

How best to nudge taxpayers?

The impact of message simplification and descriptive social norms on payment rates in a central London local authority

Peter John*, Toby Blume†

Abstract: Behavioral insights or nudges have yielded great benefits for today's public administrators by improving the quality of official messages and increasing revenue flows. In the absence of a large number of studies suitable for meta-analysis, less is known about the external validity of these interventions, their range of impact, and the exact matching of the behavioral cue to the client group and context. Factorial designs and repeated interventions, as in the study reported in this article, can add insight through respectively comparing interventions and analyzing their impacts over time. This randomized controlled trial tests whether simplification and/or a descriptive social norm can increase payment of local taxes in a central London local authority. In the first wave, a factorial design on a targeted group of residents, simplification increased the number of people paying by four percentage points, whereas the social norm did not change behavior. In wave two of the study, which was carried out across all households, the descriptive social norm backfired, reducing the rate of payment. The heterogeneous nature of the target population and the exact wording of the social norm are discussed as possible reasons for these results.

Keywords: Behavioral insights, Nudges, Social norm, Taxpayer, Payment rate

The decade of the 2010s has seen the frequent use of behavioral insights to encourage citizen compliance with the payment of taxes. Of particular interest is descriptive social norms, which became the poster child of this approach, as shown by the now famous trials carried out by the UK's tax authority, HMRC (Hallsworth, List, Metcalfe, & Vlaev, 2017). At the same time as there have been successes with these kinds of interventions, there are also other occasions when norms do not work or can backfire, which is particularly evident in the energy field where it can highlight non-conformity with desired social outcomes (Cialdini, 2003). There are some null results in the tax field too (e.g. Castro & Scartascini, 2015).

The question for both social scientists and policy-makers is whether they can be more confident in knowing when behavioral interventions, such as social norms, work or not. Rather than assuming there is a toolbox of known effective interventions to offer policy-makers, there are limits to the use of such tools, knowledge of which can help policy-makers target them to where they work best. Ideally, it would be preferable to have very large sample sizes where effects in different locations could be ascertained, or where treatment designs can be varied in their precise delivery and compared in meta-analyses. Because of the tendencies toward customized designs with specific client groups, which are not repeated, knowledge about internal and external validity can only be acquired gradually by drawing conclusions from each study. The study reported here aims to add to the knowledge base in this vein, seeking to explain null and negative results from a descriptive social norms treatment in the payment of local taxes to a central London local authority.

The study has a number of features that make the drawing of inferences superior to a single study. The experiment compared two interventions,

* Department of Political Economy, King's College London;

† Social Engine, www.social-engine.co.uk
Address correspondence to Peter John at
(peter.john@kcl.ac.uk)

Copyright: © 2018. The authors license this article under the terms of the Creative Commons Attribution 4.0 International License.

one a simplification of the documentation of the tax reminder, the other a social norms message. The simplification worked, which gives confidence about the delivery of the treatments and their likely impact on taxpayers. The second is the factorial design, which allowed for a comparison of the social norms' message across both simplified and non-simplified designs (and vice versa). The third is the repetition of the social norms' treatment for all residents in the borough, a year after the first intervention, which allowed for a second test on a larger number of taxpayers. The fourth feature is that some of the taxpayers in the first experiment were part of the second intervention, creating longitudinal aspect to the study for some taxpayers. Such features allow for questions to be asked about the null effect that would not have been possible in a simple treatment versus control design. Also, because of a similar experiment using social norms in another local authority that worked in delivering a treatment effect, it is possible to come to some tentative conclusions about the reasons for the results from this experimental design.

Background and Motivation: Simplification and Social Norms

The use of behavioral insights to improve the quality of public administration has been one of the success stories of behavioral public policy. Whereas behavioral public policy may be designed to have a direct impact on policy outcomes, such as in health or education (Oliver, 2013), its initial expansion was over improving the quality of public administration. A particular focus has been on the high transaction activities, such as revenue collection, which is related to the relative ease at which randomization can occur, the statistical power of interventions with large sample size, and clear financial benefits to the agency from the high volumes of payments as even small percentage changes in outcomes can significantly affect revenue flows. It is no surprise then that one of the first domains for large-scale randomizations of behavioral insights was on HMRC revenue collection in the UK (Hallsworth et al., 2017).

