DISCUSSION PAPER SERIES

DP13677

BLOOD DONATIONS AND INCENTIVES: EVIDENCE FROM A FIELD EXPERIMENT

Lorenz Goette and Alois Stutzer

PUBLIC ECONOMICS

ISSN 0265-8003

BLOOD DONATIONS AND INCENTIVES: EVIDENCE FROM A FIELD EXPERIMENT

Lorenz Goette and Alois Stutzer

Discussion Paper DP13677
Published 16 April 2019
Submitted 15 April 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

BLOOD DONATIONS AND INCENTIVES: EVIDENCE FROM A FIELD EXPERIMENT

Abstract

There is a longstanding concern that material incentives might undermine pro-social motivation, leading to a decrease in blood donations rather than an increase. This paper provides an empirical test of how material incentives affect blood donations in a large-scale field experiment spanning three months and involving more than 10,000 previous donors. We examine two types of rewards: a lottery ticket and a free cholesterol test. Lottery tickets significantly increase donations during the experiment, in particular among less motivated donors. Moreover, no reduction in donations is observed after the experiment. The free cholesterol test leads to no discernable impact on blood donations during and after the experiment.

JEL Classification: C93, D64, H41, I18

Keywords: blood donations, field experiment, material incentives, motivation crowding effect, pro-social behavior

Lorenz Goette - lorenz.goette@uni-bonn.de
University of Bonn and CEPR

Alois Stutzer - alois.stutzer@unibas.ch
University of Basel

Acknowledgements
We thank Tyler Williams and Lim Zhi Hao for excellent research assistance, Roland Bénabou, Anat Bracha, Tore Ellingsen, Justina Fischer, Holger Herz, Stephan Meier, Pierre-Guillaume Méon, Kurt Schmidheiny, and the participants of seminars at the University of Basel, Cornell University, the University College Dublin, the University of Freiburg i.B., the University of Lausanne, the University of Lucerne, the Singapore Management University, Stanford University, the University of St. Gallen, the University of Warwick, the University of Zurich, the annual meetings of the American Economic Association, the European Economic Association, the European Public Choice Society, the German Economic Association, the International Society of Blood Transfusion, and the End-of-Year Conference of the Network of Swiss Economists Abroad for helpful comments and the Stiftung Zürcher Blutspendedienst SRK, especially Beat Frey and Heinz Jehle, for their productive collaboration in the project.

Blood Donations and Incentives: Evidence from a Field Experiment*

Lorenz Goette† and Alois Stutzer‡

April 13, 2019

Abstract

There is a longstanding concern that material incentives might undermine pro-social motivation, leading to a decrease in blood donations rather than an increase. This paper provides an empirical test of how material incentives affect blood donations in a large-scale field experiment spanning three months and involving more than 10,000 previous donors. We examine two types of rewards: a lottery ticket and a free cholesterol test. Lottery tickets significantly increase donations during the experiment, in particular among less motivated donors. Moreover, no reduction in donations is observed after the experiment. The free cholesterol test leads to no discernable impact on blood donations during and after the experiment.

Keywords: blood donations, field experiment, material incentives, motivation crowding effect, pro-social behavior

JEL classification: C93, D64, H41, I18

*We thank Tyler Williams and Lim Zhi Hao for excellent research assistance, Roland Bénabou, Anat Bracha, Tore Ellingsen, Justina Fischer, Holger Herz, Stephan Meier, Pierre-Guillaume Méon, Kurt Schmidheiny, and the participants of seminars at the University of Basel, Cornell University, the University College Dublin, the University of Freiburg i.B., the University of Lausanne, the University of Lucerne, the Singapore Management University, Stanford University, the University of St. Gallen, the University of Warwick, the University of Zurich, the annual meetings of the American Economic Association, the European Economic Association, the European Public Choice Society, the German Economic Association, the International Society of Blood Transfusion, and the End-of-Year Conference of the Network of Swiss Economists Abroad for helpful comments and the Stiftung Zürcher Blutspendedienst SRK, especially Beat Frey and Heinz Jehle, for their productive collaboration in the project.

†Institute for Applied Microeconomics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany. Phone: +49 228 73 9240, email: lorenz.goette@uni-bonn.ch

‡Faculty of Business and Economics, University of Basel, Peter Merian-Weg 6, 4002 Basel, Switzerland. Phone: +41 61 207 33 61, email: alois.stutzer@unibas.ch

I. Introduction

In medical emergencies, blood transfusions are often the only way to save individuals' lives (Higgins, 1994). A sufficient supply of donated blood is thus literally a matter of life and death. Blood cannot yet be produced artificially, and some components of blood can only be stored for a short period of time. Moreover, the amount an individual can donate is limited. Thus, in order to meet the need for blood, a wide and healthy base of donors, willing to give blood when required, is needed. Historically, many blood donation services have relied on voluntary, non-remunerated donations and thus on the pro-social motivation of their donors (Slonim et al., 2014).[1] Despite the inherent free-rider problem of this policy, the arrangement seems to have worked satisfactorily most of the time. However, it also faces serious challenges. New techniques in surgeries and oncological therapies require larger amounts of blood (Davey, 2004). There is also a general tightening of eligibility criteria, such as stepped-up travel restrictions or restrictions on past blood recipients. Finally, there are widespread seasonal shortages (Gilcher and McCombs, 2005).

In this paper, we present evidence on whether rewards can be used to overcome shortages in an environment that relies primarily on the pro-social motivations of volunteer donors. Most economic models, including those incorporating pro-social preferences explicitly (for example, Andreoni, 1990), predict that selective incentives increase blood donations.[2] Yet, there is a deep-rooted skepticism about using incentives in blood donations even on a temporary basis. That skepticism is based on the conjecture that using incentives may undermine the motivation

[1] While whole-blood donations are generally unpaid in developed countries (World Health Organization, 2017), there are some countries where plasma donations, which can be offered frequently, are paid. Indeed, Trimmel, Lattacher and Janda (2005) find that, using survey measures, whole-blood donors are broadly more pro-socially motivated than plasma donors. Intriguingly, no difference was found in the survey when comparing whole-blood donors with plasma donors who would be willing to continue donating plasma if no longer paid. Similar differences were found when comparing whole-blood donors with the general population (Fernández Montoya et al., 1996). See Piliavin (1990) for additional references.

[2] We use the term incentive to describe an external intervention affecting the relative attractiveness of alternative options with no specific reference to the substance of the incentive. Instead when referring to rewards, we consider specific incentives, i.e. quid pro quo offers in a voluntary exchange.

to donate blood.[3,4]

Research in psychology and economics has proposed that incentives can lead to less pro-social behavior. If rewards are perceived as controlling, incentives may undermine pro-social motivation. This is often referred to as the motivation crowding effect (Deci and Ryan, 1985; Frey, 1997; Lepper and Greene, 1978). There is ample evidence from various contexts in support of this mechanism (see the reviews by Frey and Jegen 2001; Gneezy et al. 2011; Kamenica 2012) and it has repeatedly been discussed in the context of blood donation policy (see, e.g., Archard 2002; Lacetera et al. 2013; Stutzer and Goette 2010). However, the specific evidence is less clear. In a comprehensive review of the literature on the use of incentives to stimulate blood donation, Chell et al. (2018) conclude that many issues remain unresolved. One of the reasons they emphasize is that the research is "confounded by current operating context" (p. 251).[5]

Our natural field experiment allows us to study the effects of material rewards in a controlled context that is particularly attractive to pit the theoretical predictions of a traditional model against a model of motivation crowding.[6] In our study, we randomized the offer of rewards to roughly 10,000 blood donors who had not been offered any kind of reward in exchange for their blood before. We can thus learn about the impact of incentives on individuals who chose to donate in the absence of compensation. Our data include donation records of these individuals before, during, and after the experiment. This design allows us to evaluate how the motivation to donate changes i) when incentives are applied, but also ii) for the period after having been exposed to incentives. The data on donation behavior before the experiment provides us with a natural measure of how motivated individuals are to donate in the absence of incentives.

In our key experimental conditions, individuals were offered a lottery ticket in return for

[3]This conjecture is associated with Titmuss (1971), who famously argued, "From our study of the private market in blood in the United States, we have concluded that the commercialization of blood donor relationships represses the expression of altruism, erodes the sense of community, [...]" (p. 245).

[4]A second concern is that using incentives may attract at-risk donors. This concern holds in particular with regard to new donors (Eastlund, 1998; Sanchez et al., 2001; Van der Poel et al., 2002). As our experiment studies the behavioral reactions of previous donors, we focus on the effects of incentives on contributions and refer to the working paper (Goette and Stutzer 2008) for an analysis of the selection issue.

[5]Other reviews and narrative summaries are offered in Bagot et al. (2016), Goette et al. (2010), Lacetera et al. (2013), and Niza et al. (2013).

[6]The criticism of lab experiments by Levitt and List (2007) makes the conditions of our field experiment particularly desirable, as the participants did not know they were involved in a study.

donating over a pre-specified time window. This information was included on a postcard added to the standard invitation letter that the individuals were used to receive. In a second condition, individuals were offered a free cholesterol test. We complemented these two incentive treatments with two control conditions. In one, individuals also received an additional postcard, asking them to donate during the same time period as in the other two conditions, in order to control for possible effects of the perceived urgency of donating (Bruhin et al., 2015; Sun et al., 2016). The second control condition simply consisted of the standard invitation letter. Importantly, our subjects were not aware that a study was being conducted.

The different invitations were mailed privately to the donors, and donations occurred in the anonymity of a medical center or hospital, so that public image concerns are largely excluded in our experiment (see Lacetera and Macis, 2010, for a demonstration of the relevance of public-image concerns in blood donations). Thus, our study can be viewed as a test of psychological crowding theories related to intrinsic motivation or self-signaling (Bénabou and Tirole, 2006).

The overall picture that emerges from our experiment is that material rewards have no general negative effects, neither during the experiment, nor in the long-term. To the contrary, when looking at the overall experimental outcomes, we find that offering a lottery ticket increases donations by 5.6 percentage points over a baseline donation rate of 46 percent. In contrast, we find no economically and statistically significant effect of the free cholesterol test, and the simple appeal to donate on behavior during the experiment.[7]

We also show that the treatment effects vary between subsamples in interesting ways. Our treatment effects are entirely driven by individuals with a low previous motivation to donate blood. Even though this group has a low baseline donation rate of about 30 percent, offering a lottery ticket increases the probability of a donation by around 8 percentage points. By contrast, there is essentially no incentive effect on donors who have previously shown a high motivation to donate blood. This pattern shows that heterogeneity in the motivation for blood donation

[7] In an earlier analysis comparing attitudes towards rewards with actual behavioral reactions to them, we have already published the result for the minimal effect of free cholesterol testing on donating blood (Goette et al., 2009). In particular, we report a disconnect between attitudes and behavior when it comes to the evaluation of free cholesterol testing as a means to encourage donors to donate more often. This underscores the relevance of field experiments in comparison to survey studies on the effectiveness of rewards in motivating donors.

modulates the impact of incentives.

When we examine different splits of the sample, such as by age, gender, or regularity of the invitation schedule chosen, we find no significant differences in the responses to the treatments between these subsamples. This reinforces our interpretation that differences in the motivation to donate blood are the driving force behind the different responses.

In addition, our design allows us to examine whether the material rewards changed pro-social motivation in the long run by following the donors for up to 15 months after the conclusion of the experiment. After the experiment, donations generated from people who were incentivized during the experiment are no lower than those generated from people who received the rewards unexpectedly, and the precision of our estimates allows us to reject even small negative effects.

The findings from our study add to the understanding of which operating contexts offer a fruitful ground for applying incentives, and highlight other areas where results across studies are more discordant (Chell et al., 2018). Perhaps surprisingly, the results are consistent with previous evidence for monetary or near-monetary rewards such as stored-value cards. Lacetera et al. (2012) and Lacetera et al. (2014) show that such rewards reliably lead to higher turnout in blood drives of the American Red Cross (ARC). Importantly, the individuals in these two studies are used to receiving some form of a reward.[8] The results from our study population, on which material rewards had not been used before, are very similar.

By contrast, there are large differences in how other forms of rewards affect blood donations. As pointed out, the offer of a free cholesterol test does not have a positive effect on donation rates, either overall or in any of the subpopulations that we examined. This sharply contrasts with the findings in Leipnitz et al. (2018), which uses a difference-in-differences strategy to assess the effectiveness of offering blood screening on donations. They find large positive effects from offering a blood check as a reward.

Furthermore, we find that merely appealing to previous donors to give blood has some positive effect for individuals with a relatively low baseline motivation. Another recent study by

[8] Lacetera et al. (2014) state that "[b]ecause about 40% of ARC drives offer a promotional item, and most flyers show at least one drive with a promotion, the reward offers should not be perceived as unusal" (p. 1111).

Sun et al. (2016) finds strong positive effects of a similar message that was delivered by a text message (rather than postal mail as in our case).

