European Economic Review 64 (2013) 266–284

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

Promoting rule compliance in daily-life: Evidence from a randomized field experiment in the public libraries of Barcelona

Jose Apesteguia [a,b,c,*], Patricia Funk [b,c], Nagore Iriberri [b,d]

[a] ICREA, Spain
[b] Universitat Pompeu Fabra, Spain
[c] Barcelona GSE, Spain
[d] University of the Basque Country UPV/EHU, IKERBASQUE, Basque Foundation for Science, Spain

ARTICLE INFO

Article history:
Received 2 November 2012
Accepted 10 August 2013
Available online 27 September 2013

JEL classification:
C93
D01
D03
D63
K42

Keywords:
Rule compliance
Field experiment
Public libraries

ABSTRACT

We study how to promote compliance with rules that carry low penalties and are pervasive in all sorts of organizations. We have access to data on the users of all public libraries in Barcelona. In this setting, we test the effect of sending email messages with different contents. We find that users return their items earlier if asked to do so in a simple email, showing that a general reminder of the users' duty is effective in promoting rule compliance. Furthermore, adding other contents to the general reminder does not increase compliance significantly.

1. Introduction

Understanding compliance with rules is crucial for firms, organizations, and societies at large. While some rules are backed by severe sanctions, there are many others with relatively low punishment. To illustrate, workers are supposed to obey rules that make everyday interactions among co-workers more effective, such as attending meetings on time. Parents are supposed to pick up their kids from daycare on time, or to communicate immediately any disease that may have negative externalities on other kids. Visitors to parks and recreational facilities are not supposed to litter because it is costly for the maintenance service. Among the academic community, researchers are supposed to attend seminars and submit referee reports in a timely manner. Examples of similar nature abound.

Authorities often use and set this type of rules, in order to improve their functioning based on cost effectiveness and optimality arguments, and they often do not back them with severe sanctions because it may be too costly to monitor and

* Corresponding author at: Universitat Pompeu Fabra, Spain. Tel.: +34 93 5422521.
E-mail addresses: jose.apesteguia@upf.edu (J. Apesteguia), patricia.funk@upf.edu (P. Funk), nagore.iriberri@gmail.com (N. Iriberri).

0014-2921/$ - see front matter
http://dx.doi.org/10.1016/j.euroecorev.2013.08.010

implement the sanctions or due to institutional and legal restrictions. While these rules are pervasive and important for the functioning of organizations and societies, very little is known on how to promote compliance with them.

In this paper we contribute towards a better understanding of the mechanisms that may promote compliance with these types of rules. While economists would naturally think about monetary incentives, it has been found that they may backfire (see Benabou and Tirole, 2003, 2006, for theoretical arguments; Gneezy and Rustichini, 2000a, 2000b; Mellstroem and Johannesson, 2008, for empirical studies), or that they are not feasible due to political or legal restrictions. Therefore, it is crucial to understand whether there are other possible ways to promote compliance with rules. The goal of this paper is to analyze the effect of conveying various types of messages, in our case by email. Our interest in the potential effects of sending messages is that it offers a virtually costless and non-invasive intervention mechanism that is simple and flexible to implement. Surprisingly, despite the advantages of this message intervention, little is known about its effectiveness.

A setting that allows us to study compliance with rules with low penalties is the Network of Public Libraries in the city of Barcelona. The type of compliant behavior we analyze is whether users of the libraries return the items they borrowed on time. A user not returning an item by the due date is violating the rule, and generating a negative externality on the population of users. The penalty associated with returning an item late does not involve any monetary fines, but the exclusion from the possibility of borrowing more items for a time period equal to the number of days the item is overdue. We evaluate whether we can get users to return the items they borrow earlier, by means of different email contents that are randomly allocated.

There are important characteristics that make our study unique. First, we observe the borrowing behavior of all users of all public libraries in Barcelona over 11 months. During this time span, there were about 50,000 different users, who borrowed over a million items in the 32 different libraries spread throughout the city of Barcelona. Therefore, we have data on a large number of individuals, in a daily-life situation, taking part in their natural environment, and over an extended period of time. Second, we observe *every* borrowing-returning transaction of items made by users. This allows us to measure compliant behavior with the borrowing rules with exact precision. Third, the rules that govern the interaction between the users and the libraries are simple and well-defined. Finally, the rich data on users offers a unique opportunity to test for differential treatment effects with regard to previous compliance and demographic variables such as gender, age, and nationality.

By studying compliance in public libraries by way of the random allocation of messages, we add to a small but growing literature on messages and rule-compliance in other settings. Pomeranz (2010) analyzes firms' tax compliance in Chile. She finds that messages indicating an increased audit probability generate a strong increase in VAT payments. Fellner et al. (2013) study citizens' subscriptions to TV licenses in Austria. They find that a legal threat mailing significantly increases compliance rates, while neither a moral appeal nor a social information mailings have any effect. Cadena and Schoar (2011) analyze loan-repayments in Uganda and show that a text message increases repayment rates. Finally, Pruckner and Sausgruber (2013) establish that a moral message increases payments for newspapers sold following an honor system. In contrast to these studies, our setting includes a formal rule with relatively low penalty. What to expect a priori from the email intervention when stakes are small is open to question, since users may simply ignore the messages they receive. Our paper additionally differs in that we can perfectly measure compliance with rules, and study the effect of messages on different users on the basis of their previous compliance. This is important because there is evidence on crowding out effects of different interventions on individual behavior (see Frey and Jegen, 2001).

In our study, we randomize all users into groups receiving one of five different email messages, and study their behavior after receiving the email. One of the five email messages is a Control message that provides a link to the webpage of the Network of Libraries. All the remaining messages add content to the text in Control. The first treatment message, called Reminder, represents a general reminder of the users' duty to return the borrowed items on time. The second message, Social, adds to Reminder an appeal to the effect individual behavior can have on the overall functioning of the public library services. The last two email treatments, Late and Penalty, are targeted only at those users who have recently returned at least one item late. Both Late and Penalty add to Reminder the identification of the user as having recently returned items late. Finally, Penalty builds on Late and adds a reminder of the penalty associated with non-compliant behavior. Therefore, our email messages are designed to evaluate whether a generic reminder promotes rule-compliance and whether adding contents related to social motivations, the explicit identification for being late, or directing the attention of users towards the penalties associated for rule-breaking, make users conform with the rule to a greater extent.

In our analysis we evaluate the effect of emails on the proportion of late returned items by user, and on the number of days that elapse between the return date and the due date. The first variable measures the propensity to comply with the rule, while the second variable quantifies the positive/negative externality that is imposed on other users when a user returns the item earlier/later than the due date. Our main result is that compliant behavior can be promoted by sending an email that includes a general reminder of the users' duty to comply with the rule. All four treatment emails significantly reduce both, the proportion of late returned items, and the number of days between the return date and the due date. Furthermore, we cannot reject the hypothesis that all four different contents have the same effect, showing that the additional contents to the general reminder do not increase rule compliance significantly. These results show that a low cost intervention such as sending general reminders have a significant effect on users' behavior and the overall functioning of the libraries. Ultimately, this results in an improvement in the service offered by public libraries, as it alleviates the problem of the negative externality imposed by late returns.

As for the effectiveness over time, we show that the effect of getting one of these emails is short-term, mostly significant in the few weeks after the treatment. However, the effect is reproduced when the same email is received for a second time, in our case two and a half months later.

With regard to heterogeneous treatment effects based on previous compliance, we find that the email messages affect all users, although they are specially effective on those users who have shown a worse compliance with the rule in the past.

Last, we investigate differences in reactions according to user demographics. With respect to gender, we find no significant differences in reactions to the treatments between women and men – despite evidence on gender differences in other economically relevant situations (see Croson and Gneezy, 2009). In terms of age, we do not find any consistent differences across age groups in the reaction to the treatments either. As for users' nationalities, we study reactions of users from different geographical regions. Consistent with Fisman and Miguel (2007), we find different reactions depending on the users' nationalities. Interestingly, only Spaniards, people from English speaking countries and Asians react to the emails.

While our paper directly speaks to the before mentioned studies on rule compliance, it also relates to other strands of literature. First, our mechanism, sending email messages, belongs to the class of interventions known as nudges (see Thaler and Sunstein, 2008; Camerer et al., 2003). The paper also relates to a literature investigating communication between interacting agents in specific experimental settings that broadly relate to pro-social behavior, such as hold-up problem games (Ellingsen and Johannesson, 2004), trust games, hidden-information games (Charness and Dufwenberg, 2006, 2011) and dictator games (Ellingsen and Johannesson, 2008; Andreoni and Rao, 2010). The difference is that we are analyzing the effect of *specific message contents* on promoting pro-social behavior and compliance with rules.

The paper also differs from a strand of literature that investigates how providing information affects various individual choices such as retirement decisions (Duflo and Saez, 2003), or school choices (Jensen, 2010). These settings study the effect of providing information on very complex goods and services, whose characteristics are most likely only partially known by the individuals. In contrast, we study a familiar and an everyday scenario where there is a simple and well-defined rule, which is known by the users. Consequently, it is less clear whether information will have any effect in this situation.

Last, there is work studying the effect of messages on *norm* compliance, as opposed to rule compliance, related to home electricity consumption (Schultz et al., 2007; Ayres et al., 2009), saving commitments (Karlan et al., 2011; Kast et al., 2012), voting behavior (Gerber et al., 2008; Allison and Strauss, 2009), fundraising (Huck and Rasul, 2010), contributions to public good experiments (Dal Bó and Dal Bó, 2010), or adherence to health care recommendations like vaccinations (see, e.g., the recent editorial of Szilagyi and Adams, 2012, in the *JAMA*). The main difference between these studies and ours is that in our case there is an exogenously imposed rule that dictates clearly what to do, namely to return the items on time, as opposed to recommendations, or some unwritten, informal, or self-imposed norm. This is important since it has been shown that in many contexts behavior varies depending on whether it is regulated by internal norms or externally imposed rules (see Ostrom, 1990; Crawford and Ostrom, 1995). In particular, there is evidence that compliance with self-imposed norms is much higher than compliance with exogenously imposed rules, even when these are comparable in the final desired behavior and associated penalties (see Dal Bó et al., 2010, for recent experimental evidence). Therefore, the literature on the effect of messages on behavior regulated by norms is not necessarily informative for the library setting investigated here.

The remainder of this paper is organized as follows. Section 2 describes the setting, namely the Network of Public Libraries in the city of Barcelona, and explains the design of the field experiment, as well as the identification strategy. Section 3 is devoted to the presentation and discussion of the results. Finally, Section 4 concludes.

2. The field experiment

2.1. The setting: network of public libraries of barcelona

The Network of Public Libraries in the city of Barcelona is managed by a central body dependent on the City Hall of Barcelona and the Government of the Province of Barcelona. It encompasses 32 libraries spread throughout the city of Barcelona. Each library offers the possibility of borrowing items such as books, DVDs, CDs and magazines; other services such as internet access and exhibitions are also provided.