Two kinds of behavioral insights are particularly appropriate for the redesign of communications to improve the payment of taxes and fees. One is simplification, which reduces the cognitive burden of reading official communication, in-

creases the directness of the response, and highlights the key communication, making it more salient that the individual has to act. Reducing the amount of text can allow for better design of the page, and which can allow key messages to be placed where the eye naturally falls when individuals first look at a letter. Reducing psychological friction has been tried with tax benefits (Bhargava & Manoli, 2012, 2015). Simplification, particularly of the language used, has been a consistent theme in the redesign of health interventions (Zarcadoolas, 2011). In the UK, it has been used to redesign reminders for disability car parking badges with good effect (John & Blume, 2017). It would seem to be particularly appropriate for taxation communications because the accountancy and finance professionals in this domain may wish to express the precise nature of the request in a way that is consistent with values of transparency and full disclosure, but where the resultant official documentation can be dense and inaccessible for recipients. The conclusion to draw is that simplification designs are easy wins for local delivery organizations and provide positive effects that are transferable across many domains.

The other effective choice for a transferable intervention is social norms, in particular descriptive social norms. This is a powerful message because it is assumed that people will update their behavior toward the norm out of fear of non-conformity and from social pressure. There is a considerable evidence from across a range of domains, such as energy, littering, voting, and recycling waste, that it works (Allcott, 2011; Cialdini, 2003; Gerber & Rogers, 2009; Schultz, 1999). Tax compliance is analogous to these other outcomes in the sense that the norm of compliance is non-controversial and has a public benefit. The difference is that tax compliance is legally sanctioned, and the conclusion of much of the literature is that the communication of the likelihood of enforcement is more likely to be successful than messages with normative content (Blumenthal, Christian, Slemrod, & Smith, 2001; Coleman, 1996; Iyer, Reckers, & Sanders, 2010). For example, communicating the likelihood of audit has been shown to work, either from RCTs (Kleven, Knudsen, Kreiner, Pedersen, & Saez, 2011; Slemrod, Blumenthal, & Christian, 2001) or in experiments done in the laboratory (Spicer & Thomas, 1982; Tan & Yim, 2012). Moral suasion and appeals on their own do not always work (McGraw & Scholz, 1991; Torgler, 2013), though

they can in the right conditions (see Hasseldine, Hite, James, & Toumi, 2007). Appeals to fairness are sometimes effective (Wenzel, 2006), such as shaming (Coricelli, Rusconi, & Villeval, 2014). Increasing the level of punishment can even reduce compliance if people perceive it to be unfair (Murphy, 2008). There is also evidence that giving feedback to taxpayers affects their behavior (Wenzel, 2005).

Enforcement and social approaches need not be competitors, however. With tax compliance, if the norm is payment and this is high, the social norm may also indicate the probability of being caught in that authorities may be believed to be concentrating on a smaller number of non-payers. Information about the actual numbers of non-payers at a particular moment in time is not usually public (as opposed to the final accounts), so individuals can update their expectations from the new information, which are usually at a lower rate than the revealed figure. The greater awareness of the enforcement agency, and its efficiency, can be consistent with messages that convey a social element to paying fines.

Indicating the potential of social interventions, taxes and fines messaging in a range of jurisdictions have been subjected to social norm tests with positive results (Carpio, 2013; Coleman, 2007; Hallsworth et al., 2017; Kettle, Hernandez, Ruda, & Sanderson, 2016; Sigala, Burgoyne, & Webley, 1999). However, in spite of a growing number of tax experiments (for reviews, see Hallsworth, 2014; Mascagni, 2017), the number of published field experiments showing the positive effects of descriptive social norms (as opposed to moral suasion) still remains relatively few. Moreover, in some fee interventions, norms do not always work, such as when the population is heterogeneous and temporarily resident, as with students (Silva & John, 2017). An experiment in Argentina, conveying whether taxpayers were aware that only three out of ten taxpayers did not pay their tax liabilities, did not work either (Castro & Scartascini, 2015). These contrary findings should not be too surprising because the contexts within which behavioral interventions vary, such as the extent to which information flows in the social network. A large number of studies across a range of jurisdictions would be needed for a meta-analysis and at present the evidence base is not strong enough to provide one. Research does not give a strong indication of exactly when social norms work or not.