The remainder of this paper is structured as follows. Section II describes the empirical setup of our study, explains the details of the treatments, and offers some descriptive statistics. Section III discusses the behavioral predictions for the different treatments. The section also elaborates how the standard model, augmented by a dynamic mechanism to allow for habit formation or guilt, can produce longer-run responses to temporary incentives. Section IV presents the results for the donation behavior during and after the experimental intervention. Section V offers concluding remarks.

II. The Empirical Setup

We conducted a field experiment spanning three months in the summer of 2006 in four blood donation centers in the canton of Zurich, Switzerland. The study was conducted in close collaboration with the Zurich Blood Donation Service of the Swiss Red Cross (Stiftung Zürcher Blutspendedienst, henceforth SRC). The subjects participating in this field experiment were individuals registered in the database of the blood donation service.

If an individual has ever donated blood in one of the four donation services, he or she is registered in the database of the SRC. The individual is subsequently invited to donate blood again at one of the four donation centers (the center is determined by proximity to the town the individual lives in). The donors are mailed an invitation giving a specific date, approximately three weeks prior to the appointment in order to avoid congestion at the donation center. The SRC starts inviting eligible individuals four months after their last invitation (or longer, if the donors so indicate) and sends these invitations in no particular order.

If an individual has ignored seven consecutive invitations, the SRC stops sending them invitations.

Figure 1 provides an impression of the frequency with which the invited blood donors followed the invitation in the 12 months before the experiment began in May 2006. It shows that the fraction of individuals donating subsequent to an invitation is quite volatile, and between 40 and 50 percent in most months.[9] Figure 2 shows the overall donation rate during the experiment. It is 47 percent and thus slightly higher than the average donation rate during the corresponding months one year earlier. The figure also reveals important persistent differences in the motivation to donate blood across donors. Restricted to donors who have received at least four invitations prior to the experiment, we display the fraction of individuals following the invitation during the experiment as a function of how many of the previous invitations they followed. The figure shows that of those who followed none of the 4 previous invitations, only about 15 percent donated during the experiment. This fraction increases monotonically to over 80 percent for those who had followed all four previous invitations. Thus, there is strong heterogeneity between donors, and the figure suggests that the number of previous invitations followed is a strong predictor of the overall willingness to donate blood during the experiment. It is important to note that in each group, there is ample possibility to adjust donations upward or downward. Thus, mechanical "ceiling effects" are unlikely to play a role in our experiment.

A. Treatments

We implemented four experimental groups. One group served as a control to identify the effects of the control variables and was invited as usual. The remaining three groups received the standard invitation letter with an additional feature: for each of the three treatments, a postcard was included. The face side of the post card read "This summer, you can make a difference." The reverse side bore a general message as well as one specific to the treatment applied. In all treatments, it was explained that the blood donation service found it difficult to meet demand during the summer, and that this might possibly lead to significant shortages. In the appeal treatment, the card then stated

> "In order to prevent this, we are particularly relying on your voluntary donation during the summer months. We therefore especially invite you with this call to

[9] A few donations are also collected from spontaneous donors. We omit those from the statistics as they are subsequently not part of the experiment.

donate blood.

Many thanks!

Zurich Blood Donation Service SRC"

In the second treatment, in addition to the information provided in the appeal treatment, a cholesterol test was offered. Specifically, the following sentence was added to the card:

> "In appreciation of your donation, this summer we offer you the opportunity to check your cholesterol level free of charge at the blood donation center."

In the third treatment, the subjects were offered a lottery ticket. The text on the card was supplemented as follows:

> "In appreciation of your donation, this summer you will receive a lottery ticket from the Swiss State Lottery."

The retail value of the cholesterol check was CHF 15 (approx. 12 Euros), the one of the lottery ticket CHF 5 (approx. 4 Euros). The lottery ticket was low-yield, but with a higher probability of winning (the probability of winning CHF 10 or more was approximately 30 percent).

It was a requirement of the centers' board that all donors showing up at a donation center be treated equally. For this reason, we had to implement the second treatments (the cholesterol test) and the third one (the lottery ticket) in different donation centers. As the board was most interested in the effectiveness of the cholesterol test as a retention reward, the respective reward was offered in three centers. The lottery ticket was offered in only one center (the second largest). The equal treatment implies that, upon showing up at the donation center, *all* donors were offered, depending on the center, a lottery ticket or a cholesterol test irrespective of whether this was advertised to them in the invitation. The sequence of the treatments was randomized over days; i.e., only one treatment was mailed out per day due to administrative considerations. Thus, the treatment was randomized only within weeks and within donation centers. This constraint has important implications for the randomization of our treatments. It requires us to control for donation center, week, and weekday in the empirical analysis below, and has implications for the correction of standard errors.

B. Descriptive Statistics

In total, 11,320 blood donors were mailed an invitation during the experiment, which constitute are base sample. Table 2 displays information on the basic demographics as well as the past donation frequency of the blood donors who were invited during the experiment. The first line of the table shows that the mean age in the sample is about 44 years and that approximately 40 percent of the donors are female.

The second line in Table 2 restricts attention to the blood donors who have received at least four previous invitations. This subsample will be used to study heterogeneous treatment effects between frequent and infrequent donors. As can be seen, there are only minor differences between the donors in the base sample and the ones in the subsample: the main difference is that donors in the latter sample are about one year older than the average. The table also shows the frequency with which individuals have donated in response to the last four invitations (or the last four or fewer invitations in the base sample). The average frequency of donation in response to four invitations is 1.97. The standard deviation of 1.55 shows that there is strong heterogeneity between donors in the probability with which they react to an invitation. Looking further into the heterogeneity (not reported in the table), we find that 27 percent have not responded to any of the four previous invitations. About 15 percent have responded once, twice or thrice, respectively, to the past four invitations, and nearly 25 percent have donated each of the four times they were invited to. The subset of donors with at least four previous invitations is our preferred sample, as it allows us to control for heterogeneity in baseline motivation in a parsimonious way.

Turning to the distribution of treatments, Table 3 gives an overview of the number of invitations sent out per treatment overall and by center. A free cholesterol test was offered to about one third of the invited people. About one tenth received an invitation offering a lottery ticket. As the latter reward was only offered in one of the four donation centers, the respective experimental group is smaller.

The bottom part of Table 2 provides a randomization check for the observable characteristics. The table shows that there are small differences across treatments with respect to age and previous donations, possibly due to the coarse randomization procedure. In order to examine this observation in more detail, we perform balance checks for the observable characteristics. As the treatment was randomized within centers and between days, we control for center-specific week effects in these randomization checks. Moreover, the opening hours of the centers differ by weekday. We therefore also control for center $\times$ weekday effects. The randomization check is implemented by estimating a set of seemingly unrelated regressions taking into account the correlation in the residuals across equations when calculating the covariance matrix. The p-values from these regressions are shown in Table 2. We detect that there are small and marginally statistically significant imbalances with respect to age and previous donations. However, the overall F-test of balance across all four equations, also taking into account cross-correlations in the four characteristics, shows no statistically significant evidence of imbalance. This holds, in particular, for the sample restricted to individuals with four previous invitations for donations.

III. Behavioral Predictions

This section examines the predictions of different theories regarding the response of blood donations to incentives. The predictions concern (i) how individuals respond to material rewards and other kind of incentives in general, (ii) how differential donor motivation moderates the response to incentives during the experiment, and (iii) how the predictions for behavior subsequent to the experiment differ. In the first subsection, we examine different versions of the standard model in economics and explain what their predictions are concerning the use of incentives to donate blood. We also consider theories of "motivational crowding" (Frey, 1997), inspired by research in psychology suggesting that the use of monetary incentives can backfire in a sense to be made precise. In the second subsection, we describe the empirical strategy.

A. Theory

The Baseline Model: We begin by discussing a stripped-down economic model of blood donations. To fix ideas, suppose that the utility from donating blood is given by

$$u_i(d_i) = (v_i - c)d_i \tag{1}$$

where $d_i = 1$ indicates that individual i donates blood and $d_i = 0$ indicates that she does not. The parameter v_i is the individual's benefit from donating blood as derived from altruism or warm glow (see, e.g., Andreoni, 1990, 2007; Wildman and Hollingsworth, 2009). v_i may differ between individuals reflecting heterogeneity in pro-social motivations. The parameter c is the cost of donating blood as reflected in the time costs and other possible utility costs associated with donating blood. The cost c is random with cumulative distribution function $F_c()$ with support $[0, \bar{c}]$ reflecting that not each day may be equally convenient to donate blood.

In the absence of incentives, the fraction of individuals donating blood is the fraction of individuals for which $c < v_i$. This is given for the fraction $q = \int_0^{\bar{v}} F_c(z)dG(z)$, where $G()$ is the distribution function of v_i. The model naturally predicts an increase in donations as positive incentives are used. Denote the utility from the incentive by m. Then, individuals will donate blood if $c < v_i + m$, thus the fraction of individuals donating blood will change by $\Delta q = \int_0^{\bar{v}} \left(F_c(z+m) - F_c(z)\right) dG(z) \geq 0$. The model, as such is silent on what constitutes an incentive. It could be anything that has consumption value for the individual. In our experiment, we offer two types of potential incentives: a lottery ticket and a cholesterol test. As such, it is not clear which of the two is expected to have a stronger impact on behavior, as the utility of the two rewards to the individuals would need to be known.

The model can also be used to derive conditions under which highly motivated donors (deriving a high benefit v from donating blood) respond systematically differently to incentives than donors who are less motivated. As the model makes clear, differences in response to the same incentive are related to the distribution of opportunity costs as well as of the benefits from donating blood. Consider the case of a negative correlation between costs and benefits. For example, there are two types of individuals with benefits v and $v' > v$ and the density of the cost function is higher at $f_c(v)$ than at $f_c(v')$, then the change in the probability to donate blood will be higher for individuals of type v, as $F_c(v+m) - F_c(v) = \int_0^m f_c(v+z)dz \geq \int_0^m f_c(v'+z)dz = F_c(v'+m) - F_c(v')$. The reverse holds for the case of a positive correlation between costs and benefits.

Importantly, the standard model also makes the prediction that incentives only affect behavior while they are in place. Removing them causes donations to return to their baseline level.

Intertemporal Utility Spillovers: This last property, however, need not hold in the presence of intertemporal spillovers on utility. For instance, suppose that the utility from donating in period t takes the form

$$u(d_t, d_{t-1}) = \begin{cases} d_t(v - c_t) & \text{if } d_{t-1} = 0 \\ d_t(\alpha v - c_t) & \text{if } d_{t-1} = 1 \end{cases} \tag{2}$$

ignoring possible differences in v across individuals, and assuming an infinitely-lived individual with discount factor $\delta < 1$ between periods. In contrast to equation (1), equation (2) illustrates a case in which past behavior has a direct effect on the utility of donating blood today: if the individual has donated in period $t-1$, the benefit from donating today is αv, while the benefit from donating is v if she has not donated in period $t-1$. Several specifications are possible. Consider first the case where $\alpha < 1$. This corresponds to the case of "guilt-driven" preferences, i.e. the intuition that individuals feel more strongly that they should donate the more time has elapsed since their last donation. This implies that it feels less urgent to donate in the current period if the individual has donated in the last period, with $\alpha < 1$ indicating the "discounted" motivation due to the previous donation.

Consider now the effect of incentives on behavior. As we show in the appendix, if $d_{t-1} = 0$, the individual will find it optimal to donate if

$$c \leq v + m + \delta \left[E(V^1) - E(V^0) \right]$$

where $E(V^1)$ and $E(V^0)$ are the discounted expected lifetime utility starting with $d_t = 1, 0$, respectively, in period $t+1$. Thus, just like before, the individual will be more likely to donate when an incentive is added to the blood donation than if not. However this donation is going to have an effect on the decision to donate in the next period. In period $t+1$, the individual has $d_t = 1$ and will donate if

$$c \leq \alpha v + \delta \left[E(V^1) - E(V^0) \right]$$

Thus, if the individual has donated after being exposed to incentives last period, she is *less* likely to donate this period since it now feels less urgent to donate blood. In other words, $\alpha < 1$ may lead to a behavior that looks like "intertemporal substitution" in blood donations, leading to a temporary depression in blood donations subsequent to the use of incentives. From a policy perspective, this reaction makes the use of incentives less attractive as it depresses future blood donations.

