed_payments_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1),
difference_in_means(nr_missed_payments_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data1))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value                Pr(>|t|) CI Lower CI Upper
## TREATMENT    0.349     0.0389    8.98 0.000000000000000000391    0.273    0.425
##             DF
## TREATMENT 4410
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT    0.105     0.0404    2.61  0.00913   0.0261    0.184 4931
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##       0.2438890       0.0560550       4.3508898       0.0000137

## Diff in effects              SE               t         p-value
##       0.2438890       0.0560550       4.3508898       0.0000137
```

```r
comp.eff(difference_in_means(compliance_mean_DiD_1yr ~ TREATMENT,
                    data = dd_data1),
difference_in_means(compliance_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data1))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.199    0.00935   -21.3
##
```



```
## TREATMENT 0.000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
##           CI Lower CI Upper   DF
## TREATMENT   -0.218   -0.181 4611
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.116    0.00876   -13.2
##                                               Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT 0.000000000000000000000000000000000000000432   -0.133  -0.0984 4944
## [1] "#####. Difference in Effects"
##  Diff in effects               SE                t          p-value
## -0.0837187274593  0.0128110676674 -6.5348751277173  0.0000000000665

##  Diff in effects               SE                t          p-value
## -0.0837187274593  0.0128110676674 -6.5348751277173  0.0000000000665
```

```r
# Yr 2 vs 3
comp.eff(difference_in_means(missed_payment_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data1),
difference_in_means(missed_payment_mean_DiD_1yr.yr3 ~ TREATMENT,
                    data = dd_data1))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|)  CI Lower CI Upper   DF
## TREATMENT   0.0125    0.00688    1.82   0.0688 -0.000964    0.026 4985
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0042    0.00762   0.551    0.581  -0.0107   0.0191 4555
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00833         0.01027         0.81101         0.41738

## Diff in effects              SE               t         p-value
##         0.00833         0.01027         0.81101         0.41738
```

```r
comp.eff(difference_in_means(nr_missed_payments_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data1),
difference_in_means(nr_missed_payments_mean_DiD_1yr.yr3 ~ TREATMENT,
                    data = dd_data1))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT    0.105     0.0404    2.61  0.00913   0.0261    0.184 4931
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT   0.0126     0.0524    0.24     0.81  -0.0901    0.115 4849
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.0927          0.0661          1.4023          0.1608
```

```
## Diff in effects              SE               t         p-value
##          0.0927          0.0661          1.4023          0.1608
```

```
comp.eff(difference_in_means(compliance_mean_DiD_1yr.yr2 ~ TREATMENT,
                    data = dd_data1),
difference_in_means(compliance_mean_DiD_1yr.yr3 ~ TREATMENT,
                    data = dd_data1))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##           Estimate Std. Error t value
## TREATMENT   -0.116    0.00876   -13.2
##                                          Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT 0.000000000000000000000000000000000000000432   -0.133  -0.0984 4944
## [1] "Difference in Means 2"
## Design:  Standard
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## TREATMENT  0.00212    0.00857   0.247    0.805  -0.0147   0.0189 4844
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.1177          0.0123         -9.6029          0.0000
```

```
## Diff in effects              SE               t         p-value
##         -0.1177          0.0123         -9.6029          0.0000
```

```
##############################################################
## TABLE 9. FIELD EXPERIMENT. Priming knowledge of punishment.
# Good and bad taxpayers, comparison of treatments 3 and 5
# (pooled) vs. the placebo control group (treatment 0).
##############################################################

difference_in_means(missed_payment_DiD ~ pooled_35_0, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0093    0.00869   -1.07    0.285  -0.0263  0.00774 7252
```

```
difference_in_means(payments_owed_DiD ~ pooled_35_0, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  0.00877     0.0227   0.386    0.699  -0.0357   0.0533 7531
```

```
difference_in_means(web_bill_DiD ~ pooled_35_0, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0193    0.00818   -2.36   0.0184  -0.0353 -0.00325 7066
```



```
difference_in_means(compliance_DiD ~ pooled_35_0, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0106    0.00839   -1.26    0.207   -0.027  0.00586 7348
```

```
##############################################################
# TABLE 10. FIELD EXPERIMENT: Comparison of effects for good
# and bad taxpayers: difference of the difference in means for
# the comparison of treatments 3 and 5 (priming knowledge of
# sanctions, pooled) vs. the placebo control group (treatment 0).
##############################################################

comp.eff(difference_in_means(missed_payment_DiD ~ pooled_35_0,
                                data = filter(fieldex, type == "good taxpayer")),
difference_in_means(missed_payment_DiD ~ pooled_35_0,
                    data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0 -0.00862    0.00776   -1.11    0.266  -0.0238  0.00658 3188
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0131     0.0137  -0.955    0.339  -0.0399   0.0138 4083
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00446         0.01574         0.28355         0.77676
```

```
## Diff in effects              SE               t         p-value
##         0.00446         0.01574         0.28355         0.77676
```