The rules governing the borrowing of different item types are clearly defined and are the same for all the 32 libraries. At the time of our study, a book could be borrowed for 21 days, while all other item types (DVDs, CDs and magazines) could be borrowed for 7 days. Users could also ask to extend the due date if no other user required that item and before the item was due. As for the maximum number of items to be taken, each user could simultaneously take a total of 30 items, 15 books and magazines, and 15 CDs and DVDs. The penalty associated with returning an item late involved being barred from borrowing new items for a time period equivalent to the number of days elapsed between the due date and the actual return day.[1] In particular, there was no monetary fine associated with not complying with the return policy.

[1] Given the system was managed by a central body a user who was barred from borrowing new items because she was late in a particular library could not borrow items from a different library within the Network of Public Libraries.

2.2. Data and email contents

We observed the complete borrowing/returning behavior for every single user from January 2009 until the beginning of November 2009. For every transaction we observed (i) the user code, which uniquely identifies users, as well as their gender, age, and nationality, (ii) the item code and its characteristics, that is, whether it was a book, DVD, CD, or magazine, (iii) the dates of the transaction, that is, the date when the item was borrowed and returned, and (iv) the library where the transaction took place. With this information we were able to follow the exact borrowing behavior of every single user of the Network of Public Libraries in Barcelona. Given that our design is based on emails, we concentrate on the sample of those users with a known email.[2] This gives us about 50,000 different users, who borrowed over a million items.

The Network of Public Libraries in the city of Barcelona maintains constant communication with its users via email. Most emails include information on the activities organized in the different libraries of the city, such as exhibitions, and on opening hours. In collaboration with the Network, we designed five different email messages (see Table 1) that were randomly assigned to the users.[3]

Control refers to the control treatment. It provides a link to the webpage of the Network of Public Libraries. The rationale for this treatment is that by comparing the effect of the other treatment messages relative to the Control, we are able to differentiate the effect of the content of a treatment message from that of just getting an email from the Network of Libraries. The rest of the treatment messages build on Control, adding different pieces of information. Reminder represents a general reminder of users' duty to return items on time. There is evidence that people fail to pay attention to all the relevant contingencies they are involved in (see Karlan et al., 2011; Szilagyi and Adams, 2012). In this sense, mechanisms calling the attention of individuals to a specific margin may improve their behavior related to it. Reminder, then, aims to test whether by simply recalling the duty of returning the items on time improves behavior towards it. Social builds on Reminder, adding an appeal to the influence of individual behavior on the proper overall functioning of the public system of libraries. There is a large literature in Economics showing that individuals often care about the well-being of others, even when this comes at a personal cost (see, e.g., Andreoni and Miller, 2002; Charness and Rabin, 2002; Sobel, 2005). The purpose of Social is to test whether an appeal to the influence of one's behavior on the others improves rule compliance on top of the simple reminder. Email Late adds to the content of Reminder a statement that identifies the user as having recently returned an item late. It has been shown (see Gerber et al., 2008) that explicitly revealing that one has been observed breaking a norm reduces the norm breaking. This may be due to triggering feelings of guilt or shame (see, e.g., Battigalli and Dufwenberg, 2007). Then, the purpose of Late is to test whether the explicit mentioning of having been observed as a non-compliant improves behavior. Finally, Penalty builds on Late adding a reminder of the actual penalty associated with returning an item late. Here, we aim to direct users' attention to their own cost when breaking the rules. Our inspiration is the large literature in Law and Economics on the deterrence effect of punishment (see Becker, 1968). The purpose here is to evaluate whether focusing the attention of users towards the penalties improves rule-compliance on top of Late.

It has been our aim to design emails with general contents that could be applied to many settings of interest beside libraries. For this reason, no email makes any reference to particular items that may have been borrowed at the time of receiving the email so users with and without a borrowed item are equally likely to receive the email treatments. Moreover, not all settings permit the type of precise data on individual behavior that we had at the moment of treatment (e.g., identifying users as late and non-late). In this vein, three of our emails, Control, Reminder and Social, are general in the sense of not using any information on the behavior of users prior to the treatment, and can therefore easily be adapted to other settings. In many settings, individual behavior is not observable, so one needs to use interventions that do not require this type of information. On top of that, for cases where information on individual behavior is available to the policy maker, it is important to analyze potential effects of using such specific information. In our case, Late and Penalty use information on user history in order to directly target non-compliant individuals.

2.3. Randomization

We considered users independent of whether they had any borrowed items at the moment of the email intervention. We sent emails in two different waves. Wave 1 was sent on July 1, 2009, when we reached about 36,700 users. Wave 2 was sent on September 15, 2009, when we reached about 38,300 users. Overall, we reached about 50,000 *different* users.[4]

In Wave 1 (resp. Wave 2) we considered all the active users between January 1 and May 5, 2009 (resp. between March 1 and July 31, 2009), and classified them into two categories: late users and non-late users. An active user is a user who borrowed at least one item during the time interval mentioned. A late user is a user who returned an item after the due date

[2] The Network of Public Libraries knows the email addresses of around 40% of the registered users. Those with an email address are younger, more likely to be female and foreign. The biggest difference in magnitude is found for age, which is to be expected.

[3] The subject lines of all the five email messages contained the same text: “Important Notice.” By writing so in the subject line we aimed to distinguish our emails from the ones users receive regularly. The language used in the emails was Catalan, as this is the official language used by the Network when communicating with the library users.

[4] In each wave we sent about 50,000 emails but not all emails were actually delivered. About 30% of email addresses turned out to be invalid and the email messages were returned to the server as messages that were never delivered. We therefore restrict our analysis to those users to whom the message was delivered.

Table 1
Email Messages.

E-mail	Text
Control	Dear User, In the next webpage you will find information on the services and activities offered by the Libraries of Barcelona: http://www.bcn.es/biblioteques/ Best wishes, Libraries of Barcelona
General ***Reminder***	Dear User, **If at some point you borrow an item from the library, please remember that you have to return it on time**. Best wishes, Libraries of Barcelona In the next webpage you will find information on the services and activities offered by the Libraries of Barcelona: http://www.bcn.es/biblioteques/
Social *Motivation*	Dear User, **For a good functioning of the Public Libraries it is important to return the items that are borrowed on time**. If at some point you borrow an item from the library, please remember that you have to return it on time. Best wishes, Libraries of Barcelona In the next webpage you will find information on the services and activities offered by the Libraries of Barcelona: http://www.bcn.es/biblioteques/
Identification ***Late***	Dear User, **In the last months you have returned an item late**. If at some point you borrow an item from the library, please remember that you have to return it on time. Best wishes, Libraries of Barcelona In the next webpage you will find information on the services and activities offered by the Libraries of Barcelona: http://www.bcn.es/biblioteques/
Identification Late and Reminder of the ***Penalty***	Dear User, In the last months you have returned an item late. If at some point you borrow an item from the library, please remember that you have to return it on time. **Remember that the time that a user will be excluded from the possibility of borrowing an item will be the same number of natural days elapsed since the day that the item should have been returned. The maximum period for exclusion is one year**. Best wishes, Libraries of Barcelona In the next webpage you will find information on the services and activities offered by the Libraries of Barcelona: http://www.bcn.es/biblioteques/

Notes: The text in bold refers to the new addition of the treatment email. The words in bold in the first column represent the labels we will use in the paper.

at least once during the time interval. A non-late user is a user who did not return any item late during the time interval. Late users were randomly assigned to the five different treatments, while non-late users were randomly assigned to Control, Reminder and Social only.[5] The randomization was carried out at the user level and in order to ensure balance across different libraries, we stratified the randomization using the library at which users signed up.

Note that in Wave 2 we have users who were already active in Wave 1 and new active users, namely those users who were active only between May and July. With regard to the new active users, we repeated the randomization procedure as in Wave 1. The active users in Wave 2 who were also active in Wave 1 received exactly the same email as in Wave 1.[6] Exceptions were those users who were allocated to Late or Penalty in Wave 1 but who, during the interval between March 1, 2009 and July 31, 2009, were never late again. There were about 700 of such users, who were excluded from the randomization, and hence received no email in Wave 2.

Table 2 reports the descriptive statistics of all users, both non-late and late, who were randomly assigned to treatments Control, Reminder and Social in Waves 1 (top) and 2 (bottom). Table 3 reports the descriptive statistics of late users only, randomly assigned to the five treatments in Waves 1 (top) and 2 (bottom). Note that late users appear in both tables.

[5] There were 21,571 late users and 15,106 non-late users in Wave 1.

[6] There were 10,492 new active users in Wave 2, of which 5,191 were late users and 5,301 were non-late users. There were about 28,500 users who were also active in Wave 1.

Table 2
Randomization of all users into treatments Control–Reminder–Social.

	Control			Reminder			Social			p–Value
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Equ. Means
Wave 1 (***active between 1st January–5th May***)										
Male	9438	**0.42**	0.49	9059	0.42	0.49	9423	0.42	0.49	0.67
Age	9448	**32.71**	13.83	9062	**32.76**	13.89	9434	**33.07**	13.78	0.16
Foreign	9467	**0.28**	0.45	9080	**0.30**	0.46	9452	**0.30**	0.46	0.07
Proportion Late	9467	**0.33**	0.39	9080	**0.33**	0.39	9452	**0.33**	0.39	0.94
"Actual-Due" Date	9376	**1.74**	16.75	8995	**1.53**	16.37	9349	**1.31**	15.86	0.19
Number of Loans	9467	**11.92**	18.80	9080	**11.89**	19.01	9452	**12.48**	22.21	0.08
Book	9467	**0.60**	0.42	9080	**0.60**	0.42	9452	**0.61**	0.42	0.30
CD	9467	**0.09**	0.23	9080	**0.10**	0.23	9452	**0.09**	0.23	0.49
DVD	9467	**0.28**	0.37	9080	**0.28**	0.37	9452	**0.27**	0.36	0.32
Magazine	9467	**0.03**	0.13	9080	**0.02**	0.12	9452	**0.02**	0.11	0.12
Wave 2 (***Active between 1st March–31st July***)										
Male	10,037	**0.42**	0.49	9758	**0.41**	0.49	10,151	**0.41**	0.49	0.81
Age	10,049	**32.74**	14.09	9763	**32.49**	13.98	10,157	**32.79**	13.73	0.28
Foreign	10,064	**0.28**	0.45	9782	**0.29**	0.45	10,180	**0.30**	0.46	0.01
Proportion Late	10,063	**0.35**	0.39	9782	**0.36**	0.39	10,180	**0.35**	0.39	0.92
"Actual-Due" Date	9923	**1.58**	14.69	9639	**1.54**	14.82	10,047	**1.48**	14.33	0.89
Number of Loans	10,064	**11.18**	18.56	9782	**11.01**	18.67	10,180	**11.67**	21.83	0.05
Book	10,064	**0.62**	0.41	9782	**0.62**	0.41	10,180	**0.61**	0.41	0.49
CD	10,064	**0.09**	0.22	9782	**0.09**	0.22	10,180	**0.09**	0.22	0.47
DVD	10,064	**0.27**	0.36	9782	**0.27**	0.36	10,180	**0.27**	0.36	0.31
Magazine	10,064	**0.03**	0.13	9782	**0.03**	0.13	10,180	**0.03**	0.12	0.23

Notes: All variables refer to all users, late and non-late, who were active in windows 1 (1st January–15th May) and 2 (1st March–31st July) and to whom the email was delivered. All variables are obtained at the user level for the pre-treatment period. *Male* takes a value of 1 in case of male, *Age* shows the user's age in years, and *Foreign* is a dummy variable taking a value of 1 in the case of Non-Spanish. *Proportion Late* measures the proportion of late returns per user, and "*Actual-Due*" *Date* measures the average number of days between the return date and the deadline per user. *Number of Loans* represents the number of loans per user. *Book*, *CD*, *DVD* and *Magazine* reflects the user's average share of Books, CD's, DVD's and Magazines. The *p*-value in the last column is for the *F*-Test of equality of variable means across all three groups.