Now consider the case where $\alpha > 1$. In this case, having donated last period raises the benefit from donating this period. This can be thought of as habit formation in the sense of Stigler and Becker (1977). The temporary use of incentives has a positive effect in the medium run. As the incentives induce more donations, more individuals will have utility $\alpha v > v$ over the next period, thus being more likely to donate. In this case, therefore, the temporary use of incentives may have medium-term benefits.[10] Figure 3 provides a quantitative illustration of the likely magnitude of such effects, calibrated to the baseline donation rate in our sample in each panel. We use a simple example of our model, where for simplicity we choose the costs c to follow a uniform distribution. Panel A displays the results for $\alpha = 2$. The panel shows the

[10] In the context of exercising, Charness and Gneezy (2009) have found that a temporary incentive can have effects on behavior beyond the period during which the incentive is offered.

donation rate at its steady-state value in period $t-1$, and then increased by 10 percentage points in response to a temporary incentive. As we explained, the behavior gradually tends back to its steady-steate value from above. One period after the incentive has been removed, donations are still 2.5 percentage points above their steady-state value and then slowly converge back to this level. Summing over all five periods after the intervention in t, habit-forming preferences increase blood donations by 3.1 percentage points, rendering the intervention 30 percent more productive in the long-run. In contrast, Panel B of Figure 3 displays the calibrations for the case of guilt-driven preferences with $\alpha = 0.5$, i.e. a case in which donating in the next period feels only half as urgent than if the individual had not donated the period before. As the figure shows, there is a sharp drop by more than 4 percentage points below the steady state in the period after the incentives have been used. Donation rates then oscillate and tend towards the steady state value, leading to a long run loss after the use of the incentive of 3.1 percentage points, thus reducing the long-run efficiency of incentives by roughly 30 percent.

Crowding-Out of Intrinsic Motivation: Research in psychology shown that when external rewards are given for an activity an individual intrinsically enjoys, this may undermine her intrinsic motivation (Deci and Ryan, 1985; Lepper and Greene, 1978). In terms of our model, this implies that an individual's utility in the absence of incentives is

$$u_i(d_t, 0) = (\nu_i h(0) - c)d_t \tag{3}$$

If incentives are introduced, this changes the intrinsic valuation of the activity according to the function $h()$:

$$u_i(d_t, m_t) = (\nu_i h(m_t) + m_t - c)d_t \tag{4}$$

The use of incentives decreases the intrinsic benefit of donating blood, captured by a decreasing function $h(m)$.[11] As a consequence, the positive relative price effect might be reversed.[12] Overall, the model with intrinsic motivation makes no clear prediction as to how behavior responds when incentives are provided. The sum of intrinsic and extrinsic benefits only increases, and hence the probability to donate blood, if

$$h(0) - h(m_t) < m_t/\nu_i \quad (5)$$

which is not a priori clear.[13] The model also implies the assertion in the psychological reasoning that those with a particularly high intrinsic benefit from the activity experience the largest drop in motivation, as the drop in the motivation $h(m) - h(0)$ is scaled by ν_i. A higher ν_i thus implies a larger drop in intrinsic motivation. So the theory predicts that highly motivated individuals respond less to incentives than less motivated ones, or even negatively, as equation (5) is more difficult to satisfy for a high ν_i.

Motivation crowing theory makes the prediction that once incentives have been used, intrinsic motivation is reduced or destroyed and will not recover. Thus while individuals choose whether or not to donate according to (3) before period t, they will choose whether or not to donate after the use of incentives according to

$$u_i(d_s, 0) = (\nu_i h(m_s) - c)d_s \quad (6)$$

with the intrinsic motivation diminished for all future periods $s > t$. Abstracting from other forces, motivation crowing theory unambiguously predicts that donations decrease subsequent

[11] The evidence in several papers suggests that $h(m)$ may not be differentiable, but rather drops discontinuously as rewards are used (Deci et al., 1999; Gneezy and Rustichini, 2000)

[12] A similar exposition is offered in Frey and Oberholzer-Gee (1997).

[13] This simple model of motivation crowding also captures the essence of richer models based on image concerns. The self-signaling interpretation of the models by Bénabou and Tirole (2006) and Bénabou and Tirole (2011) emphasizes that an individual wishes to see him- or herself as someone who likes contributing to a pro-social activity and who is not mercenary (that is, as someone who has a low marginal utility of money). Thus, in choosing an action, an individual also takes account of what this action signals about his or her character. In this setting, it is possible for incentives to reduce the level of an activity. The reason being that a positive response to such incentives may communicate an individual's positive marginal utility of money and thereby convey a selfish signal to his or her future selves. There are other recent models in economics with a similar flavor whereby behavioral reactions are derived from *public* image motivation (Ariely et al., 2009; Ellingsen and Johannesson, 2008). We do not refer to these models as donations in donation centers can be considered largely private.

to the use of incentives. Importantly, the theory is also specific that only *contingent* incentives are expected to crowd out intrinsic motivation (Deci et al., 1999). This is particularly relevant for our setting as all individuals who donated received a reward. The emphasis on contingent incentives implies that those people for whom the reward was not announced are understood not to experience an impact on their intrinsic motivation. They thus continue to be the relevant control group.

Summary: The behavioral consequences of incentives are separately modeled here, taking intertemporal spillovers as well as the possibility of motivation crowding effects into account. The two forces might well affect behavior simultaneously. Due to the natural setting, the theories have to be assessed based on the observed net effects on donation behavior. Ideally, the theories captured in our simple models would make distinct predictions that allow us to discriminate between them when confronted with the empirical evidence. The differences turn out to be rather subtle when the predictions from both models are combined.

Six combinations can meaningfully be discussed between three cases of intertemporal utility spillovers, i.e. $\alpha < 1$ (guilt), $\alpha = 1$ (no intertemporal effects) and $\alpha > 1$ (habit formation), and motivation crowding being effective or not. For all of them, there are predictions for the behavioral consequences during the intervention and after the intervention. During the experiment, utility spillovers obviously do not matter and a positive incentive effect is expected when there is no motivation crowding while the effect turns out ambiguous if motivation crowding is relevant. After the intervention, predictions for the six cases are as follows: For $\alpha = 1$, donations are at the baseline level with no motivation crowding but below baseline with motivation crowding. If $\alpha < 1$, there is a negative intertemporal spillover effect. With no motivation crowding, the negative effect vanishes over time. With motivation crowding, the negative effect also gets smaller over time but remains negative. In the case of $\alpha > 1$, there is a positive intertemporal spillover effect that decays over time. Combined with motivation crowding, it is ambiguous whether immediately after the intervention the donation level is above or below baseline. However, any positive effect is decreasing and turning into a negative effect over time. An overview of the predictions is presented in Table 1.

In addition to the listed predictions for the overall effects before and after the intervention, there are the differential predictions for more or less motivated donors. During as well as after the provision of incentives, relatively worse outcomes are expected for the more motivated donors if motivation crowding plays a role. However, this same prediction is also consistent with the costs and benefits of donating blood being negatively correlated. A negative correlation might not be expected a priori though as pro-social behavior is often observed to be positively correlated with the level of education and thus with the opportunity costs of time. While this argument suggests a positive rather than a negative correlation, we cannot assess whether the respective relationship also holds in our context. The unobserved heterogeneity in costs and benefits from donating blood thus impedes the interpretation of differential consequences of groups of donors solely in terms of motivation crowding theory. Of course, we still learn about the behavioral reactions of different groups.

B. Empirical Strategy

For the outcome during the experiment, we model the probability of a donation as

$$\Pr(d_i = 1) = \Pr(T_{ct}'\gamma + x_i'\beta + \delta_c(t) + \omega_c(t) + \epsilon_i > 0) \tag{7}$$

where T_{ct} is a vector of binary variables, indicating which treatment is being mailed out in center c on date t. We naturally choose the condition referring to the mailings of the standard invitation letters as our control condition. The vector γ contains the associated coefficients, indicating how the different treatments affect the index function that determines the probability to donate blood. The vector x_i contains individual-level control variables with coefficients β. In our core specifications, x_i also contains our measure of past donation intensity, indicating how many times out of the four previous invitations the individual showed up to donate blood. For this measure to be comparable across individuals, an individual needs to have received at least four invitations prior to the experiment. We will thus restrict the sample to individuals who fulfill this criteria. We also estimate a set of center-specific weekday effects $\delta_c(t)$, since the different centers have different opening hours. Finally, we also estimate center-specific week fixed effects $\omega_c(t)$. We assume that ϵ_i follows a logistic distribution and estimate a logit model by maximum likelihood. For the most part, we calculate the marginal effects $\frac{\partial \Pr(d_i=1)}{\partial z_i}$ and report those rather than the coefficient estimates of the index function. As we explained in the previous subsection, the experiment is randomized at the day $\times$ center level. We therefore follow the recommendation to cluster the standard errors at the level of randomization (Abadie et al., 2017).

In order to examine how the response to the experiment depends on the individual's prior frequency of donation, we estimate equation (7) separately for two groups of individuals: those having donated zero, once or twice in response to the last 4 invitations, and those having donate three or four times in response to the last four invitations.

We analyze the effects of the treatments on donations after the experiment was concluded by estimating an equation of the form

$$d_i^k = T_{ct}'\gamma^k + x_i'\beta^k + \delta_c^k(t) + \omega_c^k(t) + \epsilon_i^k \tag{8}$$

where d_i^k indicates the number of blood donations in the k months after the experiment had been concluded. In contrast to the experimental period, we are not able to condition on who received an invitation for this later period, but simply count the number of times we see that the individual donated. We estimate equation (8) for five time horizons of $k = 3, 6, 9, 12$ and 15 months after the experiment. We estimate the equations by OLS and cluster, again, at the level of the randomization (date of the original invitation to the experiment at the particular center). As a robustness check, we repeated the analysis using poisson regressions, and reach the same conclusions. The results are available on request.

IV. Results

A. Response During the Experiment

Overall Treatment Effects: Figure 4 offers a first descriptive presentation of the outcomes in the different experimental conditions. It shows the differences in donation rates between an intervention treatment and the baseline treatment in which the standard invitation was sent out. Thereby the outcome variable is normalized by donation center times donor type mean values. The standard errors are displayed as vertical lines around the differences in donation rates. The figure reveals a moderate increase of 1.3 percentage points in the donation rate when the appeal alone is sent out to donors, with the standard error indicating that the difference to the baseline invitation is likely not statistically significant. Turning to the two rewards used in the experiment, the results show an increase of 2.3 percentage points in the donation rate when a free cholesterol test is offered to donors. For the lottery ticket, the descriptive evidence suggests that it raises donation rates quite substantially by approximately 4 percentage points, and that the effect is likely statistically significant.

As the experiment is only randomized within centers across days, the averages across periods may, however, overstate the precision with which the differences across treatments are measured. In order to provide a formal statistical test, we estimate equation (7) in three steps. In the first column of Table 4, we only include the controls ensuring randomization (but do not take into account the donor type). We add demographic controls in the second column, and finally constrain the sample to the individuals who have received at least four previous invitations in the third column. The first and the second estimation are thus for the full sample, while the third estimation is restricted to donors with a known donor history. As in the figure above, the reference condition is the standard invitation. Differences in the number of observations across estimations arise due to the dropping of observations if information for the control variables is missing or due to perfect prediction in the logit models.

The results in the table show that the statistical precision is increased once we condition on previous donations and the center-level controls required by the level of randomization. Our preferred specification is presented in the third column of Table 4, with all the controls in place. While the point estimates are similar across all specifications, the third column offers the best statistical power.

Turning to the key treatment of interest, we find that the lottery ticket significantly increases blood donations. The point estimate is comparable across the specifications, and statistically significant at the 5 percent level in column (3), our preferred specification. The effect size of a 5.6 percentage points is considerable given that the baseline donation rate is approximately 50 percent. However, the lottery ticket reward was combined with an appeal to donate. We can control for this by comparing the effect of the lottery ticket to the appeal treatment as both treatments contained the same message. As can be seen in the table, the appeal treatment is not statistically significant in any of the specifications. In column (3), its point estimate is near 1 percentage point and with the added precision of the controls, we are able to reject the hypothesis that the lottery ticket and the appeal have the same effect ($p = 0.05$, see the bottom panel of the table). The cholesterol test always has a small positive point estimate going up to 1.7 percentage points. However, it is not statistically significant, but also not statistically distinguishable from the effect of the lottery ticket.

Overall, the treatment that significantly increases blood donations is the lottery ticket. Thus far, this result is consistent with either model described above, the standard model in economics and motivational crowding.

Differential Treatment Effects by Donor Motivation: In order to discriminate between different theories, we explore in more detail their predictions. In particular, motivational crowding predicts that highly motivated donors are more negatively affected by the use of incentives. We use our measure of previous donations in response to the last four invitations as an indicator of donor motivation. We group donors into frequent and infrequent donors, depending on whether they responded to at least three of the previous four invitations. The treatment effects, based on the same normalizations as in Figure 2, are displayed for the two groups in Figure 5. The figure suggests a clear qualitative difference in how frequent and infrequent donors respond to the experiment. Frequent donors seem to be more or less unmoved by any of the treatments. If anything the appeal appears to reduce turnout among frequent donors. In contrast, the treatment effects displayed in panel A of the figure suggests more sensitivity to the experimental treatments by infrequent donors. They seem somewhat more likely to donate even when receiving the appeal alone. The point estimate of the treatment effect of the appeal and the cholesterol test are both around 3 percentage points. Interestingly, the lottery ticket evokes an even larger response and donation rates increase by around 6 percentage points. Given that the mean donation rate for the less motivated donors is only about 26 percent, this is a large increase in relative terms.