```
comp.eff(difference_in_means(web_bill_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "good taxpayer")),
difference_in_means(web_bill_DiD ~ pooled_35_0,
                    data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0258     0.0133   -1.93   0.0533   -0.052 0.000372 3004
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0134     0.0102   -1.31    0.189  -0.0334   0.0066 4079
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0124          0.0168         -0.7388          0.4600
```

```
## Diff in effects              SE               t         p-value
##         -0.0124          0.0168         -0.7388          0.4600
```

```r
################################################################
# TABLE 11. FIELD EXPERIMENT. Positive vs negative incentives.
# Good and bad taxpayers, comparison of treatments 1, 2 and 4
# (positive incentives, pooled) vs 3 and 5 (negative incentives, pooled).
# Test using compliance conditional on significant effects for missed payment,
# number of payments owed or total debt.
################################################################

difference_in_means(missed_payment_DiD ~ pooled_124_35, data = fieldex)
```

```
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper    DF
## pooled_124_35  0.00855    0.00716    1.19    0.232 -0.00548   0.0226 14260
```

```r
difference_in_means(payments_owed_DiD ~ pooled_124_35, data = fieldex)
```

```
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper    DF
## pooled_124_35  0.00429     0.0195    0.22    0.826  -0.0339   0.0425 14356
```

```r
difference_in_means(web_bill_DiD ~ pooled_124_35, data = fieldex)
```

```
## Design:  Standard
##               Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper    DF
## pooled_124_35  -0.0265    0.00633   -4.19 0.0000278  -0.0389  -0.0141 14290
```

```r
################################################################
# TABLE 12. FIELD EXPERIMENT. Comparison of effects of positive
# vs negative incentives for good and bad taxpayers: difference
# of the difference in means for the comparison of treatments
# 1, 2 and 4 (pooled) and 3 and 5 (pooled).
################################################################

comp.eff(difference_in_means(missed_payment_DiD ~ pooled_124_35,
                             data = filter(fieldex, type == "good taxpayer")),
         difference_in_means(missed_payment_DiD ~ pooled_124_35,
                             data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_35  0.00878    0.00636    1.38    0.168 -0.00369   0.0213 5972
## [1] "Difference in Means 2"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_35    0.011     0.0113   0.978    0.328   -0.011   0.0331 8252
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##        -0.00223         0.01293        -0.17255         0.86301

## Diff in effects              SE               t         p-value
##        -0.00223         0.01293        -0.17255         0.86301
```



```r
comp.eff(difference_in_means(payments_owed_DiD ~ pooled_124_35,
                             data = filter(fieldex, type == "good taxpayer")),
         difference_in_means(payments_owed_DiD ~ pooled_124_35,
                             data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower  CI Upper   DF
## pooled_124_35  -0.0082    0.00404   -2.03   0.0427  -0.0161 -0.000271 5818
## [1] "Difference in Means 2"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_35   0.0172     0.0332   0.519    0.603  -0.0478   0.0823 8335
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0254          0.0334         -0.7609          0.4467

## Diff in effects              SE               t         p-value
##         -0.0254          0.0334         -0.7609          0.4467
```

```r
comp.eff(difference_in_means(web_bill_DiD ~ pooled_124_35,
                             data = filter(fieldex, type == "good taxpayer")),
         difference_in_means(web_bill_DiD ~ pooled_124_35,
                             data = filter(fieldex, type == "bad taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_35  -0.0365     0.0101   -3.61 0.000313  -0.0563  -0.0166 5884
## [1] "Difference in Means 2"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_35  -0.0202    0.00806    -2.5   0.0124   -0.036 -0.00436 8363
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0163          0.0129         -1.2603          0.2076

## Diff in effects              SE               t         p-value
##         -0.0163          0.0129         -1.2603          0.2076
```

```r
################################################################
# TABLE 13. FIELD EXPERIMENT. Marginal taxpayers. Good taxpayers.
# Heterogeneous effects, taxpayers at risk. Comparison of
# treatment effect of 1, 2 and 4 (pooled) vs control
# (A-Information about the tax lottery), on one test and 3
# and 5 (pooled) vs control on another (B-Information about sanctions).
################################################################

# Identifying good taxpayers with history of debt
# Ever owed a bill since March 2009?
names(fieldex)[grepl("adeudadas", names(fieldex))]
```

```
##  [1] "adeudadas_2014_MAR" "adeudadas_2009_MAR" "adeudadas_2009_JUL"
##  [4] "adeudadas_2009_NOV" "adeudadas_2010_MAR" "adeudadas_2010_JUL"
##  [7] "adeudadas_2010_NOV" "adeudadas_2011_MAR" "adeudadas_2011_JUL"
## [10] "adeudadas_2011_NOV" "adeudadas_2012_MAR" "adeudadas_2012_JUL"
## [13] "adeudadas_2012_NOV" "adeudadas_2013_MAR" "adeudadas_2013_JUL"
## [16] "adeudadas_2013_NOV"
```

```
sum_bills_owed <- apply(fieldex[,grepl("adeudadas", names(fieldex))], 1, sum)
fieldex$goodtp_at_risk <- ifelse(sum_bills_owed>0, 1, 0)
table(fieldex$goodtp_at_risk[fieldex$type=="good taxpayer"])
```