The choice of showing the randomization for all users (Table 2) and late users only (Table 3) corresponds to our design (see the discussion at the end of Section 2.2) as well as our subsequent analysis. The last column in Tables 2 and 3 report the *p*-values for the *F*-Test of equality of variable means across all groups.

Consistent with the random assignment of users to treatments, the average user has similar values in the observable characteristics across the different treatments.[7]

In Tables 2 and 3, we can see the magnitude of the problem of late returns. Considering those users who have been late at least once, around 60% of the loans per user are returned after the due date. Furthermore, the typical late user returns the borrowed items on average 6.5 days later than the due date.

In our analysis, we concentrate on the post-treatment period, that is, on the behavior of users after the email intervention. For those users who received the email message in Wave 1, the post-treatment starts on the 1st of July. For those users who got the email for the first time in Wave 2, the post-treatment starts on the 15th of September.

Not all users who received the email treatment appear in the post-treatment period; some users neither borrow nor return any item at any time in the post-treatment period. It is important to investigate whether the email treatments affected borrowing behavior, and whether conditional on borrowing, the randomization is still valid. As a first step, we investigate, separately for Waves 1 and 2, the users' probabilities of being active in the post-treatment period, as well as their borrowing intensity measured by the number of loans taken. In particular, we test whether the variables "*Active*" (a dummy taking a value of 1 if the user borrows/returns any item in the post-treatment period, and 0 otherwise) and "*Number of Loans*" (which measures the number of loans in the post-treatment period and takes the value of 0 for those users who are not active) show significant differences across treatments.

As can be seen from Table 4, around 50% of emailed users were active in the post-treatment period (see the constant in columns (1) and (5)), and the number of borrowed items was between 4 and 6 (see the constant in columns (3) and (7)). More importantly, we cannot reject the null that the probability of being active and the number of loans taken is equal between the control and the treatment groups (see *p*-values at the bottom of each column in Table 4). Therefore, the email interventions did neither affect differentially the likelihood of being active, nor the borrowing intensity compared

[7] An exception is the proportion of foreigners, possibly due to the valid email address correction (see footnote 4). However, the mean values do not show sizable differences.

Table 3
Randomization of late users into treatments Control–Reminder–Social–Late–Penalty.

	Control			Reminder			Social			Late			Penalty			*p*−Value
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Equ. Means
Wave 1 (***active between 1st January–5th May***)																
Male	4315	**0.43**	0.49	4182	**0.43**	0.50	4351	**0.43**	0.50	4333	**0.43**	0.50	4304	**0.42**	0.49	0.97
Age	4321	**32.20**	12.78	4187	**32.30**	12.83	4355	**32.48**	12.55	4343	**32.41**	12.76	4312	**32.23**	12.58	0.82
Foreign	4331	**0.33**	0.47	4195	**0.35**	0.48	4367	**0.35**	0.48	4355	**0.34**	0.48	4323	**0.35**	0.48	0.08
Proportion Late	4331	**0.59**	0.33	4195	**0.58**	0.33	4367	**0.59**	0.33	4355	**0.58**	0.33	4323	**0.59**	0.33	0.22
“Actual-Due” Date	4270	**6.61**	20.79	4143	**5.94**	20.42	4301	**5.68**	19.37	4288	**5.91**	19.08	4269	**5.75**	18.22	0.20
Number of Loans	4331	**17.48**	24.19	4195	**17.91**	24.63	4367	**18.50**	29.57	4355	**17.88**	25.38	4323	**18.35**	27.10	0.39
Book	4331	**0.50**	0.39	4195	**0.49**	0.39	4367	**0.50**	0.39	4355	**0.49**	0.39	4323	**0.48**	0.39	0.20
CD	4331	**0.12**	0.24	4195	**0.12**	0.24	4367	**0.12**	0.24	4355	**0.13**	0.24	4323	**0.12**	0.24	0.77
DVD	4331	**0.35**	0.36	4195	**0.36**	0.37	4367	**0.35**	0.36	4355	**0.36**	0.36	4323	**0.36**	0.36	0.48
Magazine	4331	**0.03**	0.12	4195	**0.03**	0.13	4367	**0.03**	0.12	4355	**0.03**	0.12	4323	**0.03**	0.12	0.69
Wave 2 (***active between 1st March–31st July***)																
Male	4069	**0.43**	0.49	4014	**0.43**	0.49	4178	**0.42**	0.49	4158	**0.42**	0.49	4060	**0.42**	0.49	0.94
Age	4078	**32.18**	12.85	4019	**31.82**	12.74	4180	**32.12**	12.43	4166	**32.31**	12.81	4066	**31.78**	12.46	0.25
Foreign	4086	**0.33**	0.47	4029	**0.34**	0.47	4186	**0.36**	0.48	4178	**0.35**	0.48	4077	**0.37**	0.48	0.01
Proportion Late	4086	**0.62**	0.32	4029	**0.62**	0.32	4186	**0.61**	0.33	4178	**0.61**	0.32	4077	**0.61**	0.33	0.17
“Actual-Due” Date	3989	**6.55**	17.96	3940	**6.50**	18.20	4108	**6.02**	16.93	4067	**6.27**	17.70	3996	**6.23**	16.97	0.65
Number of Loans	4086	**17.51**	24.55	4029	**17.67**	25.10	4186	**18.54**	30.35	4178	**17.77**	26.05	4077	**18.38**	27.46	0.32
Books	4086	**0.50**	0.39	4029	**0.49**	0.39	4186	**0.49**	0.39	4178	**0.49**	0.39	4077	**0.48**	0.39	0.63
CDs	4086	**0.12**	0.24	4029	**0.11**	0.23	4186	**0.12**	0.24	4178	**0.12**	0.24	4077	**0.12**	0.24	0.36
DVDs	4086	**0.35**	0.36	4029	**0.36**	0.37	4186	**0.36**	0.36	4178	**0.36**	0.36	4077	**0.36**	0.36	0.74
Magazines	4086	**0.03**	0.13	4029	**0.04**	0.14	4186	**0.03**	0.12	4178	**0.03**	0.12	4077	**0.03**	0.13	0.06

Notes: All variables refer to the late users who were active in windows 1 (1st January–15th May) and 2 (1st March–31st July) and to whom the email was delivered. All variables are obtained at the user level for the pre-treatment period. *Male* takes a value of 1 in case of male, *Age* shows the user's age in years, and *Foreign* is a dummy variable taking a value of 1 in the case of Non-Spanish. *Proportion Late* measures the proportion of late returns per user, and “*Actual-Due*” *Date* measures the average number of days between the return date and the deadline per user. *Number of Loans* represents the number of loans per user. *Book*, *CD*, *DVD* and *Magazine* reflects the user's average share of Books, CD's, DVD's and Magazines. The *p*-value in the last column is for the *F*-Test of equality of variable means across all five groups.

to the control email.[8] We also study whether the email intervention had any effect on the *types of users* who are active after being treated. To do so, we redid Tables 2 and 3 only for those users who were treated and active in the post-treatment period. We find no differences in observable characteristics across the treatments. To summarize, this analysis prevents possible concerns about differential selection into borrowing behavior based on being treated.

2.4. Identification strategy

We focus on two different dependent variables. First, we look at the proportion of late returned items per user (*Proportion Late*). This is a direct measure of how users comply with the rule. Second, we use the average number of days between the return date and the due date per user (“*Actual-Due*” *Date*). When this difference is positive (resp. negative) the item was returned late (resp. early) compared to the due date. In contrast to the first dependent variable, which measures late/non-late per item in a binary way, this second variable also takes into account the extent of late or early returns.

In a randomized experiment like ours, the causal effect of the treatments can be estimated as follows:

$$Y_i = \alpha + \beta_1 Reminder_i + \beta_2 Social_i + \beta_3 Late_i + \beta_4 Penalty_i + \varepsilon_i \tag{1}$$

where the dependent variable Y_i is either (i) the proportion of late returns per user or (ii) the average number of days between the return date and the due date per user.[9] $Reminder_i$, $Social_i$, $Late_i$, and $Penalty_i$ are dummy variables taking a value of 1 when user i was assigned to Reminder, or Social, or Late, or Penalty, respectively. The omitted treatment to which these variables are compared is Control.

[8] The fact that around half of the treated users are active in the post-treatment period seems to reflect natural trends over time. When computing comparable *Active* and *Number of Loans* for those users with unknown email, we get similar values.

[9] Note that the dependent variables are obtained by collapsing all the transactions at the user level. For example, a user with 5 transactions that was late with 4 of them has a proportion of late returns of 4/5. In the subsequent analysis, when we add control variables, we also collapse them at the user level.

Table 4
Probability of being active and loans taken in the post-treatment period.