We provide a formal statistical test in Table 5. The test confirms the visual impression from the graph, and provides marginal effects together with the correct standard errors. Turning first to the infrequent donors, the results show a strong increase in the donation rate in response to the lottery ticket: with all the controls in place, we estimate that donation rates increase by approximately 8 percentage points for this first group of donors. The estimations also reveal that the cholesterol test has a comparatively small effect of around 2 percentage points on blood donations whereby the hypothesis of no effect cannot be statistically rejected (see also Goette et al., 2009). A statistical test also reveals that the lottery ticket works significantly better than the cholesterol test, and also somewhat better than an appeal alone, even though the appeal has a positive effect of nearly 4 percentage points on donations as well.

Compare this to the results we find for frequent donors in the third and fourth columns of Table 5. As the figure indicated, there is no significant response to either of the reward treatments, free cholesterol test or lottery ticket. All the point estimates are small, and, for the lottery ticket, negative, but not statistically significant. If anything, the appeal alone seems to marginally decrease the chance to donate blood. While we reject the hypothesis of no effect of the treatments on infrequent donors ($p \leq 0.01$ in both specifications), we cannot reject non-response to any treatment at conventional significance levels ($p \approx 0.08$ in the third and fourth column). Importantly, we can also test whether the responses to the treatments in the two subgroups of donors are identical by comparing the estimated coefficients. We clearly reject that they are the same ($p < 0.01$ in either case). Thus, innate donor motivation, measured by previous donation rates, modulates the response to the experiment.

Treatment Effects for Alternative Sample Splits: In order to interpret the latter results as donor motivation moderating the response to the experiment, we need to rule out that other factors that may be correlated with the frequency of donation explain the heterogeneity in the response to the experiment. Donation frequency may be correlated with age, gender and whether or not the individual has signed up for scheduled invitations. We address concerns that each of these factors could, in fact, moderate the response to the experiment rather than the intensity of past donations.

We approach the issue empirically by estimating equation (7) with all the control variables included for three alternative splits of the sample. Table 6 presents the results. We begin by examining whether men and women respond differently to the rewards. Since women react somewhat less often to invitations, the difference being around 3 percentage points according to Table 4, perhaps it is that women respond more strongly to rewards. The first two columns of Table 6 display the result that no such general tendency is observed.[14] While the point estimates for the treatment effects in the women sample seem larger for the appeal and the cholesterol test, there is almost no difference for the lottery ticket. However, the large standard errors of the treatment effects prevent us from drawing strong conclusions. A formal test of the equality of the treatment coefficients for men and women cannot be rejected. Next, we examine whether age moderates the differences in the response to the treatment. Age is positively correlated with the response behavior to invitations, and it is possible that young donors show a stronger response to the appeal and the rewards than elder donors, as marginal costs might hinder many to donate. The third and fourth column of Table 6 display the results for the two age groups obtained by a split at the median age of 45. It shows that the point estimates of the treatment effects are indeed larger for the young than for the old donors. However, a statistical test again cannot reject the null of no difference in treatment effects between the two groups. Finally, we split the sample by whether or not an individual has signed up to receive invitations by some periodicity

[14] Mellström and Johannesson (2008) find gender differences in the response to incentives for blood donations in a setting in which image motivation is relevant. In other research on motivational crowding, no systematic gender differences are observed (Deci et al., 1999).

or not. Individuals who have signed up for regular invitations tend to donate somewhat more often, though not much (as they can, for instance, also sign up to receive only one invitation per year). Perhaps the heterogeneity in the response comes from one set of donors being rather rigid about blood donations (those who sign up for scheduled invitations) and another set being more flexible. The fifth and the sixth column of Table 6 present the results. Again, there is no statistically discernible difference in the response to the four treatments, which is indicated by the high p-value of a formal test of the null of no treatment differences across the two groups. To summarize, none of the alternative explanations for why previous response intensity moderates the reaction to the experiment can be supported with the data.

Theoretical Interpretation of the Findings: Having set these alternative explanation aside, we conclude that it is individual differences in the motivation to donate blood that create the differential response to the experiment. However, which model do the results from this subsection support? Motivational crowding predicts that individuals with high intrinsic motivation experience the strongest drop in this motivation if extrinsic incentives are used. On the one hand, this is consistent with the result that infrequent donors respond strongly and positively to the extrinsic incentive (the lottery ticket), while there is a net zero effect on the highly motivated donors. Thereby the latter net effect can be understood as the result of some of their intrinsic motivation being destroyed and simultaneously being countered by the incentive effect of the lottery ticket. On the other hand, the results are also consistent with a standard model in economics and a situation in which the distribution of marginal costs has more density at the motivation levels of the marginal infrequent than the marginal frequent donors.

The models only make distinctive predictions for donations after the incentives are removed. Motivational crowding predicts that motivation is destroyed long-term and implies lower donation rates after the use of incentives. The standard model may also predict post-intervention effects if preferences for blood donation are habit-forming, or guilt-driven, as we explained in the previous section. However, these responses (a decreasing positive response in the case of habit formation, and a decreasing negative effect in the case of guilt) are qualitatively in sharp contrast to the predictions of the motivational crowding model which imply negative treatment effects in the longer term independently of the intertemporal utility spillovers. We turn to an empirical test of these implications in the next subsection.

B. Response After the Experiment

Average Effects: In order to examine the response to the treatments after the experiment, we estimate linear regressions for the number of donations three, six, nine, twelve and fifteen months after the experiment, as specified in equation (8). The estimated coefficients are presented in Table 7. In order to gain a more intuitive impression of the cumulative changes in the number of donations for a given time window, we have also plotted the estimates in Figure 6 for each of the treatments separately. The first observation in each of the panels is the estimated change in the probability to donate, for comparability also estimated using OLS, based on the specification in the third column in Table 4. Around the point estimates, we display the 50, 75 and 95 percent confidence intervals of the estimates.

Our main interest is to examine whether the use of incentives decreased blood donations subsequent to the experiment, and whether this effect is temporary (due to intertemporal utility spillovers) or permanent (as in the motivational-crowding model). Looking at the overall results, there is no evidence that blood donations decrease subsequent to the offering of rewards. In particular, in the case of the lottery ticket (panel A in Figure 6), the point estimates are positive. The 95% confidence interval shows that we can reject even small decreases in donations following the first-time use of incentives. If anything, the graph is mildly suggestive of habit-forming preferences as modeled and illustrated in Figure 3, with a positive effect slowly vanishing. However, besides from being able to reject small negative effects, we do not have enough precision to take a strong stance on whether preferences for blood donations are habit forming. Panels B and C in Figure 6 display the changes in donations subsequent to the intervention involving a free cholesterol test or an appeal. There are no statistically significant effects observed, with the null effect within the 95-percent confidence interval, for both of them. Taken together, the overall evaluation of the blood donations up to 15 months after the experiment shows no statistically significant effects, but with positive point estimates for the lottery ticket and rather small negative ones for the cholesterol test. The evidence thus does not support the worrisome predictions of the motivation-crowding model for blood donations.

Differential Treatment Effects by Donor Motivation: As before, we split the sample by the motivation to donate blood as measured based on behavior prior to the experiment, and estimate equation (8) for the two groups of donors. The results are reported in Tables 8 and 9 and displayed in Figures 7, 8 and 9. Formal tests comparing the two sets of estimates fail to reject the null that they are identical for any of the time windows. Thus, we find no statistically significant differences in the behavior of the individuals after the intervention from the different treatments, in contrast to what we find during the experiment.

Theoretical Interpretation of the Findings: Overall, the evidence favors the standard model in economics. While we find that previous donor motivation moderates the response to the experiment, as predicted by the motivational-crowding model, this result is also consistent with a standard model of preference heterogeneity. Furthermore, the behavioral pattern in donations subsequent to the experiment offers no systematic evidence for the predictions of the crowding model. Our best interpretation is that the qualitative features resemble most those of a model with habit formation.

V. Concluding Remarks

Whether material rewards are an effective mean to increase pro-social behavior is highly controversial. Our study offers evidence from a field experiment in the context of blood donation, often considered prototypical for an intrinsically motivated volunteer behavior. Before we put our study in perspective and briefly discuss its implications and limitations in the specific domain of blood donations, we try to precisely summarize our analysis in relation to the general academic debate.

A. Summary of the Experiment

Do private incentives in the form of material rewards offer an (extrinsic) motivation to donate blood or do they crowd out intrinsic motivation to do so, as predicted by theories of intrinsic motivation (Deci and Ryan, 1985; Frey, 1997; Gneezy et al., 2011). To address this question, in the key experimental condition of our study, individuals were offered a lottery ticket if they donate during the summer months (often plagued by a shortage of blood). Compared to individuals in the control condition who only received the standard invitation to donate (without any mention of the shortage), we find that the lottery ticket significantly increased blood donations. Individuals in the lottery ticket condition have a 5.6 percentage points higher probability of donating over a baseline probability of 46 percent in the control condition.

The announcement of the lottery was combined with the message that blood supply was expected to run low. For this reason, an additional control for the effect of the pure information about a possible shortage was implemented (similar to Sun et al., 2016). In this condition, the standard invitation was combined with a message about the special need during summer presented in the exact same way as in the main reward treatment. As offering a tangible material reward might also be seen as a form of costly signaling, we control for this interpretation in another condition in which individuals were offered a free (non-transferable) cholesterol test. We find that neither an appeal nor the offer of a cholesterol test raises the propensity to donate.[15]

While our main findings are consistent with the standard model in economics, they are also consistent with motivational crowding (with the relative price effect mitigating the impact from motivational crowing out). However, our setup allowed us to probe deeper into the underlying behavioral mechanisms for three reasons. First, donors had not been offered rewards before in return for their blood donation. Thus, using incentives on this population has the largest possible effect on intrinsic motivation. Second, the rewards were announced to the donors

[15]The result that specific material rewards are ineffective in raising donation rates is not uncommon. For instance, the offer of a free t-shirt was ineffective in Reich et al. (2006). However, the ineffectiveness of a reward in the heath domain, i.e. in the same domain as the pro-social activity, might come as a surprise. Indeed, in related research, offering a comprehensive blood check turned out effective in motivating blood donations (Leipnitz et al., 2018).

privately, and blood donations took place in the relative anonymity of a medical center. This makes it unlikely that the use of incentives creates an image concern as in Bénabou and Tirole (2006) and Ariely et al. (2009). Third, in our experiment, all donors were treated the same when at the donation center. That is, even donors who received the control (or appeal) invitation received a lottery ticket. However, to them, the lottery ticket was an unexpected gift rather than an incentive. This allows us to examine how the *contingency* of the lottery ticket affects subsequent motivation, exactly as in the experiments in social psychology (Deci et al., 1999).

Theories of intrinsic motivation make two clear predictions with respect to crowding out: first, they predict that highly motivated donors experience stronger crowding out than donors with lower baseline motivation. Second, they predict that intrinsic motivation will remain lower subsequent to the use of incentives. Our data allow tests o both predictions. We use the responsiveness to invitations prior to the experiment as our measure of intrinsic motivation. Individuals who rarely responded (low baseline motivation) reacted very strongly to the treatments. Offering a lottery ticket increased their probability to donate by 8 percentage points over a baseline probability of 30 percent. There is also a significantly positive, but smaller, response to the appeal condition, whereas the cholesterol test is ineffective. In both cases, the lottery ticket condition has a significantly stronger effect on the probability to donate than the appeal and cholesterol test condition.

In contrast, individuals with a high baseline motivation did not respond to either the lottery ticket or the cholesterol test as a reward. If anything, they respond negatively to being sent an appeal, compared to the control condition of a standard invitation.

We also followed the donation frequency of our experimental population up to 15 months after the experiment. For this period, the SRC returned to their previous regime without rewards for blood donations. The results do not suggest any significant pattern of intertemporal substitution that would follow from a standard economic model, and offers only weak and statistically insignificant hints of habit formation. Moreover, at no point in time did we find any significant evidence of crowding out of intrinsic motivation through the specific material rewards. The precision of our estimates is sufficient to reject even small negative effects. Thus, even though our setup allows us to follow almost the identical identification strategy as the lab experiments (no experience with incentives for blood donations, all individuals receive reward after donating), we do not find a pattern indicating a lower motivation once incentives are removed.

B. Implications and Limitations

Overall, our results suggest that selective incentives and pro-social motivations may coexist even in domains that heavily rely on people's intrinsic motivation *and* in which usually no rewards are offered for voluntary contributions. Thus, in light of the recurring seasonal shortages and a steady tightening of donor criteria, material rewards may prove useful to motivate previous donors to donate more blood. This finding complements previous evidence that material rewards are effective in increasing donations in the short-term in a context where rewards have become common (Lacetera et al., 2012, 2014). Our results should not, however, be construed as evidence that instituting a permanent regime of monetary incentives would have positive effects on the level of donations. In our experiment, no money was offered and there was essentially no public image concern, as in Bénabou and Tirole (2006), so an important channel by which incentives may become ineffective for the specific population of donors was shut out. Furthermore, a switch to a regime offering rewards on a permanent basis may also be interpreted by donors as evidence that the donation service is not altruistic, and may trigger effects along the line described in Ellingsen and Johannesson (2008).