```
##
##     0     1
## 10665  4119
```

```
### A. Information on the lottery - Heterogeneous effects for taxpayers at risk
comp.eff(difference_in_means(missed_payment_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==1)),
         difference_in_means(missed_payment_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_124_0  0.00963     0.0194   0.496     0.62  -0.0284   0.0477 771
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0 -0.00331    0.00764  -0.433    0.665  -0.0183   0.0117 2287
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.0129          0.0208          0.6205          0.5350

## Diff in effects              SE               t         p-value
##          0.0129          0.0208          0.6205          0.5350
```

```
comp.eff(difference_in_means(web_bill_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==1)),
         difference_in_means(web_bill_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_124_0  -0.0571     0.0248    -2.3   0.0218   -0.106 -0.00836 712
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0   -0.064      0.015   -4.28 0.0000197  -0.0934  -0.0347 1971
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00691         0.02901         0.23817         0.81176
```



```
## Diff in effects              SE               t         p-value
##         0.00691         0.02901         0.23817         0.81176
```

```
### B. Information on sanctions - Heterogeneous effects for taxpayers at risk
comp.eff(difference_in_means(missed_payment_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==1)),
         difference_in_means(missed_payment_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_35_0    0.023     0.0192     1.2     0.23  -0.0146   0.0607 745
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0    -0.02    0.00801   -2.49   0.0128  -0.0357 -0.00425 2556
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.0430          0.0208          2.0674          0.0388

## Diff in effects              SE               t         p-value
##          0.0430          0.0208          2.0674          0.0388
```

```
comp.eff(difference_in_means(web_bill_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==1)),
         difference_in_means(web_bill_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "good taxpayer" & goodtp_at_risk==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_35_0  -0.0392     0.0255   -1.54    0.124  -0.0892   0.0108 759
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0216     0.0156   -1.38    0.168  -0.0523  0.00909 2238
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0176          0.0299         -0.5903          0.5550

## Diff in effects              SE               t         p-value
##         -0.0176          0.0299         -0.5903          0.5550
```

```
##################################################################
# TABLE 14. FIELD EXPERIMENT. Marginal taxpayers. Bad taxpayers.
# Heterogeneous effects, salvageable taxpayers. Comparison of treatment
# effect of 1, 2 and 4 (pooled) vs control (A-Information about the
# tax lottery), on one test and 3 and 5 (pooled) vs control on
# another (B-Information about sanctions). Test using compliance
# conditional on significant effects for missed payment, number
# of payments owed or total debt.
```

```r
################################################################

# Identifying bad taxpayers not too in debt
# Ever owed a bill since March 2009?
fieldex$salvageable_btp <- ifelse(fieldex$adeudadas_2014_MAR>3, 0, 1)

### A. Information on the lottery - Heterogeneous effects for taxpayers at risk

# Missed payments
comp.eff(difference_in_means(missed_payment_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(missed_payment_DiD ~ pooled_124_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0  0.00433     0.0174   0.248    0.804  -0.0299   0.0385 3119
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_124_0  -0.0248     0.0151   -1.64    0.101  -0.0544   0.0048 986
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.0291          0.0231          1.2624          0.2068

## Diff in effects              SE               t         p-value
##          0.0291          0.0231          1.2624          0.2068
```

```r
# Web access
comp.eff(difference_in_means(web_bill_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(web_bill_DiD ~ pooled_124_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0  -0.0398     0.0128   -3.11  0.00188  -0.0649  -0.0147 2999
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_124_0  -0.0148      0.011   -1.34     0.18  -0.0364  0.00686 926
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0250          0.0169         -1.4803          0.1388

## Diff in effects              SE               t         p-value
##         -0.0250          0.0169         -1.4803          0.1388
```



```r
# Nr of payments owed
comp.eff(difference_in_means(payments_owed_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(payments_owed_DiD ~ pooled_124_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0   0.0334     0.0249    1.34     0.18  -0.0154   0.0821 4118
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0   0.0112      0.144  0.0781    0.938   -0.271    0.293 1026
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          0.0222          0.1458          0.1520          0.8792

## Diff in effects              SE               t         p-value
##          0.0222          0.1458          0.1520          0.8792
```

```r
# Compliance
comp.eff(difference_in_means(compliance_DiD ~ pooled_124_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(compliance_DiD ~ pooled_124_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##               Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_124_0 -0.000557     0.0171 -0.0326    0.974  -0.0341    0.033 3185
## [1] "Difference in Means 2"
## Design:  Standard
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_124_0 -0.00379    0.00611   -0.62    0.535  -0.0158   0.0082 884
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00323         0.01817         0.17782         0.85887

## Diff in effects              SE               t         p-value
##         0.00323         0.01817         0.17782         0.85887
```

```r
### B. Information on sanctions - Heterogeneous effects for taxpayers at risk

# Missed payments
comp.eff(difference_in_means(missed_payment_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(missed_payment_DiD ~ pooled_35_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0143     0.0173  -0.826    0.409  -0.0481   0.0196 3045
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0154     0.0152   -1.01    0.311  -0.0452   0.0144 1000
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00113         0.02299         0.04900         0.96092