	Control–Social–Reminder (all users)				Control–Social–Reminder–Late–Penalty (late users)			
	Active (1)	Active (2)	Number of Loans (3)	Number of Loans (4)	Active (5)	Active (6)	Number of Loans (7)	Number of Loans (8)
Wave 1 (*active between*1*st January*–5*th May*)								
Reminder	0.0115	0.0132*	0.128	0.141	0.0101	0.0117	0.219	0.230
	(0.00732)	(0.00731)	(0.155)	(0.155)	(0.0108)	(0.0108)	(0.284)	(0.284)
Social	0.00189	0.00272	0.150	0.154	−0.00771	−0.00693	0.237	0.248
	(0.00724)	(0.00723)	(0.156)	(0.156)	(0.0107)	(0.0107)	(0.290)	(0.289)
Late					0.0157	0.0170	−0.0409	−0.0304
					(0.0107)	(0.0107)	(0.280)	(0.280)
Penalty					0.0111	0.0116	0.258	0.268
					(0.0107)	(0.0107)	(0.296)	(0.295)
Constant	0.456***	0.313***	4.272***	3.571***	0.518***	0.354***	6.055***	4.648***
	(0.00512)	(0.0171)	(0.108)	(0.412)	(0.00759)	(0.0203)	(0.198)	(0.522)
Library FE	No	Yes	No	Yes	No	Yes	No	Yes
Number of users	27,999	27,999	27,999	27,999	21,571	21,571	21,571	21,571
R-squared	0.000	0.006	0.000	0.005	0.000	0.007	0.000	0.006
H0: all treatment variables = 0 (*p* -values)	0.2446	0.1615	0.585	0.5479	0.1776	0.15	0.7485	0.7336
*Wave*2 (*active between*1*st March*–31*st July*)								
Reminder	0.00330	0.00517	0.0905	0.105	−0.00275	−0.00188	0.127	0.132
	(0.00710)	(0.00708)	(0.147)	(0.147)	(0.0111)	(0.0110)	(0.296)	(0.295)
Social	−0.00908	−0.00849	0.102	0.102	−0.0208*	−0.0213*	0.157	0.147
	(0.00702)	(0.00701)	(0.148)	(0.148)	(0.0110)	(0.0109)	(0.301)	(0.301)
Late					0.00314	0.00215	−0.167	−0.189
					(0.0110)	(0.0109)	(0.290)	(0.289)
Penalty					0.00120	0.00126	0.196	0.190
					(0.0110)	(0.0110)	(0.307)	(0.306)
Constant	0.487***	0.256***	4.222***	2.346***	0.542***	0.331***	6.238***	3.215***
	(0.00498)	(0.0201)	(0.103)	(0.325)	(0.00779)	(0.0258)	(0.207)	(0.434)
Library FE	No	Yes	No	Yes	No	Yes	No	Yes
Number of users	30,032	30,032	30,032	30,032	20,556	20,556	20,556	20,556
R-squared	0.000	0.007	0.000	0.005	0.000	0.008	0.000	0.008
H0: all treatment variable = 0 (*p*- values)	0.2001	0.1533	0.7488	0.7171	0.1736	0.1631	0.7475	0.7136

Notes: *Active* is a dummy variable that takes a value of 1 if the user borrowed or returned an item in the post-treatment period and 0 otherwise. *Number of Loans* measures the number of loans that a user borrows in the post-treatment period (users who are not active in the post-treatment period have assigned a value of 0). The top panel refers to Wave 1 and the bottom panel refers to Wave 2. Columns (1)–(4) refer to all users, while columns (5)–(8) refer to the previous late users only. Robust standard errors in parenthesis: *$p<0.1$, **$p<0.05$, ***$p<0.001$.

Consistent with our design, we will estimate Eq. (1) in two different ways. First, we compare Reminder and Social to Control for all users, independent of whether they were late or not in the pre-treatment period. Second, we compare Reminder, Social, Late and Penalty to Control, restricted to all users who were late at least once in the pre-treatment period.

3. Results

3.1. Average treatment effects

We estimate Eq. (1). Table 5 reports the results for Control, Reminder, and Social, covering all users, both late and non-late users, who got one of these emails in Waves 1 and 2. Table 6 reports the results for all five treatments restricted to the late users only. The first four columns in both tables show the OLS estimates focusing on the post-treatment period, while the last two columns show difference-in-differences estimates using both pre and post-treatment periods.[10] Columns (1), (2) and (5) refer to the proportion of late returns per user, while columns (3), (4) and (6) refer to the average number of days between the return date and due date per user.

We start commenting the results for the post-treatment period (columns (1)–(4) in Tables 5 and 6). In both tables, the first column reports the results of estimating Eq. (1) without any controls. The second column adds all available controls.

[10] The results in Tables 5 and 6 are highly robust with regard to other specifications. For OLS, if instead of collapsing the data at the user level, we estimate random effects with transaction level data, the results are both qualitatively and quantitatively similar. Also, for the dependent variable *Proportion Late*, if we use probit instead of a linear probability model, the results are very much the same.

Table 5
Control–Reminder–Social (all users).

	Proportion Late		**Actual-Due Date**			**Proportion Late**	**Actual-Due Date**
	(1)	(2)	(3)	(4)		(5)	(6)
Reminder	−0.0129*	−0.0138*	−0.387*	−0.468**	Post T.*Reminder	−0.0109	−0.463**
	(0.00781)	(0.00717)	(0.202)	(0.188)		(0.00854)	(0.234)
Social	−0.0167**	−0.0184***	−0.314	−0.409**	Post T.*Social	−0.0173**	−0.327
	(0.00773)	(0.00708)	(0.207)	(0.192)		(0.00837)	(0.227)
					Post-treatment	0.141***	3.736***
						(0.0157)	(0.396)
CD		0.0850***		3.434***	CD	0.0712***	3.455***
		(0.0151)		(0.292)		(0.0184)	(0.480)
DVD		0.0924***		3.455***	DVD	0.0826***	2.819***
		(0.00937)		(0.234)		(0.0122)	(0.342)
Magazine		0.0776***		3.734***	Magazine	0.0350	3.223***
		(0.0238)		(0.526)		(0.0295)	(0.765)
August		0.0681***		0.525	February	−0.0194	0.846
		(0.0155)		(0.499)		(0.0277)	(0.768)
September		0.00730		−2.514***	March	−0.0122	−0.0647
		(0.0121)		(0.338)		(0.0257)	(0.663)
October		0.0253**		−3.414***	April	−0.0218	0.483
		(0.0105)		(0.311)		(0.0254)	(0.641)
November		−0.375***		−11.91***	May	0.0174	0.0170
		(0.0124)		(0.553)		(0.0245)	(0.636)
Age 20−40		−0.00423		0.00343	June	−0.0245	−0.453
		(0.00952)		(0.286)		(0.0252)	(0.667)
Age 40−60		−0.0610***		−1.172***	July	−0.101***	−1.007*
		(0.0102)		(0.289)		(0.0248)	(0.610)
Age over 60		−0.104***		−1.620***	August	−0.0541**	−1.571**
		(0.0143)		(0.382)		(0.0251)	(0.615)
Male		−0.00146		−0.225	September	−0.126***	−4.754***
		(0.00597)		(0.157)		(0.0281)	(0.674)
Foreign		0.0328***		0.683***	October	−0.111***	−5.611***
		(0.00692)		(0.186)		(0.0277)	(0.693)
Number of Loans		−0.00321***		−0.0698***	November	−0.458***	−12.91***
		(0.000163)		(0.00350)		(0.0326)	(0.929)
Prior Late		0.317***			Number of Loans	−0.00203***	−0.0570***
		(0.00968)				(0.000168)	(0.00514)
Prior “Actual-Due”				0.201***			
				(0.0186)			
Constant	0.358***	0.379**	−0.583***	4.803	Constant	0.354***	−0.340
	(0.00551)	(0.159)	(0.149)	(4.347)		(0.0433)	(1.115)
Library FE	No	Yes	No	Yes	Library FE	Yes	Yes
R-squared	0.00	0.17	0.00	0.14	*R*-squared	0.042	0.057
Number of users	14,605	14,442	14,157	13,990	Number of users	14,552	14,110
H0: Reminder = Social (*p*-value)	0.62	0.52	0.72	0.76		0.45	0.55
H0: Reminder = Social = 0 (*p*-value)	0.08	0.03	0.13	0.03		0.11	0.12

Notes: *Proportion Late* measures the proportion of late returns per user, columns (1), (2) and (5), and “*Actual-Due*” *Date* measures the average number of days between the return date and the due date per user, columns (3), (4) and (6). Columns (1)–(4) are estimates when the sample is restricted to the post-treatment period. Columns (5) and (6) are estimates for a difference-in-differences approach with individual fixed effects, using pre and post-treatment period observations. See different email messages in Table 1. *CD*, *DVD* and *Magazine* are dummy variables for the item type (omitted category: Book), *February*–*November* are month dummies (omitted category: July for the post-treatment period and January when pre- and post-treatment periods are considered), and *Age* 20–40, *Age* 40–60 and *Age* over 60 are age dummies (omitted category: Age under 20). *Male* takes a value of 1 in case of male, *Foreign* a value of 1 in case of non-Spanish, and *Number of Loans* is the average number of loans per user. *Prior Late* and *Prior* “*Actual-Due*” refer to proportion of late returns per user and the average number of days between the return date and the due date, both prior to the treatment. Robust standard errors in parenthesis: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

The controls for users' demographics include gender, whether the user is foreign or not, and different age intervals. It also includes controls for different months between July and November, as well as the number of borrowed items. We also add controls for users' behavior prior to the treatment, which measures their propensity for late returns and for the average number of days between the return date and the due date per user, prior to the treatment. Finally, it includes controls regarding the item type (whether it refers to a book, DVD, CD or magazine), as well as library fixed effects.[11]

In Table 5, both Reminder and Social are significant and negative, showing that both email treatments significantly reduce the proportion of late returns and the number of days between the return date and the due date. Taking the estimates of the second column, receiving a Reminder email decreases the proportion of late returns by 1.4% points (compared to Control), and

[11] In all specifications, we discard transactions that were due on a holiday, when the library was closed.

Table 6
Control–reminder–social–late – penalty (late users).

	Proportion Late		Actual – Due Date			Proportion Late	Actual – Due Date
	(1)	(2)	(3)	(4)		(5)	(6)
Reminder	−0.0266**	−0.0239**	−0.535*	−0.614**	Post T.*Reminder	−0.0193	−0.695*
	(0.0109)	(0.0101)	(0.300)	(0.281)		(0.0117)	(0.363)
Social	−0.0275**	−0.0297***	−0.402	−0.549*	Post T.*Social	−0.0317***	−0.575
	(0.0108)	(0.00990)	(0.313)	(0.292)		(0.0114)	(0.350)
Late	−0.0252**	−0.0271***	−0.423	−0.549**	Post T.*Late	−0.0247**	−0.567
	(0.0109)	(0.00992)	(0.304)	(0.280)		(0.0114)	(0.349)
Penalty	−0.0391***	−0.0433***	−0.674**	−0.879***	Post T.*Penalty	−0.0420***	−0.957***
	(0.0108)	(0.00996)	(0.297)	(0.278)		(0.0115)	(0.341)
					Post-treatment	−0.0611***	1.118*
						(0.0209)	(0.617)
CD		0.0863***		3.100***	CD	0.0666***	1.960***
		(0.0159)		(0.323)		(0.0195)	(0.660)
DVD		0.109***		3.173***	DVD	0.0795***	1.938***
		(0.0102)		(0.271)		(0.0130)	(0.414)
Magazine		0.109***		3.175***	Magazine	0.0689**	2.027**
		(0.0252)		(0.549)		(0.0342)	(0.920)
August		0.0508***		0.542	February	0.0249	1.541
		(0.0175)		(0.589)		(0.0292)	(1.011)
September		0.0212		−2.629***	March	0.00739	0.543
		(0.0139)		(0.400)		(0.0268)	(0.862)
October		0.0316***		−4.124***	April	0.00326	0.829
		(0.0120)		(0.366)		(0.0278)	(0.887)
November		−0.502***		−14.69***	May	−0.0934***	−2.224***
		(0.0148)		(0.774)		(0.0295)	(0.820)
Age 20–40		−0.00112		0.515	June	−0.0380	0.219
		(0.0109)		(0.341)		(0.0282)	(0.889)
Age 40–60		−0.0549***		−0.786**	July	−0.00722	1.203
		(0.0120)		(0.352)		(0.0278)	(0.822)
Age over 60		−0.0843***		−0.727	August	0.00106	−0.0230
		(0.0186)		(0.499)		(0.0291)	(0.855)
Male		−0.00151		−0.139	September	−0.0246	−3.274***
		(0.00646)		(0.176)		(0.0316)	(0.916)
Foreign		0.0377***		0.719***	October	−0.0254	−4.454***
		(0.00703)		(0.196)		(0.0311)	(0.939)
Number of Loans		−0.00298***		−0.0727***	November	−0.536***	−15.24***
		(0.000173)		(0.00350)		(0.0371)	(1.472)
Prior Late		0.293***			Number of Loans	−0.00407***	−0.0751***
		(0.0113)				(0.000217)	(0.00532)
Prior "Actual-Due"				0.156***			
				(0.0158)			
Constant	0.440***	0.158	0.759***	4.478		0.537***	3.435**
	(0.00770)	(0.119)	(0.225)	(3.958)		(0.0427)	(1.467)
Library FE	No	Yes	No	Yes	Library FE	Yes	Yes
R-squared	0.00	0.16	0.00	0.14	*R*-squared	0.111	0.089
Number of users	12,286	12,205	11,846	11,750	Number of users	12,229	11,794
H0: Reminder=Social (*p*-value)	0.93	0.56	0.66	0.81		0.29	0.74
H0: Reminder=Late (*p*-value)	0.90	0.75	0.70	0.81		0.64	0.73
H0: Reminder=Penalty (*p*-value)	0.25	0.05	0.64	0.34		0.05	0.46
H0: Late=Penalty (*p*-value)	0.19	0.10	0.39	0.23		0.13	0.25
H0: Reminder=Social=Late=Penalty (*p*-value)	0.55	0.22	0.78	0.58		0.24	0.63
H0: Reminder=Social=Late=Penalty=0 (*p*-value)	0.01	0.00	0.21	0.03		0.01	0.08