However, while our evidence suggests that incentives can be used as a stop-gap measure when shortages occur, further research is needed to address how permanently switching to reward schemes affects the pro-social motivation of blood donors. Neither our results nor others (e.g. Lacetera et al., 2014) should be taken as evidence that a policy shift from an all-volunteer system would be desirable. Such a shift may well induce changes in the composition of the individuals willing to donate. This may help to develop a common framework to also understand problems of recruitment (see, e.g., Iajya et al., 2013; Stutzer et al., 2011) and retention of first time donors (see, e.g., Bagot et al., 2016) as well as the retention of long-term donors (see, e.g., Chell et al., 2018).

Furthermore, the fact that incentives are not harmful doesn't imply that they are the most effective policy choice. Other mechanisms to stimulate blood donations in situations of shortages may be more cost effective. For instance, Bruhin et al. (2015) show that a phone call, aimed at overcoming shortages in particular blood types, is a highly effective tool to increase turnout. They show that the phone call increases turnout by 10 percentage points uniformly in the population of donors. The comparison is particularly interesting, as the recipients of the phone call were individuals registered with the same blood donation service as in our study. The much larger overall effect compared to our lottery ticket is striking, and suggests that the phone call taps into different motivations. Further research is needed into which types of policy interventions are relatively more effective in the context of blood donations.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge, "When Should You Adjust Standard Errors for Clustering?" NBER Working Paper w24003, National Bureau of Economic Research 2017.

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *Economic Journal*, 1990, *100* (401), 464–477.

_ , "Philanthropy," in Serge-Christophe Kolm and Jean Mercier Ythier, eds., *Handbook of the Economics of Giving, Reciprocity and Altruism*, Vol. 2, Holland: North-Holland, 2007.

Archard, David, "Selling Yourself: Titmuss's Argument Against a Market in Blood," *The Journal of Ethics*, 2002, *6* (1), 87–102.

Ariely, Dan, Anat Bracha, and Stephan Meier, "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially" *American Economic Review*, 2009, *99* (1), 544–555.

Bagot, Kathleen L., Andrea L. Murray, and Barbara M. Masser, "How Can We Improve Retention of the First-Time Donor? A Systematic Review of the Current Evidence" *Transfusion Medicine Reviews*, 2016, *30* (2), 81–91.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, 2006, *96* (5), 1652–1678.

_ **and** _ , "Identity, Morals, and Taboos: Beliefs as Assets," *The Quarterly Journal of Economics*, 2011, *126* (2), 805–855.

Bruhin, Adrian, Lorenz Goette, Adrian Roethlisberger, Alexander Markovic, Regula Buchli, and Beat M. Frey, "Call of Duty: the Effects of Phone Calls on Blood Donor Motivation," *Transfusion*, 2015, *55* (11), 2645–2652.

Charness, Gary and Uri Gneezy, "Incentives to Exercise," *Econometrica*, 2009, *77* (3), 909–931.

Chell, Kathleen, Tanya E Davison, Barbara Masser, and Kyle Jensen, "A Systematic Review of Incentives in Blood Donation," *Transfusion*, 2018, *58* (1), 242–254.

Davey, Richard J., "Recruiting Blood Donors: Challenges and Opportunities," *Transfusion*, 2004, *44* (4), 597–600.

Deci, Edward L. and Richard M. Ryan, *Intrinsic Motivation and Self-Determination in Human Behavior*, New York: Plenum Press, 1985.

_ **, Richard Koestner, and Richard M. Ryan**, "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation," *Psychological bulletin*, 1999, *125* (6), 627.

Eastlund, Ted, "Monetary Blood Donation Incentives and the Risk of Transfusion-Transmitted Infection," *Transfusion*, 1998, *38* (9), 874–882.

Ellingsen, Tore and Magnus Johannesson, "Pride and Prejudice: The Human Side of Incentive Theory," *American Economic Review*, 2008, *98* (3), 990–1008.

Frey, Bruno S., *Not Just for the Money: An Economic Theory of Personal Motivation*, Cheltenham: Edward Elgar Publishing, 1997.

_ **and Felix Oberholzer-Gee**, "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out," *American Economic Review*, 1997, *87* (4), 746–755.

_ **and Reto Jegen**, "Motivation Crowding Theory: A Survey of the Empirical Evidence," *Journal of Economic Surveys*, 2001, *15* (5), 589–611.

Gilcher, Ronald O. and S. McCombs, "Seasonal Blood Shortages Can Be Eliminated," *Current Opinion in Hematology*, 2005, *12* (6), 503–508.

Gneezy, Uri and Aldo Rustichini, "A Fine is a Price," *Journal of Legal Studies*, January 2000, *29*, 1–18.

_ , Stephan Meier, and Pedro Rey-Biel, "When and Why Incentives (Don't) Work to Modify Behavior," *Journal of Economic Perspectives*, 2011, *25* (4), 191–210.

Goette, Lorenz, Alois Stutzer, and Beat M. Frey, "Prosocial Motivation and Blood Donations: A Survey of the Empirical Literature," *Transfusion Medicine and Hemotherapy*, 2010, *37* (3), 149–154.

_ , _ , Gürcan Yavuzcan, and Beat M. Frey, "Free Cholesterol Testing as a Motivation Device in Blood Donations: Evidence from Field Experiments," *Transfusion*, 2009, *49* (3), 524–531.

Higgins, Christopher, "Blood Transfusion: Risks and Benefits," *British Journal of Nursing*, 1994, *3* (19), 986–991.

Iajya, Victor, Nicola Lacetera, Mario Macis, and Robert Slonim, "The Effects of Information, Social and Financial Incentives on Voluntary Undirected Blood Donations: Evidence from a Field Experiment in Argentina," *Social Science & Medicine*, 2013, *98*, 214–223.

Kamenica, Emir, "Behavioral Economics and Psychology of Incentives," *Annual Review of Economics*, 2012, *4* (1), 427–452.

Lacetera, Nicola and Mario Macis, "Social Image Concerns and Prosocial Behavior: Field Evidence from a Nonlinear Incentive Scheme," *Journal of Economic Behavior & Organization*, 2010, *76* (2), 225–237.

_ , _ , and Robert Slonim, "Will There Be Blood? Incentives and Displacement Effects in Pro-Social Behavior" *American Economic Journal: Economic Policy*, 2012, *4* (1), 186–223.

_ , _ , and _ , "Economic Rewards to Motivate Blood Donations," *Science*, 2013, *340* (6135), 927–928.

_ , _ , and _ , "Rewarding Volunteers: a Field Experiment," *Management Science*, 2014, *60* (5), 1107–1129.

Leipnitz, Sigrun, Martha de Vries, Michel Clement, and Nina Mazar, "Providing Health Checks as Incentives to Retain Blood DonorsEvidence from Two Field Experiments," *International Journal of Research in Marketing*, 2018, *35* (4), 628–640.

Lepper, Mark R. and David Greene, *The Hidden Costs of Reward: New Perspectives on the Psychology of Human Motivation*, Hillsdale, N.J.: Erlbaum, 1978.

Levitt, Steven and John A. List, "What do Laboratory Experiments Measuring Social Preferences Tell Us about the Real World?" *Journal of Economic Perspectives*, 2007, *21* (2), 153–174.

Mellström, Carl and Magnus Johannesson, "Crowding Out in Blood Donation: Was Titmuss Right?" *Journal of the European Economic Association*, 2008, *6* (4), 845–863.

Montoya, A. Fernández, J. de Dios Luna del Castillo, A. López Berrio, and A. Rodríguez Fernández, "Attitudes, Beliefs, and Motivations in Blood Donors and Non-Donors," *Sangre*, 1996, *41* (6), 427–440.

Niza, Claudia, Burcu Tung, and Theresa M Marteau, "Incentivizing Blood Donation: Systematic Review and Meta-Analysis to test Titmuss' Hypotheses," *Health Psychology*, 2013, *32* (9), 941.

Piliavin, Jane Allyn, "Why Do They Give the Gift of Life - a Review of Research on Blood-Donors Since 1977," *Transfusion*, 1990, *30* (5), 444–459.

Reich, Pascale, Paula Roberts, Nancy Laabs, Artina Chinn, Patrick McEvoy, Nora Hirschler, and Edward L. Murphy, "A Randomized Trial of Blood Donor Recruitment Strategies," *Transfusion*, 2006, *46* (7), 1090–1096.

Sanchez, Ana M., Dannie I. Ameti, George B. Schreiber, Ruth A. Thomson, Annie Lo, James Bethel, and Alan E. Williams, "The Potential Impact of Incentives on Future Blood Donation Behavior," *Transfusion*, 2001, *41* (2), 172–178.

Slonim, Robert, Carmen Wang, and Ellen Garbarino, "The Market for Blood," *Journal of Economic Perspectives*, 2014, *28* (2), 177–96.

Stigler, George J. and Gary S. Becker, "De Gustibus Non Est Disputandum," *American Economic Review*, 1977, *67* (2), 76–90.

Stutzer, Alois and Lorenz Goette, "Blood Donor Motivation: What is Ethical? What Works?" *ISBT Science Series*, 2010, *5* (n1), 244–248.

_ , _ , and Michael Zehnder, "Active Decisions and Prosocial Behaviour: A Field Experiment on Blood Donation," *Economic Journal*, 2011, *121* (556), F476–F493.

Sun, Tianshu, Susan Feng Lu, and Ginger Zhe Jin, "Solving Shortage in a Priceless Market: Insights from Blood Donation," *Journal of Health Economics*, 2016, *48*, 149–165.

Titmuss, Richard Morris, *The Gift Relationship: From Human Blood to Social Policy*, New York: Random House, 1971.

Trimmel, Michael, Helene Lattacher, and Monika Janda, "Voluntary Whole-Blood Donors, and Compensated Platelet Donors and Plasma Donors: Motivation to Donate, Altruism and Aggression," *Transfusion and Apheresis Science*, 2005, *33* (2), 147–155.

Van der Poel, C. L., E. Seifried, and W. P. Schaasberg, "Paying for Blood Donations: Still a Risk?" *Vox Sanguinis*, 2002, *83* (4), 285–293.

Wildman, John and Bruce Hollingsworth, "Blood Donation and the Nature of Altruism," *Journal of Health Economics*, 2009, *28* (2), 492–503.

World Health Organization, "Blood Safety and Availability," https://www.who.int/newsroom/fact-sheets/detail/blood-safety-and-availability, Geneva 2017.

A An Illustrative Model of Blood Donations

Suppose that an individual's period utility is given by

$$u(d_t, d_{t-1}) = \begin{cases} d_t(v - c_t) & \text{if } d_{t-1} = 0 \\ d_t(\alpha v - c_t) & \text{if } d_{t-1} = 1 \end{cases}$$

where $d_t \in \{0, 1\}$ indicates whether the individual donates blood ($d_t = 1$) or not. As before, v is the utility from donating blood, and c_t is the cost of donating blood in this period with distribution function $F_c(x)$. The parameter $\alpha \neq 1$ captures the way in which a donation in the last period affects the utility from donating blood in this period. If $\alpha > 1$, this can be interpreted as habit formation: donating is more pleasant if one does it regularly. The model also allows for the possibility that donating is less pleasant or feels less urgent if one has done it recently ($\alpha < 1$), for example because people feel guilt in reaction to the time that has elapsed since their last donation.

Consider now an individual who did not donate in $t - 1$. Her lifetime utility is given by

$$V_t^0 = \delta E(V^0) + \max_{d_t} \left((v - c_t) + \delta(E(V^1) - E(V^0))d_t\right) \tag{9}$$

where we impose stationarity right from the beginning. Similarly, an individual who donated in $t - 1$, lifetime utility is

$$V_t^1 = \delta E(V^0) + \max_{d_t} \left((\alpha v - c_t) + \delta(E(V^1) - E(V^0))d_t\right) \tag{10}$$

Applying the expectation operation to (9), we obtain

$$E(V^0) = \delta E(V^0) + p^0 \left((v - c^0) + \delta(E(V^1) - E(V^0))\right) \tag{11}$$

where $p^0 = \Pr(c_t \leq v + \delta(E(V^1) - E(V^0))$ and $c^0 = E(c_t | c_t \leq v + \delta(E(V^1) - E(V^0))$. To keep this example simple, assume that c_t is uniform on $[0, 1]$. In this case, $c^0 = 0.5(v + \delta\Delta V)$ and

$p^0 = 2c^0 = (v + \delta\Delta V)$, where $\Delta V \equiv E(V^1) - E(V^0)$. Thus,

$$E(V^0) = \delta E(V^0) + 0.5(v + \delta\Delta V)^2 \tag{12}$$

Similarly, we take the expectation of (10) to obtain

$$E(V^1) = \delta E(V^0) + p^1 \left((\alpha v - c^1) + \delta(E(V^1) - E(V^0))\right)$$

where $c^1 = \alpha v + \delta\Delta V$ and $p^1 = 2c^1$. Thus,

$$E(V^1) = \delta E(V^0) + 0.5(\alpha v + \delta\Delta V)^2 \tag{13}$$

Subtracting (12) from (13), we obtain

$$\Delta V = 0.5v^2(\alpha^2 - 1) + (\alpha - 1)\delta v\Delta V \tag{14}$$

$$\Leftrightarrow \Delta V = 0.5v^2 \frac{\alpha^2 - 1}{1 - \delta v(\alpha - 1)}$$

Notice that as $1 > \alpha > 0$, $\Delta V < 0$ since the reduction in the utility α reduces the utility from donating blood this period. As $\alpha > 1$, $\Delta V > 0$, reflecting the higher utility if the individual starts from a situation of having donated in the period before.