Experiment 1: Study Design

The study site was the London Borough of Lambeth, an inner-London local authority on the south bank of the River Thames, three miles (4.8 km) wide and seven miles (11 km) long, with 219,396 residents. The borough is highly heterogeneous: 57.1% have white ethnicity (it is the 11th most diverse local authority in England). Typical of London, there is a large variation in levels of income and wealth. Local tax bills (called the council tax) are sent to individual households, which the council selected as the appropriate unit to measure to the impact of the behavioral interventions. The experiment was run for the 2014-2015 financial year across three wards within the borough. In co-design sessions with the council, a factorial design was agreed that deployed commonly used behavioral insights of simplification and social norms. These designs are captured in Figures 1-3. The exact level of social norm was decided by examining revenue payments across borough, including late payments and those in debt recovery so as to produce a truthful figure.

It was agreed that the experiment would focus solely on accounts that were 'cash payers' (those which were not currently using automated payments such as direct debit), since collection rates among automated payers were already very high. Since the treatments were based on modifying the paper council tax bill, people who paid through e-billing were not considered within the scope of the experiment, as the e-mails they receive are not the same as standard paper bills.

The three wards were selected on the basis of current payment rates, which were below average for the borough, but not at the lowest level. The rationale for this was that they gave the greatest scope for achieving a statistically significant effect with a low ceiling, but where non-payment was not the norm. The wards selected were Thornton, Ferndale, and Brixton Hill, which had collection rates at 7th January 2014 of 84.48, 84.44, and 84.29% respectively. The collection rate for the borough as a whole was 84.83% for the same period. Overall borough collection rates for the previous year, including late-payments and debt recovered, was 95%.

Households in the three wards were randomly allocated into four groups using Excel's random seed generator (rand): the first group received

Figure 1
Simplification treatment

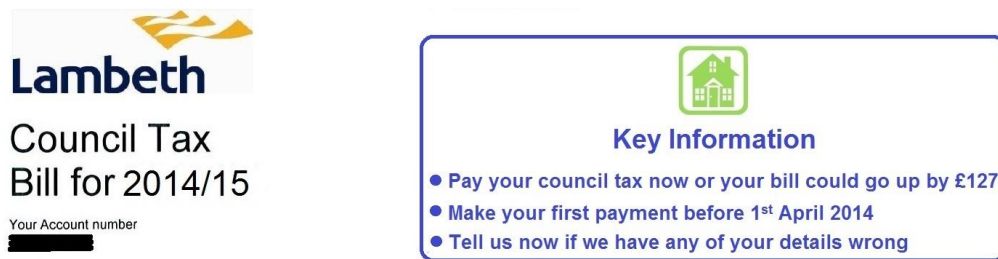
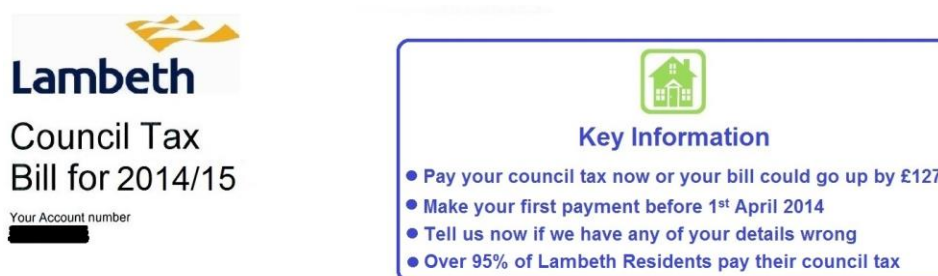


Figure 2
Social norm treatment



Figure 3
Combined (social norm + simplification) treatment



the normal annual bill (the control group); the second group received the simplified bill; the third group were sent the social norm bill; and the fourth group received the combined simplification and social norm bill. As each council taxpayer has an individual account reference, it was straightforward to monitor the integrity of the data in tracking the treatment effect. The allocation to the treatment groups is balanced as shown in Table 1, which reports the results of regression of treatment allocation on council ward covariates and the council tax

band (there are eight bands according to the property value of the household).

At the end of February, a list of accounts for the three selected wards was compiled and those accounts that were not suitable for inclusion in the experiment sample were removed. These included: those with automated payment methods (e.g. direct debit) in place; accounts where an exemption from council tax had been awarded; accounts where full Council Tax Support (CTS) for 2013-14 had been awarded; those accounts with e-billing set up; and void accounts. The resulting list

Table 1
Experiment 1 - balance tests, probit
regressions on treatment allocation

VARIABLES	(1) Simplified	(2) Social norm	(3) Combined
Brixton	-0.0212 (0.0508)	0.0677 (0.0503)	0.0487 (0.0506)
Ferndale	0.0106 (0.0499)	0.0132 (0.0503)	-0.0362 (0.0507)
Tax band	0.0297 (0.0178)	0.0267 (0.0183)	0.0272 (0.0181)
Deprivation	3.68e-08 (5.81e-06)	-5.59e-06 (5.95e-06)	6.96e-07 (5.83e-06)
Constant	-0.0864 (0.0692)	-0.0530 (0.0691)	-0.0949 (0.0693)
Observations	3,952	3,985	3,943