Notes: Proportion Late measures the proportion of late returns per user, columns (1), (2) and (5), and "Actual-Due" Date measures the average number of days between the return date and the due date per user, columns (3), (4) and (6). Columns (1)–(4) are estimates when the sample is restricted to the post-treatment period. Columns (5) and (6) are estimates for a difference-in-differences approach with individual fixed effects, using pre- and post-treatment period observations. See the different email messages in Table 1. CD, DVD and magazine are dummy variables for the item type (omitted category: Book), *February–November* are month dummies (omitted category: July for the post-treatment period and January when pre- and post-treatment periods are considered), and *Age* 20–40, *Age* 40–60 and *Age over*60 are age dummies (omitted category: Age under 20). *Male* takes a value of 1 in case of male, *Foreign* a value of 1 in case of non-Spanish, and *Number of Loans* is the average number of loans per user. *Prior Late* and *Prior "Actual-Due"* refer to proportion of late returns per user and the average number of days between the return date and the due date, both prior to the treatment. Robust standard errors in parenthesis: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

1.8% points in the case of Social. Evaluated at the mean propensity of being late for the control group (approximately 36%), the reduction in late returns lies between 4% (Reminder) and 5% (Social). Moreover, receiving a treatment email also significantly decreases the number of days between the return date and the due date: the Reminder and Social emails decrease this difference on average by almost half a day with respect to Control. Note also that the coefficients of the Reminder and the Social emails are not statistically different from each other, meaning that the appeal to social preferences, through one's contribution to the functioning of public libraries, did not affect users' behavior differently from the general reminder.

Focusing on previously late users (Table 6), we see that all four email treatments are negative and significant; both for the proportion of late returns per user and for the average number of days between the return date and the due date per user. For instance, from column (2) we see that the treatment effects (compared to the control) range from -2.4% points for the Reminder to -4.3% points for the Penalty. As for the number of days between the return date and the due date, the reduction lies between 0.54 and 0.87 days. Evaluated at the means of the control group, the treatment effects correspond to a reduction in the proportion of late returns up to 10%, and a reduction of over 100% for the number of days between the return and the due date.[12] Concerning differences in the messages' effectiveness, we cannot reject the hypothesis that all four email contents affect users' behavior in an equal manner. That is, a general reminder of the users' duty to comply with the rule is enough to promote rule compliance, and the additional contents of the other email messages, using appeals to one's contribution on the functioning of public libraries or identifying users as having been late, do not additionally affect behavior. One qualification concerns the penalty treatment: in the specification with proportion late as a dependent variable, the pairwise significance test reveals a higher effectiveness of the penalty treatment compared to the general reminder. Nevertheless, there is certainly no strong evidence that the different messages affect behavior in a different way.[13]

Columns (5) and (6) in Tables 5 and 6 show the estimates for the treatment effects when using a difference-in-differences approach with individual fixed effects. The data for this analysis span both, the pre- and post-treatment periods. As can be seen, these estimates are all negative and very similar in magnitude to those estimates focusing on the post-treatment period (columns (1)–(4)), although the significance level is lower in some cases. Given that the randomization worked properly and that we include controls for the pre-treatment behavior, we from now on focus on treatment effects restricting the sample to the post-treatment period.

We proceed with a finer analysis to shed light on the mechanism behind the average treatment effects. Note first that at the time of receiving the email, there are two types of users: those who have an open balance, that is, a borrowed item from the library, and those who have a closed balance (no borrowed items at the time of the email intervention). To better understand whether the aggregate effect mainly comes from open balances, or whether the email messages also affect timely returns in users' later transactions, we analyze these two types of transactions separately.

Table 7 reports the results when splitting up the data into open and closed transactions. To save on space, we only report estimates including the full set of controls and library fixed effects from now on. We first comment on the results for the users who were randomized into the treatments Control, Reminder and Social (columns (1), (2), (5) and (6)). As can be seen therefrom, the coefficient estimates are quite similar for open and closed transactions, when *Proportion Late* is the dependent variable. For the variable "*Actual-Due*" *Date*, the treatments are more effective for open transactions. When analyzing previously late users (columns (3), (4), (7) and (8)), we see that all treatments are negative and significant for both the open and closed transactions.[14] Again, we find that the treatments are comparable for the dependent variable *Proportion Late*, but slightly more effective for open balance transactions when the dependent variable "*Actual-Due*" *Date* is used.

We conclude that, overall, the average treatment effects in Tables 5 and 6 were not only generated from the loans that were open and pending at the time of the email intervention; instead, the treatments affected all users' behavior, also the behavior of those users who had a closed balance at the time of the email intervention. This is important to understand when we think of the applicability to other settings, as to whether the target population consists of individuals who are currently subject to the rule or also those who may be in the future.

3.2. Duration of the treatment effect

Having shown that receiving an email has a significant effect on behavior, we now address the question related to the duration of the effect. To this end, we partition the post-treatment period into 8 different time windows: (i) July 1–14,

[12] The mean of the control group in the post-period is 0.44 for the proportion late (see the constant in column (1)), and 0.759 for the number of days between return and due date (see the constant in column (5)).

[13] We have also looked at other dependent variables of interest. First, *Any Late*, which takes a value 1 if a user was late at least once in the post-treatment period and 0 otherwise, and *Days Outstanding*, which measures the number of days the item is returned with delay. We find qualitatively similar results with the exception that the Penalty email seems to be slightly more effective when restricting to late users only. Second, *Proportion Renew*, which takes value 1 if a user renewed the item and 0 otherwise. It is conceivable that one way of reducing the proportion of being late is actually renewing the items with higher proportion. We find no evidence for this type of behavior. Finally, note that items vary according to their demand, as some items are taken out only once (low demand) while others are taken out 6 times or more during the post-treatment period (high demand). The problem of the negative externality imposed by late returns is more serious for the highly demanded items as these are the items for which users are more likely waiting. When we restrict the sample to the highly demanded items, we also observe significant and negative effects of the treatment. These results are available upon request.

[14] We have further investigated the effect of emails on transactions which are open and due (which include few observations). Since these users are already late, the relevant dependent variable is "*Actual-Due*" *Date*. We do not find any significant effect.

Table 7
The effect on open and closed transaction.

	Open transactions at the time of the email intervention				**Close transactions at the time of the email intervention**			
	Proportion Late All users (1)	"Actual-Due" (2)	Proportion Late Late users (3)	"Actual-Due" (4)	Proportion Late All users (5)	"Actual-Due" (6)	Proportion Late Late users (7)	"Actual-Due" (8)
Reminder	−0.0102	−0.458	−0.0250*	−0.763*	−0.0130*	−0.266*	−0.0231**	−0.416*
	(0.0106)	(0.313)	(0.0145)	(0.461)	(0.00764)	(0.149)	(0.0108)	(0.214)
Social	−0.0191*	−0.867***	−0.0228	−0.925**	−0.0120	−0.171	−0.0282***	−0.469**
	(0.0105)	(0.309)	(0.0143)	(0.457)	(0.00752)	(0.149)	(0.0105)	(0.215)
Late			−0.0425***	−0.588			−0.0227**	−0.412*
			(0.0144)	(0.470)			(0.0106)	(0.217)
Penalty			−0.0553***	−1.108**			−0.0321***	−0.463**
			(0.0144)	(0.454)			(0.0107)	(0.209)
Constant	0.494***	13.84**	0.382**	10.77**	0.170	−0.263	0.112	−0.185
	(0.140)	(6.944)	(0.149)	(6.243)	(0.167)	(3.618)	(0.186)	(5.361)
Controls and library FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.17	0.15	0.14	0.15	0.18	0.23	0.19	0.23
Number of users	8606	8162	7693	7266	12,101	11,742	10,405	10,042
H0: Reminder = Social (*p*-value)	0.40	0.19			0.89	0.53		
H0: Reminder = Social = 0 (*p*-value)	0.19	0.02			0.16	0.20		
H0: Reminder = Social = Late = Penalty (*p*-value)			0.08	0.70			0.78	0.99
H0: Reminder = Social = Late = Penalty = 0 (*p*-value)			0.00	0.14			0.03	0.14

	Proportion Late	"Actual-Due"	Proportion Late	"Actual-Due"
Cross-Equation Joint Tests (*p-Values*):	(1), (5)	(2), (6)	(3), (7)	(2), (6)
H0: Reminder Open = Reminder Close	0.64	0.01	0.58	0.02
H0: Social Open = Social Close	0.40	0.00	0.98	0.00
H0: Late Open = Late Close			0.85	0.01
H0: Penalty Open = Penalty Close			0.20	0.01

Notes: The table reports treatment effects separately for open and closed balances. The table left-hand side shows the average treatment effects for all items that were open (pending) at the time of receiving the email. The table right-hand side shows the average treatment effects for all items that were taken out after receiving the email. The full set of controls is used, as well as the library fixed effects. See the notes from Tables 5 and 6. Robust standard errors in parenthesis: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

(ii) July 15–31, (iii) August 1–14, (iv) August 15–31, (v) September 1–14, (vi) September 15–30, (vii) October 1–14, and (viii) October 15–November 2. Furthermore, to facilitate interpretation of the results, we restrict the analysis to users who get two emails. With the sample of users being fixed, one can easily distinguish between the effect of the first and the second email.[15]

Table 8 reports the estimates for Eq. (1) separately for the eight time windows. Table 8(a) refers to treatments Control, Reminder, and Social, covering all users, while Table 8(b) reports the results for all five treatments (restricted to previously late users only).

