

Blood Donations and Incentives: Evidence from a Field Experiment*

Lorenz Goette[†] and Alois Stutzer[‡]

June 18, 2008

Abstract

There is a longstanding concern that material incentives might undermine prosocial 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 incentive: a lottery ticket and a free cholesterol test. Lottery tickets significantly increase donations, in particular among less motivated donors. The cholesterol test leads to no discernable impact on usable blood donations. If anything, it creates a small negative selection effect in terms of donations that must be discarded.

Keywords: prosocial behavior, blood donations, material incentives, field experiment

JEL classification: C93, D64, H41, I18

*We thank Tyler Williams for excellent research assistance and Roland Bénabou, Anat Bracha, Justina Fischer, Holger Herz, Stephan Meier, Kurt Schmidheiny, and the participants of seminars at Cornell University, the University College Dublin, the University of Lausanne, the University of St. Gallen, the 2008 ASSA meetings in New Orleans, and the 2007 End-of-Year Conference of the Network of Swiss Economists Abroad for helpful comments. The views expressed herein do not necessarily reflect the views of the Federal Reserve Bank of Boston or the Federal Reserve System.

[†]Federal Reserve Bank of Boston, Center for Behavioral Economics and Decision Making. Email: lorenz.goette@bos.frb.org

[‡]University of Basel, Department of Business and Economics. Email: alois.stutzer@unibas.ch

I. Introduction

In medical emergencies, blood transfusions are often the only way to save an individual's life (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 a single 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 prosocial motivation of their donors.¹ Despite the inherent free-rider problem of this policy, the arrangement seems to have worked satisfactorily most of the time. Three problems have, however, emerged recently that increase the risk of blood shortages in the future. First, innovations in surgery and oncological therapies have led to more aggressive medical treatments in cases that were previously thought inoperable or incurable, increasing the number of transfusions, and requiring larger amounts of blood (Davey, 2004). Second, there is a general tightening of donation eligibility criteria, such as stepped-up travel restrictions or restrictions on past blood recipients because of Creutzfeldt-Jakob disease. Third, there are widespread seasonal shortages due to reduced supply, particularly in summer and early winter (Gilcher and McCombs, 2005).

In this paper, we ask whether selective incentives can be used to overcome shortages in an environment that relies primarily on the prosocial motivations of its donors. Most economic models, including those incorporating prosocial preferences (for example, Andreoni, 1990) explicitly, predict that selective incentives increase blood donations. 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 attract at-risk donors, and—worse—undermine the motivation to donate blood.²

Research in psychology and economics has proposed that incentives can lead to less prosocial behavior. If rewards are perceived as controlling, incentives may undermine prosocial motivation. This is often referred to as the motivation crowding effect (Deci and Ryan, 1985; Frey, 1997; Lepper and Greene, 1978). More recently, economic models have been developed that

¹While whole-blood donations are generally unpaid in developed countries (World Health Organization, 2004), there are some countries where plasma donations, which can be done frequently, are paid. Indeed, Trimmel, Lattacher, and Janda (2005) find that, using survey measures, whole-blood donors are broadly more prosocially 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.

²Both conjectures are 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).

predict similar behavioral reactions, but for different reasons. In the model by Bénabou and Tirole (2006), individuals care about appearing “prosocial” and not mercenary. Accordingly, introducing incentives dilutes the positive signal value of a prosocial activity, and, as a consequence, individuals may contribute less. In a clever experiment, Ariely, Bracha, and Meier (forthcoming) manipulate the visibility of prosocial behavior and show that such image concerns can render incentives completely ineffective. Though not directly a test of this theory, the evidence in Gneezy and Rustichini (2000a,b) is also consistent with this model. A further reason why incentives may undermine prosocial incentives is conformity, as modeled in Sliwka (2007). The introduction of incentives signals to individuals that a large fraction of the population is not behaving prosocially, and, in response, conformists reduce their contribution. Finally, Ellingsen and Johannesson (forthcoming) develop a model in which individuals care about the motives of the counterparty. Individuals are willing to act prosocially only if they perceive the counterparty as non-selfish. Interpreted in our context, using incentives may signal that the Red Cross pursues selfish goals, which then reduces the motivation to donate blood.

With regard to the quality of blood donations, there is a serious concern that incentives undermine the truthful reporting of risk factors for infections.³ Even without using incentives, the blood donation service we study in this paper has to turn away donors or discard donations of about 10 percent of repeat blood donors. Thus, even regular donors at times fail to meet the donation criteria and medical screens. The fear is that this selection problem may be exacerbated if incentives are used. Individuals will then have an additional motivation to show up, even if they know that their donation is likely to fail the screenings. Moreover, they may be reluctant to disclose information that would jeopardize the donation reward.