Note: Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

of accounts remaining in the sample were only cash payers who would be liable for a payment for 2014-15 and who received a paper bill. Council tax bills were sent out at the beginning of March, with the first payments for 2014-15 due on or before 1st April. The accounts were then monitored over a period of one month to measure the effect that the different letters had on levels of council tax payment.

The outcome data were captured on 8th April, one week after the deadline for making a first payment for the 2014-15 financial year, immediately prior to reminder notices being sent out, to account holders who had not made their first payment on time. The final sample size for the trial was 7,951 individual account holders. These were randomly assigned to one of the four groups: control group (normal bill), 1,975; simplification, 1,988; social norm, 2,015; simplification + social norm, 1,973.

After the experiment started and before the randomization was undertaken, the circumstances of some accounts had changed, such as being awarded Council Tax Support or some other arrangement that means no payment was due on the account on 1st April 2014. In these circumstances the account status was recorded as 'no payment due'. Other accounts had been closed, where the resident has moved out of the borough or the liability for payment had ended. If the liability end

date was before 1st April 2014 then no payment would be due, if it was after 1st April then the balance would be due as a single sum, but additional time is given to account holders to make this payment. These were recorded as 'vacated'. The combined number of vacated and no payment due records was within a range of 7.39-7.81% of total sample size for each group: control group, 87, 62 (7.54%); simplification, 85, 62, (7.39%); social norm, 93, 62, (7.69%); and simplification plus social norm, 86, 68, (7.81%). In total, 605 records that were listed as vacated or no payment due were removed from the analysis of payment levels in the 'adjusted sample', since no payment was any longer required on these accounts and there is no bias in the treatment allocation from removing them. The remaining 7,346 accounts, where a payment was due by 1st April 2014, were categorized with one of the following statuses: not paid – where no payment for 2014-15 had been received; paid – where the instalment due on 1st April has been paid in full; part-paid – where the instalment due on 1st April has been paid in part; and direct debit – where the account has switched from 'cash payment' to paying by direct debit. Accounts with a 'paid' or with a 'direct debit' status were regarded as being paid. Where accounts were recorded as 'part paid' it was decided to classify them as unpaid since they had not met the objective of being paid in full.

Experiment 1: Results

Table 2 contains the results from the experiment in terms of full payment outcomes and reports the p-values from a chi-squared test (also corrected for multiple comparisons). The table shows that the simplification raises payment by 3.8 percentage points for the simplification only group. The increase is 4.3 percentage points for the combined group, essentially the same treatment impact for the simplification-only group. There is no impact of a descriptive social norm, nor is there an impact of social norm combined with simplification when compared to the simplification-only condition.

There are a small number of residents (3.25%) who make part payments who are classified as non-payers in Table 1. Combining these with those who paid in full reproduces the same pattern of results. Numbers are too small to produce meaningful cross-tabulations of part-payers on their own, but the difference between 3.29% in the control and 4.40% in the combined treatment

Table 2
Experiment 1 - payment in full (%) by treatment allocation

	Simplified	Social Norm	Combined	Control	Total
Not paid in full	52.91	57.63	52.39	56.68	51.28
Paid in full	47.09	42.37	47.61	43.32	48.72
N	1,841	1,860	1,819	1,826	7,346
P-value	0.022	0.559	0.009		
Adjusted p-value	0.033	0.559	.027		

Note: p-values, chi-square on comparison with control; adjusted p-values using Benjamin & Hochberg correction

Table 3
Experiment 1 - probit regressions of paid in full on treatment allocation