The preceding calculations are based on the case that the probability of donating blood is strictly less than one, a necessary condition to this is $\delta(\alpha - 1) < 1$. However, if α is large enough, this condition may not be satisfied as the individual will donate blood in any case. Formally, this can be seen by plugging the solution for ΔV into the definition of p^0

$$p^0 = v + \delta\Delta V = v + 0.5v^2 \frac{\alpha^2 - 1}{1 - \delta v(\alpha - 1)} = v \frac{1 - \delta v(1 - \alpha) + 0.5\delta(\alpha^2 - 1)v}{1 - \delta v(\alpha - 1)}$$

$$= v \frac{1 - \delta v(\alpha - 1)(1 - 0.5(\alpha + 1))}{1 - \delta v(\alpha - 1)} = v \frac{1 + 0.5\delta v(\alpha - 1)^2}{1 - \delta v(\alpha - 1)}$$

This probability is less than one if $v\left(1+0.5\delta v(\alpha-1)^2\right) < 1-\delta v(\alpha-1)$, a necessary condition to which is $1-\delta v(\alpha-1) > 0 \Leftrightarrow \delta v(\alpha-1) < 1$.

We use this model to calibrate a numerical simulation of the use of incentives, displayed in Figure 3. We calculate Δ, p^0, p^1 in the absence of incentives using parameter values $\delta = 0.99$ and values of 2 and 0.5 for α, respectively. The steady state donation rate q in this model is given by

$$q = qp^1 + (1-q)p^0 \tag{15}$$

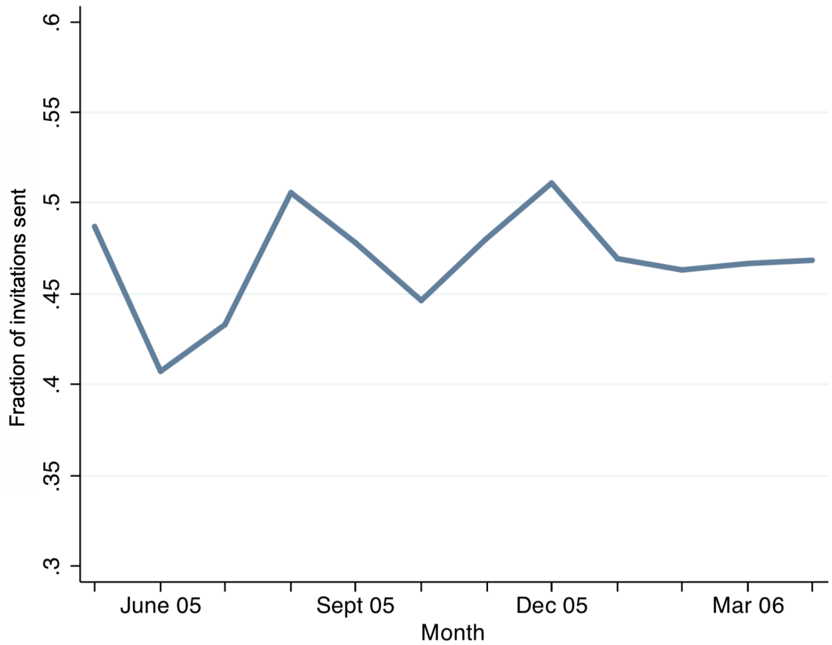
where the right-hand side of equation (15) is the composition of the population in the last period, and the left-hand side is the fraction of individuals who will be donating this period. Solving for q yields

$$q = \frac{p^0}{1-p^1+p^0}$$

For each of the two cases $\alpha = 0.5, 2$, we choose v such that q matches the average donation rate in our sample (0.477). We then shock individuals in one period with a one-time incentive of $m = 0.1$ in utility terms, equivalent to a 10 percentage point increase in the donation rate. We subsequently calculate how the donation choices tend back to the steady state. While the model could be solved analytically, we approximated the solution by simulating it numerically using $N = 1,000,000$ individuals.

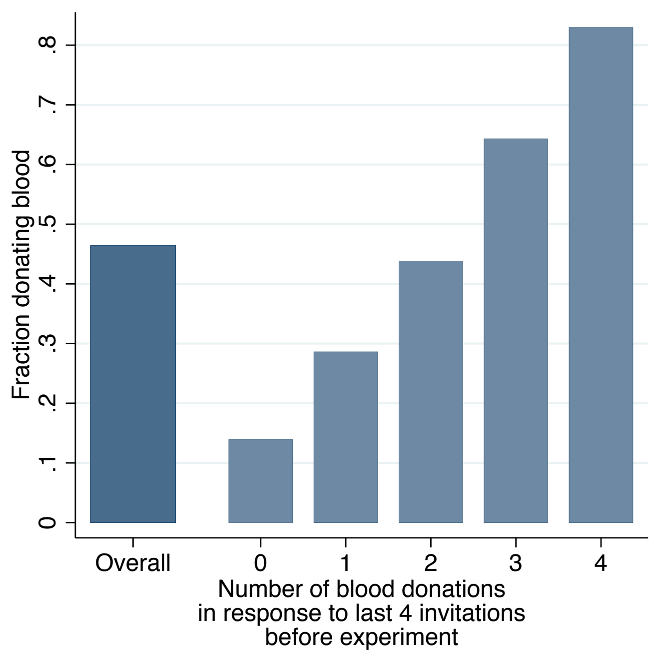
B Figures

Figure 1: Propensity to Donate Blood Following an Invitation Prior to the Field Experiment



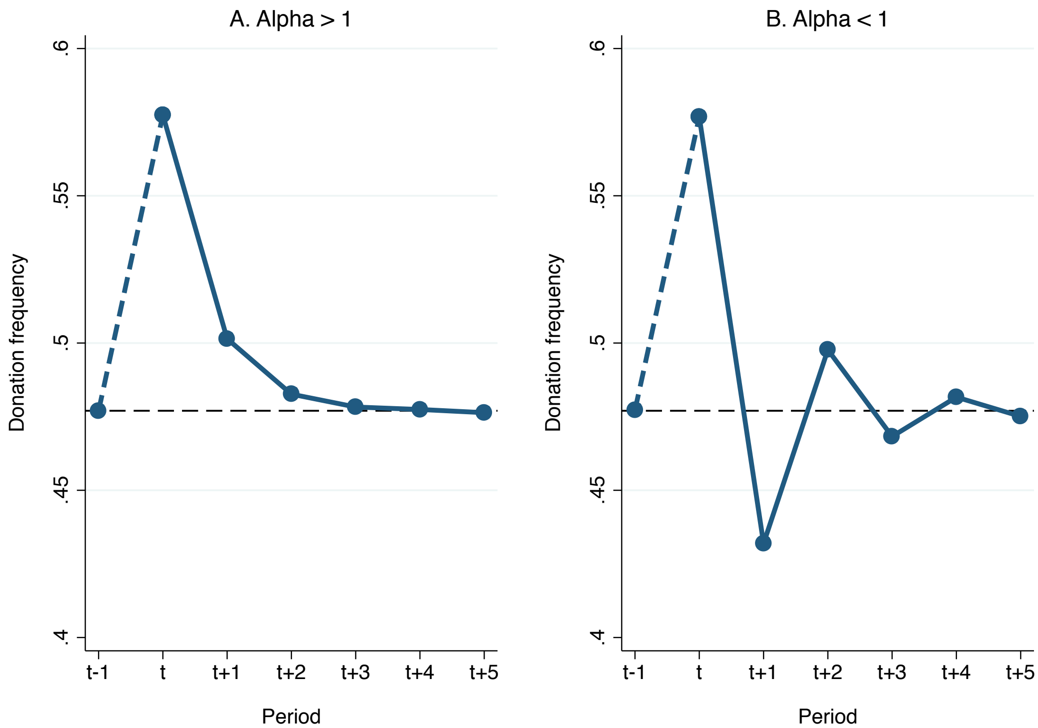
Source: Own calculations based on data from the SRC.

Figure 2: Propensity to Donate Blood During the Field Experiment



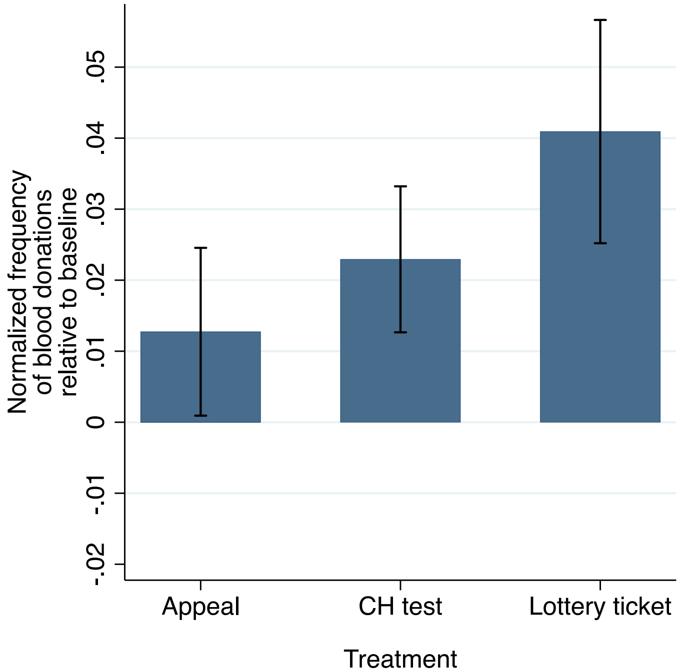
Note: The sample is restricted to the 9,731 donors who have received at least four previous invitations to donate blood.

Figure 3: Simulated Donation Patterns with Intertemporal Spillovers in Utility



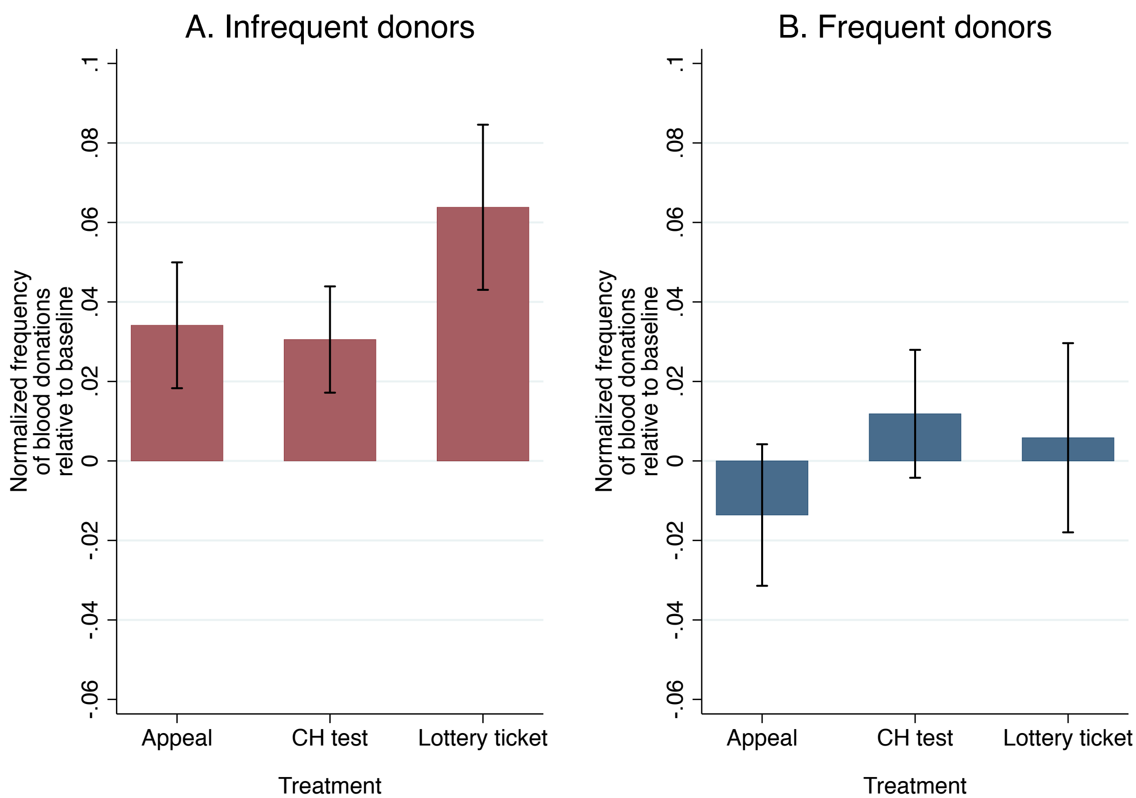
Notes: Calculations are based on the model details in the appendix. Panel A displays the response to incentives with positive habit formation. In this example, $\alpha = 2$. Panel B displays the case of diminishing utility after a donation, and $\alpha = 0.5$ in this example. In both cases, the benefit of donating blood v is calibrated such that it corresponds to the baseline donation rate of 47 percent.The incentive $m = 0.1$ in period t increases donation rates by 10 percentage points. $\delta = 0.99$.