## Diff in effects              SE               t         p-value
##         0.00113         0.02299         0.04900         0.96092
```

```
# Web access
comp.eff(difference_in_means(web_bill_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(web_bill_DiD ~ pooled_35_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0224     0.0128   -1.74   0.0813  -0.0475  0.00278 3035
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0   0.0118     0.0117    1.01    0.311  -0.0111   0.0347 1069
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0342          0.0173         -1.9722          0.0486

## Diff in effects              SE               t         p-value
##         -0.0342          0.0173         -1.9722          0.0486
```

```
# Nr of payments owed
comp.eff(difference_in_means(payments_owed_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(payments_owed_DiD ~ pooled_35_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  0.00328     0.0272   0.121    0.904    -0.05   0.0566 4610
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_35_0   0.0533      0.137   0.389    0.697   -0.216    0.322 897
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##          -0.050           0.140          -0.358           0.720
```



```
## Diff in effects              SE               t         p-value
##          -0.050           0.140          -0.358           0.720
```

```
# Compliance
comp.eff(difference_in_means(compliance_DiD ~ pooled_35_0,
                             data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==1)),
difference_in_means(compliance_DiD ~ pooled_35_0,
                    data = filter(fieldex, type == "bad taxpayer" & salvageable_btp==0)))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_35_0  -0.0217     0.0169   -1.28    0.201  -0.0549   0.0115 3110
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## pooled_35_0 -0.00122    0.00636  -0.191    0.848  -0.0137   0.0113 987
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0204          0.0181         -1.1301          0.2585

## Diff in effects              SE               t         p-value
##         -0.0204          0.0181         -1.1301          0.2585
```

```
##############################################################
# TABLE 15. FIELD EXPERIMENT. Good and bad taxpayers. Social
# vs individual rewards. Comparison of treatments 1, 2 (pooled) vs 4.
# Test using compliance conditional on significant effects for
# missed payment, number of payments owed or total debt.
##############################################################

# Missed payments
difference_in_means(missed_payment_DiD ~ pooled_12_4, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4  -0.0312     0.0101   -3.09  0.00201   -0.051  -0.0114 7128
```

```
# Web access
difference_in_means(web_bill_DiD ~ pooled_12_4, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4 -0.00308    0.00859  -0.358     0.72  -0.0199   0.0138 7185
```

```
# Nr of payments owed
difference_in_means(payments_owed_DiD ~ pooled_12_4, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4  -0.0507      0.028   -1.81   0.0707   -0.106  0.00428 6480
```

```r
# Compliance
difference_in_means(compliance_DiD ~ pooled_12_4, data = fieldex)
```

```
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4  -0.0318    0.00975   -3.26   0.0011  -0.0509  -0.0127 7169
```

```r
###############################################################
# TABLE 16. FIELD EXPERIMENT. Social (4) vs individual rewards
# (1 and 2, pooled), comparison of effect between good and bad taxpayers.
###############################################################

# Missed payments
comp.eff(difference_in_means(missed_payment_DiD ~ pooled_12_4,
                             data = filter(fieldex, type == "bad taxpayer")),
difference_in_means(missed_payment_DiD ~ pooled_12_4,
                    data = filter(fieldex, type == "good taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4   -0.044      0.016   -2.74  0.00612  -0.0754  -0.0125 4088
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4  -0.0158    0.00873   -1.81   0.0701  -0.0329   0.0013 2990
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0282          0.0183         -1.5433          0.1228

## Diff in effects              SE               t         p-value
##         -0.0282          0.0183         -1.5433          0.1228
```

```r
# Web access
comp.eff(difference_in_means(web_bill_DiD ~ pooled_12_4,
                             data = filter(fieldex, type == "bad taxpayer")),
difference_in_means(web_bill_DiD ~ pooled_12_4,
                    data = filter(fieldex, type == "good taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4 -0.00168     0.0112   -0.15    0.881  -0.0237   0.0203 4134
## [1] "Difference in Means 2"
## Design:  Standard
##             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_12_4 -0.00456     0.0133  -0.342    0.732  -0.0307   0.0215 3024
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00287         0.01741         0.16503         0.86893

## Diff in effects              SE               t         p-value
##         0.00287         0.01741         0.16503         0.86893
```



```r
###############################################################
# TABLE 17. FIELD EXPERIMENT. Good and bad taxpayers. Social vs
# individual sanctions. Comparison of treatments 3 vs 5. Test
# using compliance conditional on significant effects for missed
# payment, number of payments owed or total debt.
###############################################################

# Missed payments
difference_in_means(missed_payment_DiD ~ pooled_3_5, data = fieldex)
```

```
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5  0.00949     0.0101   0.935     0.35  -0.0104   0.0294 7123
```

```r
# Web access
difference_in_means(web_bill_DiD ~ pooled_3_5, data = fieldex)
```

```
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5   0.0211    0.00929    2.27   0.0234  0.00286   0.0393 7171
```

```r
# Nr of payments owed
difference_in_means(payments_owed_DiD ~ pooled_3_5, data = fieldex)
```