VARIABLES	(1)	(2)	(3)	(4)
Simplified	0.0954* (0.0415)	0.0864* (0.0417)	-0.0788 (0.113)	0.0687 (0.0448)
Social norm	-0.0243 (0.0415)	-0.0299 (0.0417)	-0.218 (0.114)	-0.0483 (0.0447)
Combined	0.108** (0.0416)	0.102* (0.0418)	-0.000472 (0.113)	0.0886* (0.0449)
Deprivation		1.30e-05** (4.34e-06)	1.30e-05** (4.35e-06)	1.86e-05*** (4.04e-06)
Brixton		0.0936* (0.0372)	0.0928* (0.0373)	0.106** (0.0371)
Ferndale		0.0934* (0.0373)	0.0933* (0.0373)	0.102** (0.0372)
Tax band		0.0651*** (0.0131)	0.0273 (0.0253)	
Band*simplified			0.0535 (0.0340)	
Band*norm			0.0608 (0.0344)	
Band*combined			0.0336 (0.0341)	
High-low tax				0.0556 (0.0901)
High-low*simplif				0.145 (0.122)
High-low*norm				0.155 (0.123)
High-low*combi				0.112 (0.122)
Constant	-0.168*** (0.0295)	-0.537*** (0.0568)	-0.421*** (0.0872)	-0.397*** (0.0528)
Observations	7,346	7,335	7,335	7,335

Note: Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

has a p-value of .08, which is suggestive of an impact while other cross-tabulations have negligible differences.

Regression analysis reported in Table 3 shows much the same findings, either with just the treatment allocations as covariates or with the addition of the covariates of ward and rank of deprivation. The table also allows the posing the question as to whether the impact of the treatment is conditional on the level of tax band, with higher tax bands thought to be more sensitive to the social norm on the grounds of high social status groups fearing visibility and shaming. But this is not the case—interactions of the treatment with the tax band do not prove to be significant, not even when dividing taxpayers into higher or lower tax bands and interacting on these variables. The results are very clear: there is a clear impact for simplification but no impact for a social norm, which is reinforced in the factorial design where the social norm is varied across two designs, with no impact.

Experiment 2: Study Design

On receiving the findings from the experiment, the local authority decided to simplify all the bill reminders to obtain the benefits from the design. It was decided to test the social norm treatment further. In the 2015/2016 tax year, bill reminders were redesigned accordingly. The trial sample included only accounts that were ‘cash payers’—that is those not currently using automated payments such as direct debit since collection rates among automated payers are very high. Just as in experiment 1, people who pay through e-billing were not considered within the scope of the trial as the emails they receive are not the same as the standard paper bill. Any households where no payment is due were also removed from the trial sample. As with the previous trial, every qualifying household was randomly allocated to a treatment or control group and the relevant council tax bill sent. The results were then monitored to measure any differences in payment levels between the groups. The simplification treatment used in the 2014/2015 trial was adopted as the control version of the bill. Some minor changes were made to remove superfluous text, but the main feature—a box at the top of the bill with ‘key information’—was left unchanged. The social norm treatment was unaltered from the earlier trial,

Table 4
Experiment 2 – probit regression of
treatment allocation on covariates

VARIABLES	RCT GROUP
Ward 00AYGA	0.0499 (0.0291)
Ward 00AYGB	0.0105 (0.0313)
Ward 00AYGC	0.0170 (0.0297)
Ward 00AYGD	0.0203 (0.0275)
Ward 00AYGE	0.0485 (0.0289)
Ward 00AYGF	0.0222(0.0293)
Ward 00AYGG	0.00684 (0.0296)
Ward 00AYGH	0.0693* (0.0295)
Ward 00AYGJ	0.0420 (0.0278)
Ward 00AYGK	0.0218 (0.0284)
Ward 00AYGL	0.0363 (0.0284)
Ward 00AYGM	-0.00559 (0.0291)
Ward 00AYGN	0.0329 (0.0285)
Ward 00AYGP	0.0299 (0.0296)
Ward 00AYGQ	-0.00250 (0.0304)
Ward 00AYGR	0.00712 (0.0295)
Ward 00AYGS	0.0194 (0.0303)
Ward 00AYGT	0.0561 (0.0307)
Ward 00AYGU	0.0436 (0.0285)
Ward 00AYGW	0.0147 (0.0283)
Tax band	0.00220 (0.00313)
Constant	0.984*** (0.0249)
Observations	56,568
R-squared	0.000

Note: Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

except that the 95% figure used to indicate the proportion of Lambeth residents that pay their council tax was changed to 96% to reflect the most recent figures for the borough.

The resulting sample was randomized into 28,876 residents in the social norm group, and 28,877 to the control. The resulting allocation is balanced, as indicated by Table 4, which regresses the treatment allocation on the covariates. One ward is a significant predictor, but with twenty-one wards this is consistent with a balanced sample overall.