Figure 4: Descriptive Evidence of the Overall Treatment Effects



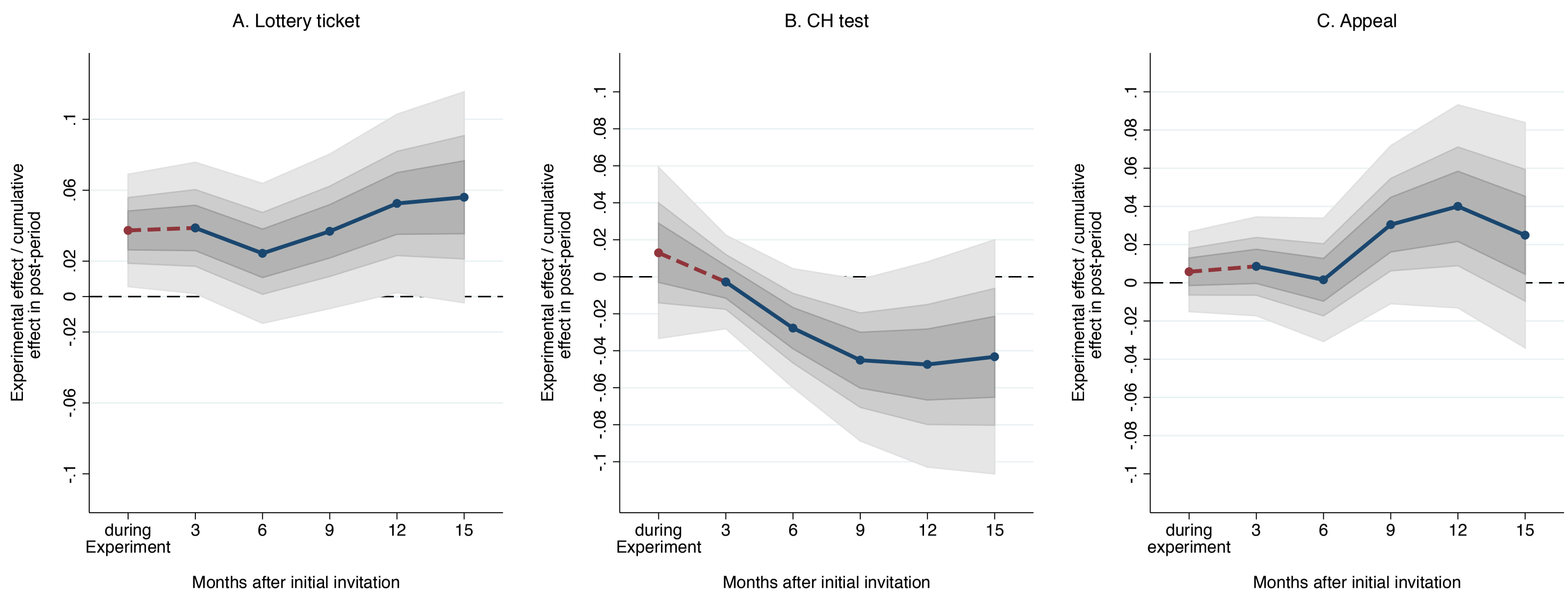
Notes: The outcome variable is normalized by donation center times donor type mean values. Treatment effects are calculated relative to the outcome with standard invitations. The vertical lines indicate the standard errors of the differences. The sample is restricted to the 9,731 donors who have received at least four previous invitations to donate blood. "Appeal" indicates that a special card was added to the invitation, calling subjects up to donate. "CH test" means that together with the card, subjects were offered a free cholesterol (CH) test if they showed up."Lottery ticket" refers to cards on which a lottery ticket was offered to donors.

Figure 5: Treatment Effects by Frequency of Prior Donations



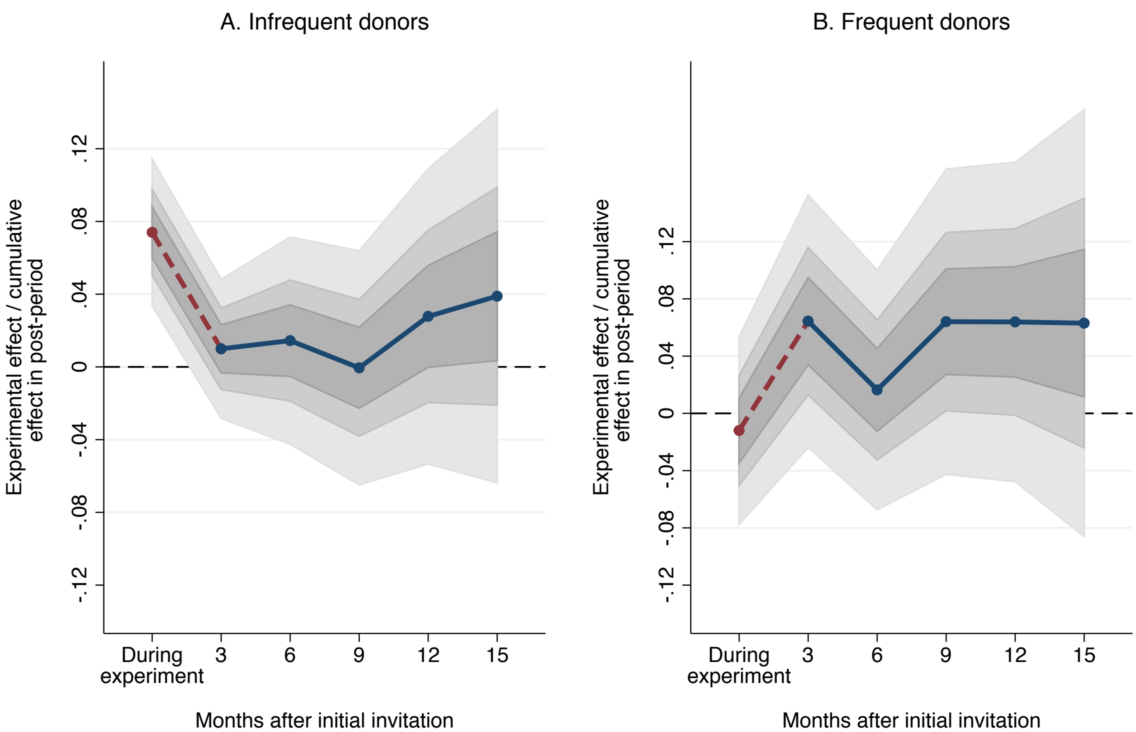
Note: See Figure 4.

Figure 6: Long-run Effects of the Experimental Interventions



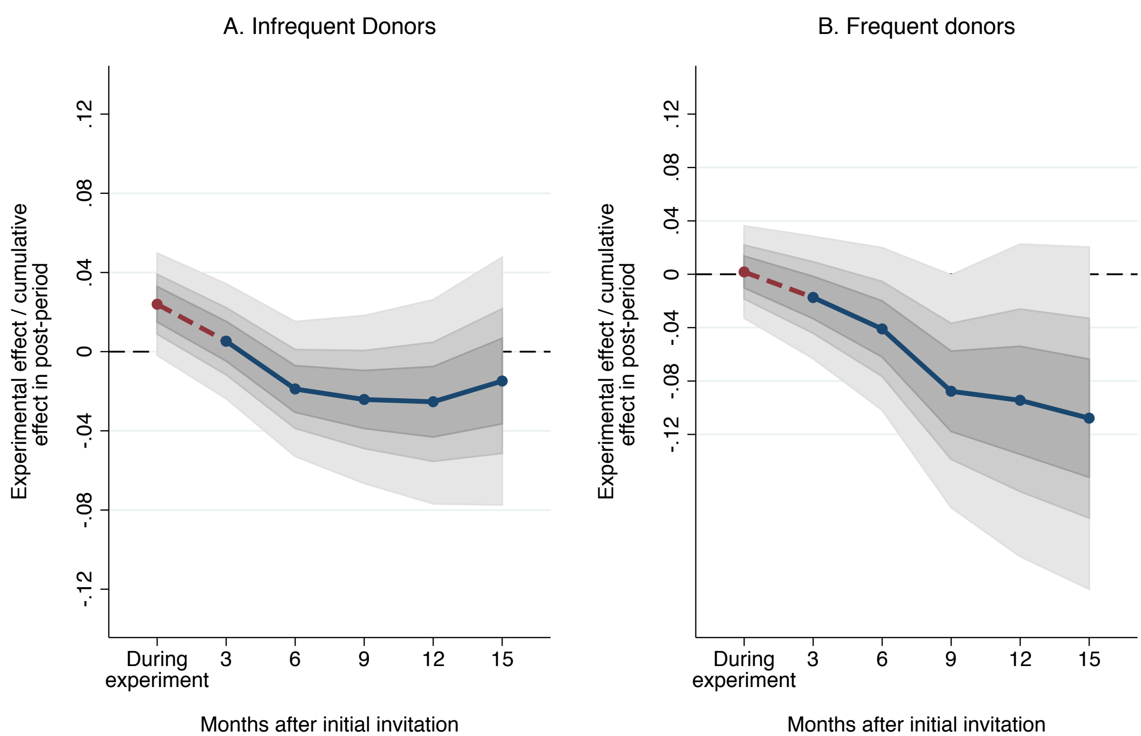
Notes: The three panels display the impact on the cumulative number of donations (relative to the main control treatment) estimated by OLS. The numbers are calculated based on the regressions in Table 7 for the number of blood donations following k months after the experiment. The shaded regions display the 50, 75 and 95 percent confidence intervals in decreasing shading, respectively. The sample includes only donors who have received at least four previous invitations to donate blood.

Figure 7: Long-run Effects of the Lottery Ticket by Frequency of Prior Donations



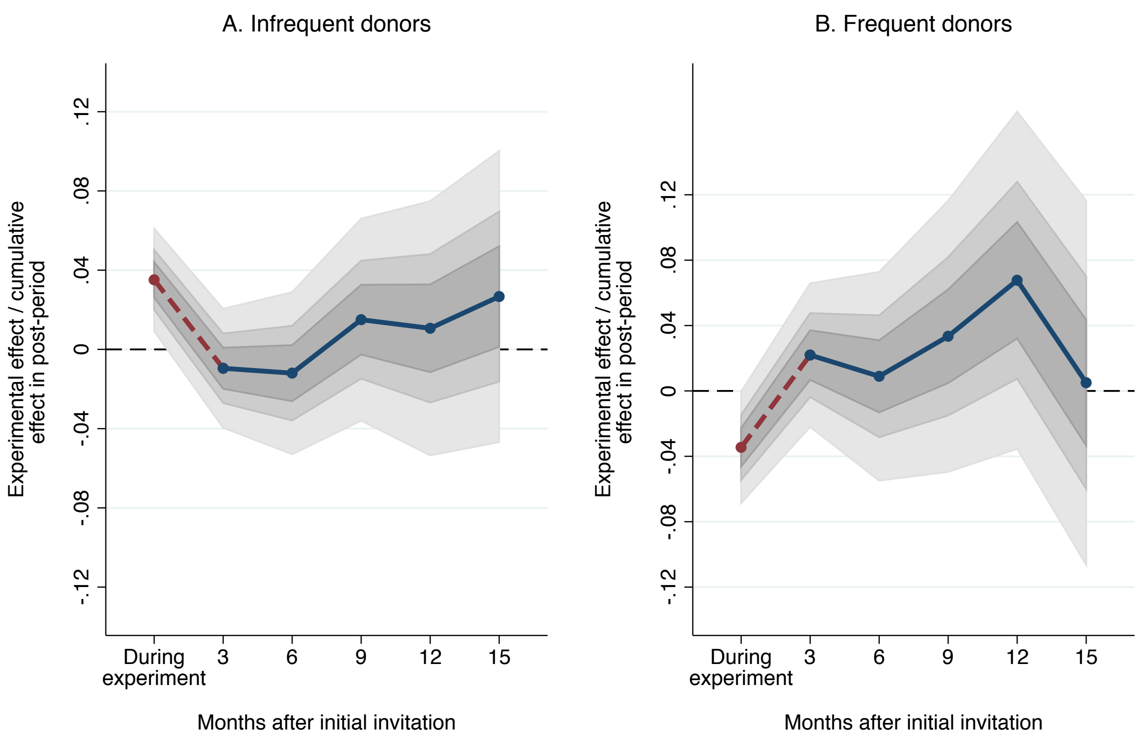
Notes: See Figure 6. The cumulative effects on donations and their confidence intervals are calculated based on the estimates in Tables 8 and 9.