The tables show that the effect of getting an email is short term, but it is replicated after getting a second email. No matter whether we use the proportion of late returns per user as a dependent variable, or the average number of days between the return date and the due date, the effect is negative and mostly significant in the short-term. The first emails that were sent on July 1 had an effect in the period July 1–14 and some of the treatments also in the period July 15–31, but the effect becomes insignificant afterwards. A similar pattern can be observed for the emails that were sent on September 15, even though the effects seem to last a bit longer here. Furthermore, for most email messages (see bottom of Table 8), we cannot reject the null that treatments are the same in the first and the sixth time window. Therefore, users who stopped reacting to the first email react again upon reception of the second message, in a comparable manner.[16]

[15] We also analyzed whether the effect on users who get two emails is significantly different from the effect on users who get only one email. We find that for the proportion of late returns, the effect is larger for users who get two emails. We find no significant differences for the dependent variable "*Actual-Due*" *Date*.

[16] We also did this analysis separately for the open and closed transactions. Interestingly, the effect in the first email is mostly coming from the open transactions at the time of the email intervention, while the effect in the second email is mostly coming from closed transactions at the time of receiving the second email.

Table 8
Treatment effect over time.

	July 1–14 (1)	July 15–31 (2)	August 1–14 (3)	August 15–31 (4)	September 1–14 (5)	**September 15–30** (6)	October 1–14 (7)	October 15–November 2 (8)
(*a* Control-Social-Reminder (all users))								
					Proportion Late			
Reminder	**−0.0372***	−0.00531	0.00729	0.0155	−0.0170	**−0.0179**	−0.0288*	−0.0153
	(**0.0142**)	(0.0139)	(0.0179)	(0.0193)	(0.0197)	(**0.0157**)	(0.0164)	(0.0132)
Social	**−0.0197**	−0.0180	0.00468	0.0222	−0.0143	**−0.0309****	−0.0241	−0.0168
	(**0.0141**)	(0.0138)	(0.0178)	(0.0190)	(0.0194)	(**0.0153**)	(0.0161)	(0.0131)
Constant	**0.752*****	0.0497	0.864***	0.724***	0.571***	**0.705*****	0.191	−0.282***
	(**0.164**)	(0.194)	(0.117)	(0.159)	(0.157)	(**0.270**)	(0.182)	(0.0210)
Controls and library FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	**0.14**	0.12	0.15	0.16	0.12	**0.14**	0.14	0.13
Number of users	**4809**	4815	3058	2768	2660	**3968**	3687	5067
					"Actual-Due" Date			
Reminder	**−1.014****	−0.308	0.175	0.217	−0.158	**−0.529****	−0.329	−0.310
	(**0.399**)	(0.405)	(0.533)	(0.439)	(0.323)	(**0.251**)	(0.243)	(0.200)
Social	**−0.623**	−0.888**	0.00344	0.751*	0.0379	**−0.822*****	−0.322	−0.254
	(**0.412**)	(0.411)	(0.492)	(0.431)	(0.316)	(0.246)	(0.240)	(0.201)
Constant	**18.08****	7.644	5.681	6.777**	2.528	**−20.05*****	−2.593	−11.03***
	(**8.922**)	(8.217)	(4.398)	(2.774)	(2.762)	(**1.485**)	(1.684)	(0.253)
Controls and library FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	**0.15**	0.14	0.20	0.19	0.22	**0.21**	0.22	0.20
Number of users	**4734**	4748	3006	2720	2615	**3893**	3584	4671

Cross-Equation Joint Tests:	Proportion Late	"Actual-Due" Date
H0: Reminder equal for all time intervals	0.82	0.21
H0: Social equal for all time intervals	0.29	0.15
H0: Reminder equal in the first and the sixth time interval	0.23	0.46
H0: Social equal in the first and the sixth time interval	0.31	0.88

	July 1–14 (1)	July 15–31 (2)	August 1–14 (3)	August 15–31 (4)	September 1–14 (5)	**September 15–30** (6)	October 1–14 (7)	October 15–November 2 (8)
(*b* Control-Social-Reminder-Late-Penalty (late users))								
					Proportion Late			
Reminder	**−0.0703*****	−0.0167	0.0159	−0.000351	0.00713	**−0.0312**	−0.0479**	−0.0174
	(**0.0195**)	(0.0190)	(0.0242)	(0.0254)	(0.0257)	(**0.0208**)	(0.0218)	(0.0179)
Social	**−0.0415****	−0.0334*	−0.00920	0.0254	0.0260	**−0.0315**	−0.00634	−0.0321*
	(**0.0194**)	(0.0190)	(0.0241)	(0.0251)	(0.0255)	(**0.0203**)	(0.0218)	(0.0176)
Late	**−0.0335***	0.0208	−0.00956	0.00471	0.0288	**−0.0172**	−0.0534**	−0.0119
	(**0.0200**)	(0.0193)	(0.0244)	(0.0261)	(0.0259)	(**0.0210**)	(0.0218)	(0.0183)
Penalty	**−0.0539*****	−0.0136	0.00614	−0.0128	0.0333	**−0.0235**	−0.0289	−0.0346*
	(**0.0197**)	(0.0193)	(0.0241)	(0.0265)	(0.0262)	(**0.0210**)	(0.0222)	(0.0185)
Constant	**0.629*****	0.212	0.713***	−0.887***	0.540***	**−0.280*****	2.454***	−0.297***
	(**0.146**)	(0.187)	(0.161)	(0.148)	(0.160)	(**0.0457**)	(0.752)	(0.0226)
Controls and library FE	Yes	Yes	Yes	Yes	Yes	**Yes**	Yes	Yes
R-squared	**0.14**	0.12	0.16	0.14	0.13	**0.13**	0.13	0.13
Number of users	**4435**	4388	2935	2646	2612	**3835**	3572	4659
					"Actual-Due" Date			
Reminder	**−1.252****	−0.764	0.589	−0.357	−0.0308	**−0.465**	−0.724**	−0.579**
	(**0.554**)	(0.557)	(0.725)	(0.587)	(0.425)	(**0.332**)	(0.325)	(0.270)
Social	**−0.680**	−1.411**	−0.131	0.660	0.261	**−0.782****	−0.133	−0.403
	(**0.582**)	(0.575)	(0.669)	(0.579)	(0.413)	(**0.333**)	(0.309)	(0.273)
Late	**0.0927**	−0.290	−0.187	0.0225	0.588	**−0.583***	−0.711**	−0.356
	(**0.596**)	(0.583)	(0.678)	(0.614)	(0.437)	(**0.330**)	(0.322)	(0.264)
Penalty	**−0.745**	−0.601	0.0597	−0.648	0.423	**−0.596***	−0.368	−0.0494
	(**0.569**)	(0.583)	(0.653)	(0.561)	(0.434)	(**0.333**)	(0.319)	(0.258)
Constant	**14.24***	16.47	9.661	−21.54**	2.482	**−5.634**	13.69	−10.94**
	(**8.508**)	(14.73)	(6.499)	(3.427)	(2.952)	(**5.497**)	(13.27)	(0.246)
Controls and library FE	Yes	Yes	Yes	Yes	Yes	**Yes**	Yes	Yes
R-squared	**0.15**	0.14	0.24	0.17	0.21	**0.20**	0.21	0.21
Number of users	**4364**	4324	2879	2605	2572	**3762**	3463	4236

Table 8 (*continued*)

Cross-Equation Joint Tests:	Proportion Late	"Actual-Due" Date
H0: Reminder equal for all time intervals	0.19	0.67
H0: Social equal for all time intervals	0.10	0.26
H0: Late equal for all time intervals	0.24	0.33
H0: Penalty equal for all time intervals	0.28	0.51
H0: Reminder equal in the first and the sixth time interval	0.63	0.84
H0: Social equal in the first and the sixth time interval	0.43	0.60
H0: Late equal in the first and the sixth time interval	0.13	0.05
H0: Penalty equal in the first and the sixth time interval	0.99	0.50

Notes: The table reports treatment effects for (a) *Control–Reminder–Social* and (b) *Control–Reminder–Social–Late–Penalty* for different time periods. The sample contains all users which received two emails, the first on July 1 and the second on September 15. The upper part (in panels (a) and (b)) of the table uses *Proportion Late* as a dependent variable and the lower part (in panels (a) and (b)) of the table "Actual-Due" Date. The full set of controls is used, as well as the library fixed effects. See the notes from Tables 5 and 6. Robust standard errors in parenthesis: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

3.3. Heterogenous treatment effects by user characteristics

After estimating the average treatment effect of sending different emails, we now proceed to the analysis of heterogeneous reactions depending on relevant user-specific characteristics. We start testing for differential treatment effects that depend on users' previous behavior. Afterwards, we study whether there are significant differences in behavior depending on gender, age and nationality.[17]

3.3.1. Previous compliance with the rule

It is conceivable that the reaction to the different treatments is related to the users' compliance history, that is, their behavior prior to the treatment. To test for this, we interact the treatment variables with two new variables: *Prior Late* and *Prior "Actual-Due*." *Prior Late* and *Prior "Actual-Due"* refer to the average user-specific proportion of late returns and days between return and due date in the pre-treatment period, respectively.

Table 9 reports the results. We observe that the interaction terms are negative (when significant), suggesting that when there is a differential treatment effect, less rule-compliant users react more strongly. In particular, when the treatment variables are interacted with the user-specific *Prior "Actual-Due"* (columns (2) and (4)), we see that the previous non-compliers have a stronger reaction to the Reminder message, but also to the Late and Penalty messages.[18]

To summarize, the email treatments are especially effective in changing the behavior of a very relevant sample of users, namely those breaking the rule more often. Also, it is important to see that there are no crowding out effects. For users who have a value of *Prior Late* and *Prior "Actual-Due"* equal to 0, the estimated treatment effects are still negative (some of them significant), suggesting a positive effect on the "good types" as well.

3.3.2. Gender, age and nationality

We focus on three demographic variables: users' gender, age and nationality. Table 10 shows the treatment effects for male and female users, for different age groups and for different nationalities. Table 10(a) refers to the treatments Control, Reminder, and Social, covering all users, while Table 10(b) reports the results for all five treatments, restricted to previously late users only. For the sake of brevity, we show results for the dependent variable *Proportion Late*, but find similar effects for the dependent variable "*Actual-Due*" *Date*.

We start commenting on the treatment effects separated by gender. Whether and why gender matters has increasingly attracted economists' attention, and our data offer a rare opportunity to investigate gender differences in the reaction to the different email treatments. Table 10(a) shows that the Reminder email works somewhat better for men than for women, and the opposite for the Social. However, the gender differences are statistically insignificant (see the *p*-values in the third

[17] We also investigated other differential treatment effects that depend on the number of borrowed items as well as the borrowed item types. We did not find any significant differences.