On 27 February the data were ‘frozen’ for annual billing purposes and the corresponding council tax bill was then sent out. Between 23 February, when the randomization was done, and 27 February, when the live system was taken offline to

Table 5
Experiment 2 - payment in full by
treatment allocation

	Social norm	Control	Total
Not paid in full	58.60	56.43	57.50
Paid in full	41.40	43.57	42.50
N	25,967	26,775	52,742
p-value	0.000		

Note: p-value, chi-square on comparison with control

process bills, some changes to the accounts occurred, as in the previous year. As a result, the final sample size changed from the original randomization. In total, 56,568 accounts were included within the trial: 27,775 received the social norm bill design and 28,793 received the standard (control) bill. Data were captured on 8th April, one week after the first payment was due, so as to analyze the responses. Accounts were coded according to responses under one of six categories: vacated—where the account holder had vacated and no payment was due on 1st April; switched to direct debit—where the account had changed to paying by direct debit; no instalment—where a change had

been made to an account, such as the award or removal of Council Tax Support, and as a result on 1 April no instalment was due; not paid—where no payment had been made against the April instalment; part paid—where a payment had been made against the April instalment but less than the full amount owed; and paid—where full payment of the April instalment had been made. As before, accounts with a ‘paid’ or with a ‘direct debit’ status were regarded as being paid. Where accounts were recorded as ‘part paid’, in line with the approach taken in the previous council tax trial, it was decided to classify part-paid accounts as ‘unpaid’ since they had not met the objective of being paid in full. Accounts where no instalment was due or where the account holder had vacated were removed from the sample for analysis purposes. This means that the final sample for analysis purposes drops to 52,742. There is no reason to expect these household moves are correlated with the treatment so biasing the experiment.

Experiment 2: Results

Table 5 contains the results for experiment 2. It shows that 41.40% in the treatment group paid in full, whereas 43.57% did so in the control, which indicates the social norm backfired with less people paying in full in the treatment group. This difference is statistically significant at $p < .001$, which is

Table 6
Experiment 2 - probit regressions of paid in full on treatment

VARIABLES	(1)	(2)	(3)	(4)
Social Norm	-0.0555*** (0.0126)	-0.0559*** (0.0126)	-0.0714*** (0.0216)	-0.0649*** (0.0121)
Tax band		0.0478*** (0.0109)	0.0455*** (0.0117)	
Norm*band			0.00478 (0.00726)	
High-low				0.108** (0.0399)
Norm*High-low				0.0561* (0.0238)
Constant	-0.162*** (0.0165)	-0.383*** (0.0363)	-0.375*** (0.0380)	-0.242*** (0.00786)
Observations	52,742	52,742	52,742	52,742

Note: Ward covariates not reported, Robust standard errors in parentheses

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 7
The impact of Wave 1 treatments on
Wave 2 paid in full, probit regressions

VARIABLES	(1)	(2)
Social norm (wave 2)	-0.0560*** (0.0126)	-0.0589*** (0.0130)
Simplification (wave 1)	0.0958*** (0.0115)	0.0585 (0.0362)
Social norm (wave 1)	0.00190 (0.0194)	-0.0309 (0.0181)
Combined (wave 1)	0.141*** (0.0132)	0.154*** (0.0330)
Band	0.0480*** (0.0109)	0.0480*** (0.0110)
Simplified*wave 2 norm		0.0751 (0.0562)
Social norm*wave 2 norm		0.0700 (0.0399)
Combined*wave 2 norm		-0.0260 (0.0639)
Constant	-0.383*** (0.0364)	-0.382*** (0.0361)
Observations	52,742	52,742

Note: Ward covariates not reported; robust standard errors in parentheses

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

different from the non-significant results on the social norm in experiment 1. There is no significant difference between the groups in part-paying the taxes ($p = .266$). Regression analysis is reported in Table 6, using robust standard errors clustered on ward and with wards as covariates (not displayed). Results using robust standard errors are similar to those in models without them. The headline result confirms the finding from the cross-tabulations. The addition of ward and council tax band covariates makes little difference to the results either. The impact of the interaction of treatment and council tax band is not significant, but the interaction of high and low taxpayers and the treatment is positive. It appears that the treatment was less likely to backfire with higher-tax payers, which is consistent with expectations.