```
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5  0.00205     0.0271  0.0755     0.94  -0.0512   0.0553 5162
```

```r
# Compliance
difference_in_means(compliance_DiD ~ pooled_3_5, data = fieldex)
```

```
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5  0.00888    0.00983   0.903    0.366  -0.0104   0.0282 7165
```

```r
###############################################################
# TABLE 18. FIELD EXPERIMENT. Social vs individual sanctions.
# Comparison of effects for good and bad taxpayers.
# Comparison of treatments 3 vs 5. Test using compliance conditional
# on significant effects for missed payment, number of payments owed or total debt.
###############################################################

# Missed payments
comp.eff(difference_in_means(missed_payment_DiD ~ pooled_3_5,
                             data = filter(fieldex, type == "bad taxpayer")),
difference_in_means(missed_payment_DiD ~ pooled_3_5,
                    data = filter(fieldex, type == "good taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
```

```
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5  0.00983     0.0158   0.623    0.533  -0.0211   0.0408 4160
## [1] "Difference in Means 2"
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5  0.00623    0.00925   0.673    0.501  -0.0119   0.0244 2965
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         0.00361         0.01829         0.19711         0.84375

## Diff in effects              SE               t         p-value
##         0.00361         0.01829         0.19711         0.84375
```

```
# Web access
comp.eff(difference_in_means(web_bill_DiD ~ pooled_3_5,
                             data = filter(fieldex, type == "bad taxpayer")),
difference_in_means(web_bill_DiD ~ pooled_3_5,
                    data = filter(fieldex, type == "good taxpayer")))
```

```
## [1] "Difference in Means 1"
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5   0.0143     0.0116    1.23    0.218 -0.00845    0.037 4205
## [1] "Difference in Means 2"
## Design:  Standard
##            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper   DF
## pooled_3_5   0.0316     0.0152    2.08   0.0378  0.00178   0.0615 2963
## [1] "#####. Difference in Effects"
## Diff in effects              SE               t         p-value
##         -0.0174          0.0191         -0.9073          0.3643

## Diff in effects              SE               t         p-value
##         -0.0174          0.0191         -0.9073          0.3643
```

```
################################################################
# PAP 2 - SURVEY EXPERIMENT
load("data/survey_data.Rda")
################################################################

################################################################
# Social vs Individual Benefit
################################################################

# Pooling the individual benefits treatments (questionaire versions 2 and 3)
# and recoding as treatment=1.
survey_data$social_individual <- ifelse((survey_data$treatment==2|
                                         survey_data$treatment==3),1,NA)
# Recoding the social benefits treatment as 0.
survey_data$social_individual <- ifelse((survey_data$treatment==1),0,
                                        survey_data$social_individual)


# Outcome: "Policies that reward good taxpayers are a waste of money"
```



```
# totally disagree (0) - totally agree (10)
ben1 <- difference_in_means(S1p1 ~ social_individual, data = survey_data)

# Outcome: "It is worth it to be up to date with ones taxes"
# totally disagree (0) - totally agree (10)
ben2 <- difference_in_means(S1p3 ~ social_individual, data = survey_data)

social_individual <- rbind.data.frame(ben1, ben2)
rownames(social_individual) <- c("Rewards are waste of money",
                                 "Worth it to be up to date")

# Adding p-value adjustments
social_individual <- social_individual[order(social_individual[,6], decreasing=F),]
# Ordering p-values in decreasing order
ordered.ps <- social_individual[, 6]
# Building reference vector to compare to ordered p-values
FDR_reference <- .05*(1:length(ordered.ps)/length(ordered.ps))
# Comparing p-values to refernce vector
FDR <- as.data.frame(cbind(ordered.ps, FDR_reference, ordered.ps<=FDR_reference))
FDR[,3] <- as.numeric(FDR[,3])
if (sum(FDR[,3])>0){
  fdr <- which(FDR[,1]==max(FDR[,1][FDR[,3]==1]))
  FDR[,4] <- c(rep("reject null", fdr), rep("do not reject", nrow(FDR)-fdr))}

if (sum(FDR[,3])==0){
  FDR[,4] <- rep("do not reject", nrow(FDR))}

Bonferroni_reference <- rep(.05/nrow(FDR), nrow(FDR))
Bonferroni_reject <- ifelse(ordered.ps<=Bonferroni_reference,
                            "reject null", "do not reject")

social_individual <- cbind(social_individual, FDR[,c(2,4)],
                           Bonferroni_reference,
                           Bonferroni_reject)

names(social_individual)[17] <- "FDR_reject"
social_individual
```

```
##                            coefficients std.error  df nobs statistic p.value
## Worth it to be up to date       -0.0518     0.234 425  830    -0.221   0.825
## Rewards are waste of money      -0.0385     0.273 445  825    -0.141   0.888
##                            conf.low conf.high alpha              term outcome
## Worth it to be up to date    -0.512     0.408  0.05 social_individual    S1p3
## Rewards are waste of money   -0.575     0.498  0.05 social_individual    S1p1
##                            condition2 condition1   vcov   design FDR_reference
## Worth it to be up to date           1          0 0.0548 Standard         0.025
## Rewards are waste of money          1          0 0.0746 Standard         0.050
##                               FDR_reject Bonferroni_reference Bonferroni_reject
## Worth it to be up to date  do not reject                0.025     do not reject
## Rewards are waste of money do not reject                0.025     do not reject
```