[18] Note that for the dependent variable *Proportion Late* this finding may partly be mechanical due to the fact that this variable is bounded between 0 and 1. However, this is not the case for the variable "*Actual-Due*" *Date*, where we see equally strong evidence that previous non-compliers react more strongly.

Table 9
Differential treatment effect with respect to prior compliance.

	Control–Reminder–Social (all users)		**Control–Reminder–Social–Late–Penalty (late users)**	
	Proportion Late (1)	"Actual-Due" Date (2)	Proportion Late (3)	"Actual-Due" Date (4)
Reminder	−0.00243	−0.516***	0.00488	−0.381
	(0.00957)	(0.194)	(0.0181)	(0.273)
Social	−0.0137	−0.401**	−0.0260	−0.450
	(0.00944)	(0.203)	(0.0177)	(0.275)
Late			−0.0321*	−0.414
			(0.0178)	(0.269)
Penalty			−0.0347*	−0.705***
			(0.0177)	(0.268)
Prior Late	0.334***		0.307***	
	(0.0159)		(0.0236)	
Reminder*Prior Late	−0.0358		−0.0581*	
	(0.0229)		(0.0338)	
Social*Prior Late	−0.0146		−0.00728	
	(0.0223)		(0.0330)	
Late*Prior Late			0.0100	
			(0.0330)	
Penalty*Prior Late			−0.0170	
			(0.0329)	
Prior "Actual-Due"		0.231***		0.230***
		(0.0279)		(0.0298)
Reminder*Prior "Actual-Due"		−0.0826**		−0.124***
		(0.0395)		(0.0426)
Social*Prior "Actual-Due"		0.00652		−0.0519
		(0.0399)		(0.0467)
Late*Prior "Actual-Due"				−0.0715
				(0.0442)
Penalty*Prior "Actual-Due"				−0.0910**
				(0.0382)
Constant	0.375**	4.878	0.157	4.410
	(0.160)	(4.337)	(0.121)	(3.955)
Controls and library FE	Yes	Yes	Yes	Yes
R-squared	0.17	0.14	0.16	0.14
Number of users	14,442	13,990	12,205	11,750
H0: Reminder = Social (*p*-value)	0.25	0.54	0.11	0.81
H0: Reminder = Late (*p*-value)			0.05	0.90
H0: Reminder = Penalty (*p*-value)			0.39	0.25
H0: Late = Penalty (*p*-value)			0.83	0.30
H0: Reminder = Social = Late = Penalty (*p*-value)			0.14	0.64
H0: Reminder = Social = Late = Penalty = 0 (*p*-value)			0.11	0.17
H0: Reminder = Social = 0 (*p*-value)	0.33	0.02		
H0: Reminder*Comp = Social*Comp (*p*-value)	0.32	0.00	0.13	0.01
H0: Reminder*Comp = Late*Comp (*p*-value)			0.04	0.05
H0: Reminder*Comp = Penalty*Comp (*p*-value)			0.21	0.25
H0: Late*Comp = Penalty*Comp (*p*-value)			0.40	0.51
H0: Reminder*Comp = Social*Comp = Late*Comp = Penalty*Comp (*p*-value)			0.20	0.07
H0: Reminder*Comp = Social*Comp = Late*Comp = Penalty*Comp = 0 (*p*-value)			0.28	0.00
H0: Reminder*Comp = Social*Comp = 0 (*p*-value)	0.24	0.00		

Notes: This table reports differential treatment effects with respect to prior compliance. *Prior Late* measures the user-specific proportion of items that were returned late in the pre-treatment period. *Prior "Actual-Due"* measures the average number of days between the return date and the due date per user in the pre-treatment period. The full set of controls is used, as well as the library fixed effects. See the notes from Tables 5 and 6. Robust standard errors in parenthesis: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

column). Table 10(b) confirms this result for all five treatments. Therefore, women and men are very much comparable when it comes to the reaction to messages aimed at promoting rule compliance.[19]

[19] Our results are in line with experimental findings that display a complex picture of gender. For instance, women do not seem to be more social or fairer per se, but it depends on the circumstances (Croson and Gneezy, 2009; Andreoni and Vesterlund, 2001).

Table 10
Heterogeneous treatment effect on Proportion Late: gender, age, nationality. (a) All users and (b) late users.

	Gender			Age				H0: (1)=(2)=(3)=(4) (*p*-Value)
	Male (1)	Female (2)	H0: (1)=(2) (*p*-Value)	Age under 20 (1)	Age 20–40 (2)	Age 40–60 (3)	Age over 60 (4)	
(*a*)								
Reminder	−0.0213*	−0.00845	0.38	0.0103	−0.0221**	−0.0160	0.00960	0.43
	(0.0109)	(0.00956)		(0.0210)	(0.00967)	(0.0135)	(0.0276)	
Social	−0.0146	−0.0204**	0.68	−0.0183	−0.0192**	−0.0257*	0.0136	0.78
	(0.0107)	(0.00944)		(0.0205)	(0.00948)	(0.0138)	(0.0273)	
Constant	0.391	0.374**		0.491	0.374	0.268	0.108*	
	(0.250)	(0.173)		(0.338)	(0.243)	(0.176)	(0.0582)	
Controls and library FE	Yes	Yes		Yes	Yes	Yes	Yes	
R-squared	0.18	0.16		0.14	0.16	0.16	0.23	
Number of users	6170	8272		1969	8325	3486	662	

	Nationalities							
	Spain (1)	North Europe (2)	South Europe (3)	English Speaking (4)	Eastern-Russia (5)	Latin America (6)	Asia (7)	H0: (1)=(2)=(3)=⋯=(7) (*p*-Value)
Reminder	−0.0111	−0.00536	−0.00472	−0.123**	0.0915	−0.0254	−0.0666	0.34
	(0.00842)	(0.0556)	(0.0316)	(0.0558)	(0.0697)	(0.0181)	(0.0836)	
Social	−0.0148*	−0.0456	0.00147	−0.207***	−0.0164	−0.0227	0.101	0.04
	(0.00833)	(0.0562)	(0.0311)	(0.0547)	(0.0634)	(0.0176)	(0.0781)	
Constant	0.260***	0.411***	0.280***	0.294***	0.182*	0.308***	−0.0764	
	(0.0141)	(0.120)	(0.0823)	(0.110)	(0.109)	(0.0323)	(0.126)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
R-squared	0.15	0.13	0.19	0.26	0.23	0.17	0.41	
Number of users	10,395	265	745	224	185	2369	84	

(*b*)

	Gender			Age				H0: (1)=(2)=(3)=(4) (*p*-Value)
	Male (1)	Female (2)	H0: (1)=(2) (*p*-Value)	Age under 20 (1)	Age 20–40 (2)	Age 40–60 (3)	Age over 60 (4)	
Reminder	−0.0295*	−0.0193	0.65	0.00779	−0.0343***	−0.0106	−0.0486	0.50
	(0.0152)	(0.0135)		(0.0321)	(0.0128)	(0.0202)	(0.0585)	
Social	−0.0231	−0.0331**	0.57	−0.0254	−0.0241*	−0.0390*	−0.0782	0.77
	(0.0146)	(0.0134)		(0.0312)	(0.0125)	(0.0206)	(0.0510)	
Late	−0.0334**	−0.0214	0.57	0.0195	−0.0328***	−0.0177	−0.107**	0.17
	(0.0147)	(0.0134)		(0.0312)	(0.0125)	(0.0206)	(0.0490)	
Penalty	−0.0377**	−0.0482***	0.56	0.00236	−0.0456***	−0.0480**	−0.0877	0.38
	(0.0151)	(0.0132)		(0.0309)	(0.0126)	(0.0209)	(0.0543)	
Controls and library FE	Yes	Yes		Yes	Yes	Yes	Yes	
Constant	0.0459	0.188		−0.0151	0.154	0.528***	−0.287	
	(0.118)	(0.162)		(0.106)	(0.197)	(0.183)	(0.968)	
R-squared	0.18	0.16		0.13	0.16	0.16	0.25	
Number of users	5311	6894		1416	7770	2609	410	

	Nationalities							
	Spain (1)	North Europe (2)	South Europe (3)	English Speaking (4)	Eastern-Russia (5)	Latin America (6)	Asia (7)	H0: (1)=(2)=(3)=⋯=(7) (*p*-Value)
Reminder	−0.0184	0.0258	0.0133	−0.224***	0.0315	−0.0401*	−0.206**	0.09
	(0.0123)	(0.0681)	(0.0395)	(0.0677)	(0.0934)	(0.0237)	(0.0913)	
Social	−0.0253**	−0.0150	0.0139	−0.295***	−0.0788	−0.0256	−0.0211	0.03
	(0.0121)	(0.0662)	(0.0393)	(0.0685)	(0.0893)	(0.0226)	(0.110)	
Late	−0.0290**	0.0840	−0.0156	−0.186***	0.0615	−0.0108	−0.207**	0.09
	(0.0121)	(0.0640)	(0.0399)	(0.0686)	(0.0887)	(0.0233)	(0.101)	
Penalty	−0.0427***	0.0263	−0.0463	−0.245***	−0.0659	−0.0111	−0.280***	0.04
	(0.0122)	(0.0707)	(0.0392)	(0.0608)	(0.0904)	(0.0233)	(0.105)	
Constant	0.274***	0.299**	0.217***	0.444***	0.112	0.361***	0.0430	
	(0.0183)	(0.146)	(0.0807)	(0.140)	(0.133)	(0.0384)	(0.131)	

Table 10 (*continued*)

	Nationalities							
	Spain (1)	North Europe (2)	South Europe (3)	English Speaking (4)	Eastern-Russia (5)	Latin America (6)	Asia (7)	H0: (1)=(2)=(3)=⋯=(7) (*p*-Value)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
R-squared	0.14	0.26	0.20	0.29	0.27	0.16	0.52	
Number of users	8198	241	778	230	195	2308	79	

Notes: The table shows heterogeneous treatment effects of (a) *Control–Reminder–Social* and (b) *Control–Reminder–Social–Late–Penalty* for the dependent variable *Proportion Late*. The sample is split with regard to gender, age and nationalities. The full set of controls is used. See the notes from Tables 5 and 6. Robust standard errors in parenthesis:* $p<0.1$, ** $p<0.05$, *** $p<0.01$.

With respect to age, we classify users in four age groups: below 20 years of age, between 20 and 40, between 40 and 60, and above 60. Note that in Table 10 most users belong to the age group between 20 and 40, while other age groups are less represented. The effect is always significant for the main age group, between 20 and 40, and less often for the other age groups. However, we cannot reject the null that the effect is the same for the different four age groups.