The final piece of analysis concerns the impact of the original treatments from experiment 1

in wave 2, the financial year after they were introduced. Table 7 contains these regressions, performed on a merged dataset using the taxpayer identifiers. The simplification treatment, both in simple form and in combination with social norm, continue to influence payment in full, whereas the social norm treatment on its own does not, which is the same pattern of results as before. It is interesting because in wave 2 all bills were simplified so all were treated irrespective as to whether they had been treated in wave 1. Perhaps the change in the design of the bill stimulated a change in behavior, which did not occur when all bills across the borough were the same—a small group effect—or an additive treatment. The second model includes interactions between wave 1 and wave 2 treatments, that is people who got treated twice with a social norm treatment in wave 2. There are no effects here, except for the receipt of the double norm, which is significant at $p < .10$, which indicates that getting the norm twice may have had an effect. If replicated in other studies, this result would be a new finding in the literature on social norms, but it is not possible to come to a firm conclusion here. In this case, the norm message was slightly different in each year with wave 1 being 95% whereas wave 2 was 96%. It is possible that the double-norm treatment caused the residents notice the increase in compliance and they changed their behavior accordingly.

Conclusion

The descriptive social norm message in this study was not successful in getting households to pay their local taxes on time. Not only did it not work, it even backfired when rolled out to all residents in the borough, which is a first in social norm studies. The findings are a puzzle because the literature suggests descriptive social norms should increase payments, especially in the UK context where other tax trials work well. The social norm has also been successful in other local council tax jurisdictions in the UK. The UK Behavioural Insights Team carried out a trial with Medway local authority using a social norm, increasing payments by 11 percentage points (Behavioural Insights Team, 2016; Michael Sanders & Miranda Jackman, 2017). This letter said, “96 per cent of council tax is paid on time. You are currently in the small minority of people who have not paid us yet”. There are thus differences to the Lambeth intervention. Lambeth used the present tense. Medway also had a personalized message

(‘you are’). The council had more space to convey the social norm and to explain it. Another explanation might be due to context in that London is different with a mobile population, which is highly diverse so that social norms do not make so much sense in comparison to a small town in the more rural location of the country of Kent. In London, the make-up of the communities is not only diverse, but often transitory, with about 12% of the population changing each year (Lambeth London Borough Council, 2016). It is important to note that the Medway trial was targeted at late taxpayers whereas the Lambeth one was directed to all cash payers, which are different populations so different compositions of households.

However, this interpretation of the results would be more consistent with the null results in wave 1; but the backfire in wave 2 suggests the norm did work, but just not in the way the local council expected. In the end, the interaction between the context of London and the particular form of social norm intervention caused people to

not to settle their bills on time. Even communicating high rates of payment may have alerted respondents to the possibility of non-payment, perhaps among a wider group of taxpayers who normally settle up on time. These findings can form part of the wider evidence base on social norms, whilst not forgetting the positive and sustained results for one-time simplification.

Acknowledgment

The authors thank the finance and revenues staff in the London borough of Lambeth, who advised and implemented the project. The project received an impact prize from the Local Government Association and the Institute for Revenues, Rates and Valuations in 2015. The findings were presented to the Randomised Controlled Trials in the Social Sciences Twelfth Annual Conference, University of York, September 2017. We thank participants for their comments.

References

- Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9), 1082–1095.
- Behavioural Insights Team. (2016). The Behavioural Insights Team’s Update Report: 2015-16 | The Behavioural Insights Team. Retrieved 6 October 2017, from <http://www.behaviouralinsights.co.uk/publications/the-behavioural-insights-teams-update-report-2015-16/>
- Bhargava, S., & Manoli, D. (2012). Why Are Benefits Left on the Table? Assessing the Role of Information, Complexity, and Stigma on Take-Up With an Irs Field Experiment. *ACR North American Advances*, NA-40. Retrieved from <http://acrwebsite.org/volumes/1012975/volumes/v40/NA-40>
- Bhargava, S., & Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *The American Economic Review*, 105(11), 3489–3529.
- Blumenthal, M., Christian, C., Slemrod, J., & Smith, M. G. (2001). Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota. *National Tax Journal*, 54(1), 125–138. <https://doi.org/10.17310/ntj.2001.1.06>
- Carpio, L. D. (2013). Are the Neighbors Cheating? Evidence from a Social Norm Experiment on. In *Property Taxes in Peru*, Unpublished Manuscript, Princeton University, 2014, http://scholar.princeton.edu/sites/default/files/Are_the_neighbors_cheating_Apr2014_0.pdf
- Castro, L., & Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization*, 116(C), 65–82.
- Cialdini, R. B. (2003). Crafting Normative Messages to Protect the Environment. *Current Directions in Psychological Science*, 12(4), 105–109.
- Coleman, S. (1996). The Minnesota Income Tax Compliance Experiment--State Tax Results (SSRN Scholarly Paper No. ID 4419). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=4419>
- Coleman, S. (2007). The Minnesota Income Tax Compliance Experiment: Replication of the Social Norms Experiment (MPRA Paper No. 5820). University Library of Munich, Germany. Retrieved from <https://ideas.repec.org/p/pra/mprapa/5820.html>
- Coricelli, G., Rusconi, E., & Villeval, M. C. (2014). Tax evasion and emotions: An empirical test of re-integrative shaming theory. *Journal of Economic Psychology*, 40(Supplement C), 49–61.