```
################################################################
# Discretionary vs lottery allocation of benefits
################################################################

# Generating discretionary benefits dummy where surveys with the discretionary version
# of the survey are 1.
survey_data$treat_discretion <- ifelse((survey_data$treatment==4),1,0)

# Outcome: "In Montevideo, rewards for good taxpayers go to the same people as always"
# totally disagree (0) - totally agree (10)
discretion1 <- difference_in_means(S1p4 ~ treat_discretion,
                                   data = survey_data)
discretion1
```

```
## Design:  Standard
##                  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## treat_discretion   0.0478      0.322   0.148    0.882   -0.587    0.682 223
```

```
# Outcome: "Policies that reward good taxpayers are a waste of money"
# totally disagree (0) - totally agree (10)
discretion2 <- difference_in_means(S1p1 ~ treat_discretion,
                                   data = survey_data)
discretion2
```

```
## Design:  Standard
##                  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## treat_discretion  -0.0315      0.245  -0.129    0.898   -0.513     0.45 330
```

```
# Outcome: "It is worth it to be up to date with ones taxes"
# totally disagree (0) - totally agree (10)
discretion3 <- difference_in_means(S1p3 ~ treat_discretion,
                                   data = survey_data)
discretion3
```

```
## Design:  Standard
##                  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## treat_discretion  -0.0726      0.194  -0.373    0.709   -0.455     0.31 349
```

```
# Outcome: "In general, the municipal government does a good job"
# totally disagree (0) - totally agree (10)
discretion4 <- difference_in_means(S1p2 ~ treat_discretion,
                                   data = survey_data)
discretion4
```

```
## Design:  Standard
##                  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## treat_discretion    -0.21      0.194   -1.08    0.282   -0.592    0.173 342
```

```
# Outcome: "How would you classify the taxes that the municipal
# government charges in gene
# very just (1) - not just at all (4)
discretion5 <- difference_in_means(S1p5 ~ treat_discretion,
                                   data = survey_data)
discretion5
```



```
## Design:  Standard
##                  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## treat_discretion  -0.0249     0.0482  -0.516    0.606    -0.12     0.07 350
```

```
discretion <- rbind.data.frame(discretion1, discretion2, discretion3, discretion4, discretion5)
rownames(discretion) <- c("Rewards go to the same people as always",
                          "Rewards are waste of money",
                          "Worth it to be up to date",
                          "Mun.gov. does a good job",
                          "Mun. taxes are just")


# Adding p-value adjustments
discretion <- discretion[order(discretion[,6], decreasing=F),]
# Ordering p-values in decreasing order
ordered.ps <- discretion[,6]
# Building reference vector to compare to ordered p-values
FDR_reference <- .05*(1:length(ordered.ps)/length(ordered.ps))
# Comparing p-values to referece vector
FDR <- as.data.frame(cbind(ordered.ps, FDR_reference, ordered.ps<=FDR_reference))
FDR[,3] <- as.numeric(FDR[,3])
if (sum(FDR[,3])>0){
  fdr <- which(FDR[,1]==max(FDR[,1][FDR[,3]==1]))
  FDR[,4] <- c(rep("reject null", fdr), rep("do not reject", nrow(FDR)-fdr))}

if (sum(FDR[,3])==0){
  FDR[,4] <- rep("do not reject", nrow(FDR))}

Bonferroni_reference <- rep(.05/nrow(FDR), nrow(FDR))
Bonferroni_reject <- ifelse(ordered.ps<=Bonferroni_reference,"reject null", "do not reject")

discretion <- cbind(discretion, FDR[,c(2,4)],
                          Bonferroni_reference,
                          Bonferroni_reject)

names(discretion)[17] <- "FDR_reject"
discretion
```

```
##                                         coefficients std.error  df nobs
## Mun.gov. does a good job                     -0.2096    0.1944 342 1798
## Mun. taxes are just                          -0.0249    0.0482 350 1785
## Worth it to be up to date                    -0.0726    0.1944 349 1767
## Rewards go to the same people as always       0.0478    0.3220 223 1204
## Rewards are waste of money                   -0.0315    0.2450 330 1741
##                                         statistic p.value conf.low conf.high
## Mun.gov. does a good job                   -1.078   0.282   -0.592     0.173
## Mun. taxes are just                        -0.516   0.606   -0.120     0.070
## Worth it to be up to date                  -0.373   0.709   -0.455     0.310
## Rewards go to the same people as always     0.148   0.882   -0.587     0.682
## Rewards are waste of money                 -0.129   0.898   -0.513     0.450
##                                         alpha             term outcome
## Mun.gov. does a good job                 0.05 treat_discretion    S1p2
## Mun. taxes are just                      0.05 treat_discretion    S1p5
## Worth it to be up to date                0.05 treat_discretion    S1p3
```