Figure 8: Long-run Effects of the Cholesterol Test by Frequency of Prior Donations



Notes: See Figure 6. The cumulative effects on donations and their confidence intervals are calculated based on the estimates in Tables 8 and 9.

Figure 9: Long-run Effects of the Appeal by Frequency of Prior Donations



Notes: See Figure 6. The cumulative effects on donations and their confidence intervals are calculated based on the estimates in Tables 8 and 9.

C Tables

Table 1: Summary of the Behavioral Predictions

			Intertemporal utility spillovers		
			$\alpha < 1$ (Guilt)	$\alpha = 1$	$\alpha > 1$ (Habit formation)
Motivation crowding out	No	During experiment	Homogenous increase of donations across donors		
		After experiment	Decreasing negative effects	No effect	Decreasing positive effects
	Yes	During experiment	Ambiguous effect on the level of blood donations Relatively less effective for highly motivated donors		
		After experiment	Decreasing negative effects with remaining negative level effect	Negative level effect	Decreasing positive effects (or increasing negative effects) turning into a negative level effect

Table 2: Descriptive Statistics

	Age	Gender (female = 1)	Donations following last four invitations	Signed up for fixed interval of inv. (=1)	N
Overall sample					
Full sample	43.52	0.41	–	0.33	11,141
	(14.14)	(0.49)	–	(0.47)	
Sample with at least four previous invitations	44.77	0.39	1.97	0.35	9,731
	(13.87)	(0.49)	(1.55)	(0.48)	
By treatment					
Baseline	44.98	0.40	1.98	0.37	2,973
	(13.73)	(0.49)	(1.54)	(0.48)	
Appeal	45.66	0.40	2.11	0.35	2,224
	(13.84)	(0.49)	(1.54)	(0.48)	
Cholesterol test	44.12	0.38	1.90	0.26	3,544
	(13.96)	(0.49)	(1.56)	(0.44)	
Lottery ticket	44.47	0.38	1.83	0.63	990
	(13.92)	(0.49)	(1.53)	(0.48)	
Balance checks					
Test for full sample (p-value)	0.04	0.59	0.03	0.58	
Overall p-value for all four characteristics	0.19				
Test for sample with at least four previous invitations (p-value)	0.06	0.86	0.07	0.86	
Overall p-value for all four characteristics	0.42				

Notes: Numbers in parentheses are standard deviations. Descriptive statistics by treatment condition are calculated based on the sample with at least four previous invitations. The p-values for balance checks are obtained from seemingly unrelated regression (SUR) model for the four characteristics.

Table 3: The Distribution of Treatments
(Number of Subjects)

	Donation center				
Treatment	1	2	3	4	Total
Baseline	1,987	512	393	790	3,682
Appeal	1,378	373	601	261	2,613
Cholesterol test	2,911	615	579		4,105
Lottery ticket				1,160	1,160
Total	6,185	1,392	2,521	1,222	11,320

Notes: The sample refers to the full sample in Table 2.

Table 4: Overall Treatment Effects on Blood Donations
Dependent variable: Solicitation resulted in blood donation (=1)
Marginal effects from logit models

Appeal	0.031	0.026	0.008
	(0.022)	(0.022)	(0.015)
Cholesterol test	0.008	0.012	0.017
	(0.022)	(0.021)	(0.017)
Lottery ticket	0.041*	0.048**	0.056**
	(0.025)	(0.022)	(0.024)
Control variables			
Gender (female = 1)		–0.039***	–0.025*
		(0.012)	(0.015)
No. prev. inv. followed			
1 out of 4			0.251***
			(0.019)
2 out of 4			0.376***
			(0.016)
3 out of 4			0.499***
			(0.012)
4 out of 4			0.639***
			(0.010)
Differences between treatments			
Lottery ticket vs. ... (p-values)			
Appeal	0.705	0.357	0.049
Cholesterol test	0.294	0.201	0.159
Other controls?	No	Yes	Yes
Received at least 4 previous invitations	No	No	Yes
Pseudo-R^2	0.019	0.101	0.262
N	11,319	11,319	9,723

Notes: All specifications include a full set of dummy variables for donation center × week and donation center × weekday. Other controls include a full set of dummy variables indicating the periodicity of invitations as required by the donors. Baseline treatment is the standard invitation. "Appeal" indicates that a special card was added to the invitation, calling subjects up to donate. "Cholesterol test" means that together with the card, subjects were offered a free cholesterol test if they showed up. "Lottery ticket" refers to cards on which a lottery ticket was offered to donors.

Table 5: Treatment Effects for Frequent and Infrequent Donors
Dependent variable: Solicitation resulted in blood donation (=1)
Marginal effects from logit models

	Infrequent donors		Frequent donors	
Appeal	0.039**	0.037**	–0.032*	–0.036**
	(0.020)	(0.014)	(0.016)	(0.017)
Cholesterol test	0.026	0.023	0.000	0.002
	(0.019)	(0.015)	(0.018)	(0.018)
Lottery ticket	0.077***	0.082***	–0.035	–0.013
	(0.027)	(0.023)	(0.037)	(0.034)
Control variables				
Gender (female =1)		–0.015		–0.021
		(0.013)		(0.017)
No. prev. inv. followed				
1 out of 4		0.201***		
		(0.021)		
2 out of 4		0.331***		
		(0.023)		
4 out of 4				–0.153***
				(0.013)
Differences between treatments				
Lottery ticket vs. ... (p-values)				
Appeal	0.084	0.028	0.467	0.252
Cholesterol test	0.058	0.015	0.180	0.336
Different response to treatments?	$p < 0.008^a$	$p < 0.003^b$		
Other controls?	No	Yes	No	Yes
Pseudo-R^2	0.026	0.114	0.030	0.092
N	5,647	5,647	4,079	4,072

Notes: See Table 4. [a] $\chi^2(3)$-test for equality of treatment effects in the first and the third column. [b] $\chi^2(3)$-test for equality of treatment effects in the second and fourth column.

Table 6: Treatment Effects in Different Subsamples
Dependent variable: Solicitation resulted in blood donation (=1)
Marginal effects from logit models

	Gender		Age		Fixed interval inv.	
	Female	Male	Young	Old	No	Yes
Appeal	0.022	–0.005	0.015	–0.004	0.003	0.011
	(0.029)	(0.025)	(0.023)	(0.025)	(0.025)	(0.031)
Cholesterol test	0.026	0.003	0.037	–0.009	0.032	–0.026
	(0.029)	(0.025)	(0.023)	(0.025)	(0.023)	(0.034)
Lottery ticket	0.049	0.060	0.063	0.034	0.041	0.053
	(0.052)	(0.043)	(0.042)	(0.042)	(0.060)	(0.041)
Control variables						
Gender (female = 1)			–0.020	–0.027	–0.047***	–0.015
			(0.016)	(0.018)	(0.018)	(0.022)
No. prev. inv. followed						
1 out of 4	0.250***	0.250***	0.239***	0.198***	0.261***	0.237***
	(0.032)	(0.024)	(0.027)	(0.023)	(0.025)	(0.058)
2 out of 4	0.375***	0.372***	0.381***	0.296***	0.388***	0.354***
	(0.027)	(0.019)	(0.025)	(0.017)	(0.020)	(0.120)
3 out of 4	??	0.372?***				**
	(0.022)	(0.015)	(0.021)	(0.016)	(0.018)	(0.187)
4 out of 4	0.638***	0.643***	0.656***	0.596***	0.667***	0.587**
	(0.017)	(0.013)	(0.016)	(0.015)	(0.014)	(0.227)
Diff. response to treatm.?	$p = 0.86$		$p = 0.53$		$p = 0.44$	
Pseudo-R^2	0.251	0.270	0.219	0.223	0.279	0.244
N	3,792	5,931	4,782	4,940	6,296	3,433

Notes: See Table 4. The age cutoff is 45. The specifications for gender, age and individuals signed up for a fixed interval of invitations include controls for the required periodicity of invitations.

Table 7: Long-run Effects of the Experiment
Number of donations within different periods after the intervention
OLS estimates

	1 to 3 months	1 to 6 months	1 to 9 months	1 to 12 months	1 to 15 months
Appeal	0.009	0.002	0.030	0.040	0.025
	(0.013)	(0.016)	(0.021)	(0.027)	(0.030)
CH Test	–0.003	–0.028*	–0.045**	–0.047*	–0.043
	(0.013)	(0.016)	(0.022)	(0.028)	(0.032)
Lottery Ticket	0.039**	0.024	0.037*	0.053**	0.056*
	(0.019)	(0.020)	(0.022)	(0.026)	(0.030)
Age	0.007***	0.013***	0.024***	0.035***	0.050***
	(0.002)	(0.003)	(0.005)	(0.006)	(0.007)
Age squared	–0.000**	–0.000**	–0.000***	–0.000***	–0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender (female = 1)	–0.018*	–0.048***	–0.092***	–0.141***	–0.187***
	(0.010)	(0.013)	(0.016)	(0.022)	(0.028)
No. prev. inv. followed					
1 out of 4	0.137***	0.299***	0.451***	0.595***	0.710***
	(0.016)	(0.025)	(0.033)	(0.041)	(0.051)
2 out of 4	0.230***	0.520***	0.832***	1.102***	1.344***
	(0.017)	(0.023)	(0.033)	(0.041)	(0.050)
3 out of 4	0.330***	0.768***	1.176***	1.599***	1.978***
	(0.018)	(0.027)	(0.037)	(0.047)	(0.056)
4 out of 4	0.470***	1.044***	1.557***	2.091***	2.588***
	(0.018)	(0.027)	(0.036)	(0.048)	(0.058)
R^2	0.246	0.393	0.451	0.488	0.507
N	9593	9593	9593	9593	9593

Notes: See Table 4. All specifications include controls for the required periodicity of invitations.

Table 8: Long-run Effects of the Experiment for Infrequent Donors
Number of donations within different periods after the intervention
OLS estimates

	1 to 3 months	1 to 6 months	1 to 9 months	1 to 12 months	1 to 15 months
Appeal	–0.010	–0.012	0.015	0.011	0.027
	(0.015)	(0.021)	(0.026)	(0.033)	(0.037)
CH Test	0.006	–0.019	–0.024	–0.025	–0.014
	(0.015)	(0.017)	(0.022)	(0.026)	(0.032)
Lottery Ticket	0.010	0.015	0.000	0.028	0.041
	(0.019)	(0.029)	(0.032)	(0.041)	(0.052)
Age	0.005*	0.006	0.011*	0.015**	0.023***
	(0.003)	(0.004)	(0.006)	(0.007)	(0.008)
Age squared	–0.000	–0.000	–0.000	–0.000	–0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender (female = 1)	–0.004	–0.028*	–0.049**	–0.083***	–0.114***
	(0.010)	(0.015)	(0.020)	(0.025)	(0.028)
No. prev. inv. followed					
1 out of 4	0.134***	0.282***	0.417***	0.546***	0.654***
	(0.015)	(0.023)	(0.030)	(0.037)	(0.045)
2 out of 4	0.231***	0.504***	0.802***	1.061***	1.299***
	(0.016)	(0.022)	(0.031)	(0.038)	(0.045)
R^2	0.116	0.215	0.278	0.298	0.307
Obs	5564	5564	5564	5564	5564

Note: See Table 7.

Table 9: Long-run Effects of the Experiment for Frequent Donors
Number of donations within different periods after the intervention
OLS estimates

	1 to 3 months	1 to 6 months	1 to 9 months	1 to 12 months	1 to 15 months
Appeal	0.022	0.009	0.033	0.068	0.005
	(0.022)	(0.032)	(0.042)	(0.052)	(0.057)
CH Test	–0.017	–0.041	–0.088**	–0.094	–0.108*
	(0.023)	(0.031)	(0.044)	(0.059)	(0.065)
Lottery Ticket	0.064	0.016	0.064	0.064	0.063
	(0.045)	(0.042)	(0.054)	(0.057)	(0.076)
Age	0.005	0.019***	0.041***	0.067***	0.096***
	(0.005)	(0.006)	(0.009)	(0.011)	(0.013)
Age squared	–0.000	–0.000**	–0.000***	–0.001***	–0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender (female = 1)	–0.037*	–0.073***	–0.144***	–0.219***	–0.286***
	(0.019)	(0.025)	(0.030)	(0.040)	(0.052)
No. prev. inv. followed 3 out of 4	–0.111***	–0.230***	–0.312***	–0.405***	–0.499***
	(0.016)	(0.022)	(0.029)	(0.036)	(0.043)
R^2	0.216	0.194	0.222	0.232	0.247
Obs	4029	4029	4029	4029	4029

Note: See Table 7.