- Gerber, A. S., & Rogers, T. (2009). Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should You. *The Journal of Politics*, 71(1), 178–191.
- Hallsworth, M. (2014). The use of field experiments to increase tax compliance. *Oxford Review of Economic Policy*, 30(4), 658–679.
- Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148(Supplement C), 14–31.
- Hasseldine, J., Hite, P., James, S., & Toumi, M. (2007). Persuasive Communications: Tax Compliance Enforcement Strategies for Sole Proprietors. *Contemporary Accounting Research*, 24(1), 171–194.
- Iyer, G. S., Reckers, P. M. J., & Sanders, D. L. (2010). Increasing Tax Compliance in Washington State: A Field Experiment. *National Tax Journal*, 63(1), 7–32.
- John, P., & Blume, T. (2017). Nudges That Promote Channel Shift: A Randomized Evaluation of Messages to Encourage Citizens to Renew Benefits Online. *Policy & Internet*, 9(2), 168–183.
- Kettle, S., Hernandez, M., Ruda, S., & Sanderson, M. A. (2016). Behavioral Interventions in Tax Compliance: Evidence from Guatemala (SSRN Scholarly Paper No. ID 2811337). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=2811337>
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, 79(3), 651–692.
- Lambeth London Borough Council. (2016). The State of the Borough 2016. Retrieved from <https://www.lambeth.gov.uk/sites/default/files/State%20of%20Borough%202016%20-%20v3.pdf>
- Mascagni, G. (2017). From the Lab to the Field: A Review of Tax Experiments. *Journal of Economic Surveys*, n/a-n/a. <https://doi.org/10.1111/joes.12201>
- McGraw, K. M., & Scholz, J. T. (1991). Appeals to Civic Virtue versus Attention to Self-Interest: Effects on Tax Compliance. *Law & Society Review*, 25(3), 471–498.
- Michael Sanders and Miranda Jackman. (2017). Behavioural Insights, Tax, and TEST and Build.
- Murphy, K. (2008). Enforcing Tax Compliance: To Punish or Persuade? *Economic Analysis and Policy (EAP)*, 38, 113–135.
- Oliver, A. (2013). Behavioural Public Policy. Cambridge, United Kingdom; New York: Cambridge University Press.
- Schultz, P. W. (1999). Changing Behavior With Normative Feedback Interventions: A Field Experiment on Curbside Recycling. *Basic and Applied Social Psychology*, 21(1), 25–36.
- Sigala, M., Burgoyne, C. B., & Webley, P. (1999). Tax communication and social influence: evidence from a British sample. *Journal of Community & Applied Social Psychology*, 9(3), 237–241.
- Silva, A., & John, P. (2017). Social norms don't always work: An experiment to encourage more efficient fees collection for students. *PLOS ONE*, 12(5), e0177354.
- Slemrod, J., Blumenthal, M., & Christian, C. (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, 79(3), 455–483.
- Spicer, M. W., & Thomas, J. E. (1982). Audit probabilities and the tax evasion decision: An experimental approach. *Journal of Economic Psychology*, 2(3), 241–245.
- Tan, F., & Yim, A. (2012). Can Strategic Uncertainty Help Deter Tax Evasion? An Experiment on Auditing Rules (SSRN Scholarly Paper No. ID 1977420). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=1977420>
- Torgler, B. (2013). A Field Experiment in Moral Suasion and Tax Compliance Focusing on Underdeclaration and Overdeduction (Vol. 69). <https://doi.org/10.1628/001522113X675647>
- Wenzel, M. (2005). Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology*, 26(6), 862–883.
- Wenzel, M. (2006). A Letter from the Tax Office: Compliance Effects of Informational and Interpersonal Justice. *Social Justice Research*, 19(3), 345–364.
- Zarcadoolas, C. (2011). The simplicity complex: exploring simplified health messages in a complex world. *Health Promotion International*, 26(3), 338–350.