Finally, let us turn to nationalities.[20] Our database allows us to distinguish between the users' countries of origin. Hence, we can evaluate whether the behavior of users differs by nationality, and whether there are differential reactions to receiving an email. We classify users into 7 geographical areas according to their nationality: (i) Spain, (ii) Northern and Central Europe (Germany, Belgium, Denmark, Finland, Netherlands, Norway, Sweden, Switzerland, and Austria), (iii) Southern and Western Europe (France, Italy, Greece, and Portugal), (iv) English speaking countries (UK, US, Canada, Ireland, and Australia), (v) Eastern Europe and Russia (Bulgaria, Croatia, Slovakia, Estonia, Hungary, Lithuania, Poland, Rumania, Russia, Czech Republic, Ukraine, Georgia, and Armenia), (vi) Latin America (Argentina, Bolivia, Brazil, Colombia, Cuba, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Paraguay, Peru, El Salvador, Uruguay, Venezuela, Chile, Costa Rica, and Panama), and (vii) Asia (Philippines, Japan, Nepal, China, India, and South Korea). Spain accounts for the vast majority of users (around 70%), followed by Latin America with 19%, Southern and Western Europe with 6%, and at the bottom of the distribution is Asia with 1%.

Table 10 shows that there are remarkable differences in the reaction to the treatments. First, users from English speaking countries react significantly to every single treatment. They reduce the proportion of late returns by up to 30% points.[21] Previously late users from Asia also react significantly, in particular to the treatments Reminder, Late and Penalty. With the exception of Spain, we do not find consistent and significant effects for the other nationality groups.

Furthermore, when comparing the effects found for Spaniards with the effects found for English speaking and Asian countries, controlling for different initial propensities of being late, as well as different reactions depending on prior propensity to be late, we still find that users from English speaking countries and Asian users react significantly more than Spaniards. Therefore, culture seems to influence how much users react to messages targeted at rule compliance.

4. Conclusions

In this paper we study the effect of a very simple, versatile, and virtually costless mechanism, such as sending email messages, on promoting compliance with rules. The study was conducted in the Public Libraries of Barcelona, where compliance with rules means returning items on time. What makes our setting unique is that we observe a large number of users in a daily-life situation, where rules are simple, well-defined, with low attached penalties, and where compliance is perfectly measurable. Our paper involves, therefore, the study of compliance with a rule where the consequences of rule-breaking are relatively minor. We believe that these characteristics of rule compliance are pervasive in day-to-day life, and therefore crucial for a proper functioning of organizations and institutions

Using the methodology of a randomized field experiment, we show that sending email messages helps to promote compliance with rules. The treatments reduce the proportion of late returned items by up to 10% and the number of days between the return and the due date by up to almost one day. These effects are not only statistically significant but also economically relevant, especially in light of the negligible costs associated with the intervention. A general reminder of the users' duty to comply with the rule is effective in promoting rule compliance. Furthermore, adding other contents to the

[20] There is sound evidence that nationality is an important determinant of behavior in a variety of settings. Most related to our study, there are interesting nationality differences in the determinants of corruption (Fisman and Miguel, 2007) and cooperation (Herrmann et al., 2008; Gächter et al., 2010).

[21] Note that this result is not driven by one single country. The same result holds even if we analyze UK/Ireland and US/Canada separately.

general reminder, appealing to one's contribution on the functioning of public libraries or identifying users as having been late in the past, do not bring a significant additional increase in rule compliance.

The messages provoke the biggest reactions in users who complied the least in the past. Interestingly, even the “good citizens” react positively to receiving an email. Hence, the email treatment is more effective precisely with those users whose compliance prior to the treatment was lower, and, importantly, does not generate crowding-out effects in those users that were complying with the rule before the intervention.

A natural interpretation of our results is that individuals pay limited attention to the duty to return items on time, and an email reminding them about this duty increases rule compliance (see Karlan et al., 2011, for a model of limited attention, and empirical evidence supporting it). This interpretation is consistent with our main results. Adding extra messages to the reminder email does not significantly change behavior, and precisely those users who were complying the least are the ones that react the most. An alternative interpretation of our results is that users interpret the treatment emails as a signal that the libraries care about rule compliance (maybe to a greater extent than expected), and that this triggers a positive reaction on them. According to this interpretation, users react to their beliefs about what the authorities expect from them. Although the result that a simple reminder increases rule compliance has important implications for organizations and authorities, future research should be directed to disentangle among the two possible interpretations given above.

Acknowledgments

We are grateful to Anna Surroca (*Serveis de Biblioteques of the Diputacio de Barcelona*), Judit Terma (*Consorci de Biblioteques de Barcelona*), and their respective teams, for their collaboration and assistance in conducting this study. We thank Antonio Cabrales, Pablo Casas, Vincent P. Crawford, Laura Gee, Rosemarie Nagel, Imran Rasul, Pedro Rey-Biel, Carmit Segal, Alessandro Tarozzi, Christian Traxler and seminar participants at various universities for helpful comments. Special thanks go to Stephan Litschig for constant feedback and discussions. Sergio Correia provided excellent research assistance. Financial support from the Spanish Commission of Science and Technology (ECO2010-09555-E, ECO2011-25272, ECO2011-25295, ECO2012-34202, ECO2012-31626) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.euroecorev.2013.08.010.

References

Allison, Dale, Strauss, Aaron, 2009. Don't forget to vote: text message reminders as a mobilization tool. American Journal of Political Science 53 (4), 787–804.
Andreoni, James, Miller, John, 2002. Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica 70 (2), 737–753.
Andreoni, James, Rao, Justin M., 2010. The Power of Asking: How Communication Affects Selfishness, Empathy and Altruism, mimeo.
Andreoni, James, Vesterlund, Lise, 2001. Which is the fair sex? Gender differences in Altruism. Quarterly Journal of Economics 116, 293–312.
Ayres, Ian, Raseman, Sophie, Shi, Alice, 2009. Evidence From Two Large Field Experiments that Peer Comparison Feedback Can Reduce Energy Usage, NBER W.P. 15386.
Battigalli, Pierpaolo, Dufwenberg, Martin, 2007. Guilt in games. American Economic Association, Papers and Proceedings 97 (2), 170–176.
Becker, Gary, 1968. Crime and punishment: an economic approach. Journal of Political Economy 76, 169–217.
Benabou, Roland, Tirole, Jean, 2003. Intrinsic and extrinsic motivation. Review of Economic Studies 70, 489–520.
Benabou, Roland, Tirole, Jean, 2006. Incentives and pro-social behavior. American Economic Review 96 (5), 1652–1678.
Cadena, Ximena, Schoar, Antoinette, 2011. Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments, mimeo.
Camerer, Colin, Issacharoff, Samuel, Loewenstein, George, O'Donoghue, Ted, Rabin, Matthew, 2003. Regulation for conservatives: behavioral economics and the case for asymmetric paternalism. University of Pennsylvania Law Review 1151 (3), 1211–1254.
Charness, Gary, Dufwenberg, Martin, 2006. Promises and partnership. Econometrica 74 (6), 1579–1601.
Charness, Gary, Dufwenberg, Martin, 2011. Participation. American Economic Review 101, 1213–1239.
Charness, Gary, Rabin, Matthew, 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics 117 (3), 817–869.
Crawford, Sue E.S., Ostrom, Elinor, 1995. A grammar of institutions. American Political Science Review 89 (3), 582–600.
Croson, Rachel, Gneezy, Uri, 2009. Gender differences in preferences. Journal of Economic Literature 47 (2), 1–27.
Dal Bó, Ernesto, Dal Bó, Pedro, 2010. Do the Right Thing: The Effect of Moral Suasion on Cooperation, mimeo.
Dal Bó, Pedro, Foster, Andrew, Putterman, Louis, 2010. Institutions and behavior: experimental evidence on the effects of democracy. American Economic Review 100, 2205–2229.
Duflo, Esther, Saez, Emmanuel, 2003. The role of information and social interactions in retirement plan decision: evidence from a randomized experiment. Quarterly Journal of Economics 118, 815–842.
Ellingsen, Tore, Johannesson, Magnus, 2004. Promises, threats and fairness. Economic Journal 114 (495), 397–420.
Ellingsen, Tore, Johannesson, Magnus, 2008. Anticipated verbal feedback induces altruistic behavior. Evolution and Human Behavior 29 (2), 100–105.
Fellner, Gerlinde, Sausgruber, Rupert, Traxler, Christian 2013. Testing enforcement strategies in the field: threat, moral appeal and social information. Journal of the European Economic Association, 11 (3), 634–660.
Fisman, Ray, Miguel, Edward, 2007. Corruption, norms, and legal enforcement: evidence from diplomatic parking tickets. Journal of Political Economy 115 (6), 1020–1048.
Frey, Bruno, Jegen, Reto, 2001. Motivation crowding theory: a survey of empirical evidence. Journal of Economic Surveys 15 (5), 589–611.
Gächter, Simon, Herrmann, Benedikt, Thöni, Christian, 2010. Culture and cooperation. Philosophical Transactions of the Royal Society B—Biological Sciences 365 (1553), 2651–2661.

Gerber, Alan S., Green, Donald P., Larimer, Christopher W., 2008. Social pressure and voter turnout: evidence from a large-scale field experiment. American Political Science Review 102, 33–48.
Gneezy, Uri, Rustichini, Aldo, 2000a. Pay enough or don't pay at all. Quarterly Journal of Economics 115 (3), 791–810.
Gneezy, Uri, Rustichini, Aldo, 2000b. A fine is a price. Journal of Legal Studies 29 (1), 1–18.
Herrmann, Benedikt, Thöni, Christian, Gächter, Simon, 2008. Antisocial punishment across societies. Science 319, 1362–1367.
Huck, Steffen, Rasul, Imran, 2010. Transaction costs in charitable giving: evidence from two field experiments. B.E. Journal of Economic Analysis and Policy, 10(1) (Advances, Art. 31).
Jensen, Robert, 2010. The (perceived) returns to education and the demand for schooling. Quarterly Journal of Economics 125 (2), 515–548.
Karlan, Dean, McConnell, Margaret, Mullainathan, Sendhil, Zinman, Jonathan, 2011. Getting to the Top of Mind: How Reminders Increase Saving, mimeo.
Kast, Felipe, Meier, Stephan, Pomeranz, Dina, 2012. Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device, mimeo.
Mellstroem, Carl, Johannesson, Magnus, 2008. Crowding out in blood donation: was titmuss right?. Journal of the European Economic Association 6 (4), 845–863.
Ostrom, Elinor, 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press.
Pomeranz, Dina, 2010. No Taxation without Information, mimeo.
Pruckner, Gerald J., Sausgruber, Rupert, 2013. Honesty on the streets: a field study on newspaper purchasing. Journal of the European Economic Association 11 (3), 661–679.
Schultz, P. Wesley, Nolan, Jessica M., Cialdini, Robert B., Goldstein, Noah J., Griskevicius, Vladas, 2007. The constructive, destructive, and reconstructive power of social norms. Psychological Science 18, 429–434.
Sobel, Joel, 2005. Interdependent preferences and reciprocity. Journal of Economic Literature 53, 392–436.
Szilagyi, Peter G., Adams, William G., 2012. Text messaging a new tool for improving preventive services. Journal of the American Medical Association 307 (16), 1748–1749.
Thaler, Richard H., Sunstein, Cass R., 2008. Nudge: Improving Decisions About Health, Wealth, and Happiness. Yale University Press.