```
## Rewards go to the same people as always  0.05 treat_discretion    S1p4
## Rewards are waste of money               0.05 treat_discretion    S1p1
##                                        condition2 condition1    vcov   design
## Mun.gov. does a good job                        1          0 0.03778 Standard
## Mun. taxes are just                             1          0 0.00233 Standard
## Worth it to be up to date                       1          0 0.03779 Standard
## Rewards go to the same people as always         1          0 0.10368 Standard
## Rewards are waste of money                      1          0 0.06002 Standard
##                                        FDR_reference    FDR_reject
## Mun.gov. does a good job                        0.01 do not reject
## Mun. taxes are just                             0.02 do not reject
## Worth it to be up to date                       0.03 do not reject
## Rewards go to the same people as always         0.04 do not reject
## Rewards are waste of money                      0.05 do not reject
##                                        Bonferroni_reference Bonferroni_reject
## Mun.gov. does a good job                               0.01     do not reject
## Mun. taxes are just                                    0.01     do not reject
## Worth it to be up to date                              0.01     do not reject
## Rewards go to the same people as always                0.01     do not reject
## Rewards are waste of money                             0.01     do not reject
```

```
##############################################################
# Fines and charges vs. benefits of tax holidays
##############################################################

# Creating dataframe with relevant treatments and outcomes.
# Keeping outcomes we want for benefits
ben <- survey_data[(survey_data$treatment!=4),names(survey_data) %in%
                     c("treatment", "S1p2", "S1p3","S1p5" )]
names(ben)
```

```
## [1] "S1p2"      "S1p3"      "S1p5"      "treatment"
```

```
# Benefits pooled (S1p2) (A, B and C)
ben$benefits_punishments <- 1

# versus fines and charges pooled (M1p3)  (A, B , C and D)
fin <- survey_data[,names(survey_data) %in% c("treatment", "M1p2", "M1p3","M1p6")]
names(fin)
```

```
## [1] "M1p2"      "M1p3"      "M1p6"      "treatment"
```

```
# Pooling punishments
fin$benefits_punishments <- 0

# For three questions the outcomes are the same in the punishments and benefits conditions
# but have different survey question numbers. Here we rename the variables so
# that we can bind the datasets into one.
names(ben)
```

```
## [1] "S1p2"                 "S1p3"                 "S1p5"
## [4] "treatment"            "benefits_punishments"
```



```
names(ben)[1:3] <- c("M1p3", "M1p2", "M1p6")
pooled <- rbind(ben,fin)


# Outcome: "In general, the municipal government does a good job"
# totally disagree (0) - totally agree (10)
benefits_punishments1 <- difference_in_means(M1p3 ~ benefits_punishments, data = pooled)
benefits_punishments1
```

```
## Design:  Standard
##                      Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper
## benefits_punishments   -0.202     0.0965    -2.1    0.036   -0.392  -0.0132
##                        DF
## benefits_punishments 3389
```

```
# Outcome: "In Montevideo, it is worth it to be up to date on ones taxes"
# totally disagree (0) - totally agree (10)
benefits_punishments2 <- difference_in_means(M1p2 ~ benefits_punishments, data = pooled)
benefits_punishments2
```

```
## Design:  Standard
##                      Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper
## benefits_punishments   -0.379     0.0961   -3.94 0.0000837   -0.567    -0.19
##                        DF
## benefits_punishments 3072
```

```
# Outcome: "How would you classify the taxes that the municipal government charges?"
# very just (1) - not just at all (4)
benefits_punishments3 <- difference_in_means(M1p6 ~ benefits_punishments, data = pooled)
benefits_punishments3
```

```
## Design:  Standard
##                      Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper
## benefits_punishments   0.0387     0.0243    1.59    0.112 -0.00898   0.0863
##                        DF
## benefits_punishments 3249
```

```
benefits_punishments <- rbind.data.frame(benefits_punishments1,
                                         benefits_punishments2,
                                         benefits_punishments3)
rownames(benefits_punishments) <- c("Mun. gov. does a good job",
                                    "Worth it to be up to date",
                                    "Mun. taxes are just")

# Adding p-value adjustments
benefits_punishments <- benefits_punishments[order(benefits_punishments[,6], decreasing=F),]
# Ordering p-values in decreasing order
ordered.ps <- benefits_punishments[,6]
# Building reference vector to compare to ordered p-values
FDR_reference <- .05*(1:length(ordered.ps)/length(ordered.ps))
# Comparing p-values to referece vector
```

```r
FDR <- as.data.frame(cbind(ordered.ps, FDR_reference, ordered.ps<=FDR_reference))
FDR[,3] <- as.numeric(FDR[,3])
if (sum(FDR[,3])>0){
  fdr <- which(FDR[,1]==max(FDR[,1][FDR[,3]==1]))
  FDR[,4] <- c(rep("reject null", fdr), rep("do not reject", nrow(FDR)-fdr))}

if (sum(FDR[,3])==0){
  FDR[,4] <- rep("do not reject", nrow(FDR))}

Bonferroni_reference <- rep(.05/nrow(FDR), nrow(FDR))
Bonferroni_reject <- ifelse(ordered.ps<=Bonferroni_reference,"reject null", "do not reject")

benefits_punishments <- cbind(benefits_punishments, FDR[,c(2,4)],
                    Bonferroni_reference,
                    Bonferroni_reject)

names(benefits_punishments)[17] <- "FDR_reject"
benefits_punishments
```

```
##                            coefficients std.error   df nobs statistic   p.value
## Worth it to be up to date      -0.3787    0.0961 3072 3877     -3.94 0.0000837
## Mun. gov. does a good job      -0.2024    0.0965 3389 3927     -2.10 0.0360038
## Mun. taxes are just             0.0387    0.0243 3249 3914      1.59 0.1116414
##                            conf.low conf.high alpha                 term outcome
## Worth it to be up to date -0.56716   -0.1902  0.05 benefits_punishments    M1p2
## Mun. gov. does a good job -0.39155   -0.0132  0.05 benefits_punishments    M1p3
## Mun. taxes are just       -0.00898    0.0863  0.05 benefits_punishments    M1p6
##                            condition2 condition1     vcov   design FDR_reference
## Worth it to be up to date          1          0 0.009242 Standard        0.0167
## Mun. gov. does a good job          1          0 0.009308 Standard        0.0333
## Mun. taxes are just                1          0 0.000591 Standard        0.0500
##                              FDR_reject Bonferroni_reference Bonferroni_reject
## Worth it to be up to date   reject null               0.0167       reject null
## Mun. gov. does a good job do not reject               0.0167     do not reject
## Mun. taxes are just       do not reject               0.0167     do not reject
```

G.5 Reconciliation of registered analyses and final report

In this section, we discuss the results presented in the paper in relation to the analyses registered in the PAP and also discuss several minor deviations. We base the discussion on the registered hypotheses, tests, and outcome measures found in Table 0.3 in the second amendment to our registered PAP (see section G.3 above).[10] This reflects our most up-to-date plan for testing the registered hypotheses, prior to running the tests.

G.5.1 Code for analysis of mock data

In our first amendment to the PAP, we registered analysis using data with the treatment labels reshuffled to allow a mock analysis (Section G.2 in this appendix). In Section G.4, we reproduce this full analysis using the real data, including the accurate treatment labels.

We note here, however, that while the analysis is as pre-registered, we have modernized some of the code. For example, we use functions such as `difference_of_means` from the `R` package `estimatr` (Blair et al., , which was not available at the time we pre-registered our study, rather than the proprietary functions we wrote for the mock analysis.

G.5.2 Natural experiment

For the analysis of persistence of effects over time, we use plot in Figure 2 in the paper, as registered in our mock code. For readibility and because of minimal variance in the pre-treatment means across the treatment and control groups, in Figure 2 we present differences in levels rather than differences in differences. The full set of results of the difference-in-differences analyses is available in section G.4. See also Table A3, which shows p-values from the one year difference-in-difference analysis for our three registered outcomes. All our results are similar with analysis in levels or first differences. As also discussed in Section B2, we were not able to gather data (in particular from our survey) that would allow us to conduct tests for the full set of pre-registered hypotheses using the natural experiment. Our adjustment

[10]Table 0.3 in the second amendment revises Table 7.3 in our original PAP and reflects choices made in building the mock code for our first amendment.

for multiple comparisons in Table A3 uses all registered tests that we were able to conduct and that were discussed on pp. 43-44 of the original PAP.

G.5.3 Field experiment

For the field experiment, the only discrepancy we have found between our registered analysis and the analysis reported in the paper is that we did not pre-specify adjustment for differing probabilities of treatment assignment among eligible and ineligible taxpayers, when pooling results for both taxpayer types. The reason is that we initially specified an experiment with only a placebo control group (Table 3.2 in the original PAP) and had equal numbers of eligible and ineligible taxpayers assigned to each treatment condition. Our pre-analysis plan thus registered comparisons of outcomes in each treatment condition to those in the placebo control group, separately for eligible and ineligible taxpayers. The municipality subsequently allowed us to access payment data for an additional random sample of eligible and ineligible taxpayers who were not exposed to any of our informational flyers, allowing for a pure control condition and the full experimental design depicted in Online Appendix Table A4. This however introduced probabilities of assignments to the treatment conditions that differed for eligible and ineligible taxpayers, which we failed to register in a subsequent amendment.

In the analysis reported in Figure 7 of the paper, we adjust for the unequal probabilities of treatment assignment across eligible and ineligible taxpayers using inverse probability weights. However, in the Online Appendix we also present several other strategies for estimating average treatment effects. In Appendix Table A5, we present results of a block difference-of-means approach, in which we estimate effects separately for eligible and ineligible taxpayers and then take a weighted average of the block-specific estimated effects, weighting by the relative size of each taxpayer block in the study group. In Figure A16, we estimate effects of the information relative to the placebo control (rather than the pure control), pooling across ineligible and eligible taxpayers. And Figure 7 in the paper presents results of comparisons to the placebo control (as registered) as well as the pure control group.