In our study, we examine the effectiveness of selective incentives, based on a large-scale field experiment we conducted with over 10,000 blood donors.⁴ Previous studies have mainly surveyed donors’ attitudes toward blood donation incentives, or have conducted experiments on convenience samples that are not representative of the population of blood donors. The largest survey study (the Retrovirus Epidemiology Donor Study, REDS), based on more than 45,000 blood donors in the United States, analyzes donors’ responses to whether they would be encouraged, discouraged or remain neutral if they were offered some incentives (Glynn et al., 2003). Blood credits and free cholesterol testing were found to be the two incentives most likely to encourage net donations. Sixty-one percent of the donors responded that they would be encouraged to donate blood if offered a cholesterol test. Reported discouragement was highest

³Even though blood donations are screened for a variety of infectious diseases, the presence of many bacteria and viruses can only be established after a so-called window period following the infection. During the window period, it is thus crucially important that donors truthfully answer questions about their exposure to risks.

⁴See Andreoni (2007) and Meier (2007) for surveys on experimental evidence on prosocial preferences. Field experiments on charitable donations include Frey and Meier (2004), Karlan and List (2007), Landry et al. (2006), Croson and Shang (2008), and Stutzer, Goette, and Zehnder (2007).

for compensatory cash and lottery and/or raffle tickets. In an earlier part of REDS, reported attitudes were linked to the donor status of people at risk (Sanchez et al., 2001). It was found that cash seems relatively more attractive to donors who are HIV test-seeking and who have incurred health risks in the past. Free medical testing was found more likely to encourage individuals seeking HIV tests. There are no significant differences in the encouragement of at-risk donors when attitudes towards lottery tickets are assessed. However, both of the studies are based on reported attitudes, and it is not clear how these translate into behavior.

Studies that measure individual behavior and apply a randomized design are rare and often involve non-representative samples. From the available studies, no consensus finding emerges. One experiment finds that incentives work and increase blood donations (Ferrari et al., 2001), one that they are ineffective (Reich et al., 2006), and one even finds that incentives reduce participation in a preparatory health examination, at least in one subsample (Mellström and Johannesson, forthcoming). Yet, it is not clear to what extent these differences are driven by differences in samples, or by the environment in which the incentives are administered. This may be particularly important in light of the different theoretical conditions predicting the potentially detrimental effects of incentives.⁵

In this study, we examine the impact of selective incentives on donors who normally donate for prosocial motivations alone. In our experiment, some donors are offered a free cholesterol test in return for a blood donation, others are offered a lottery ticket. One group is not offered any incentives. It serves as a control. Importantly, our subjects were not aware that a study was being conducted.⁶ The treatments were mailed privately to the donors, so that public image concerns were excluded in our experiment. Thus, our study can be viewed as a test of psychological theories of intrinsic motivation (for example, Frey, 1997) or of self-signaling (Bénabou and Tirole, 2006). Furthermore, in one treatment, we send an appeal with the identical message as in the incentive treatments, but do not offer a reward. This allows us to examine whether information conveyed through the mailing alone affects blood donations (possibly about the behavior of others, as stressed in Sliwka (2007)). Our data contain information on basic demographics as well as on individuals' donation histories. This allows us to look in a systematic way at how the responses to the treatments vary between subsamples for which theories predict different responses. Further, the study allows us to analyze the impact of incentives on two different outcomes. The first outcome we consider is for usable donations, the second outcome for donations that must be rejected or discarded by the donation service. By separating the

⁵For example, in Ellingsen and Johannesson (forthcoming), the key parameter is the extent to which blood donors believe that the solicitor of the blood donation has altruistic motives. Plausibly, this may vary depending on whether lab experiments are conducted, in which case the blood donation is solicited by the experimenter, or whether donations are made in a natural environment, with the Red Cross soliciting the blood donation.

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

two outcomes, we can distinguish between the impact of incentives on donor motivation and the impact on donor selection.

The overall picture that emerges from our experiment is that material incentives have no general negative effects. To the contrary, when looking at the overall experimental outcomes, we find that offering a lottery ticket increases usable donations by 5 percentage points over a baseline donation rate of 42 percent. By contrast, we find no economically and statistically significant effect of the free cholesterol test on usable donations. We further find that offering the lottery ticket is significantly more effective than appealing to donors without giving incentives or offering a free cholesterol test.

We also show that the treatment effects vary between subsamples in interesting ways. It is mainly subjects who donated only infrequently in the past who respond to the experimental treatments and increase donations. There is essentially no incentive effect on donors who have shown a strong preference for donating blood in the past. This pattern is evidence of heterogeneity in the motivations for blood donations. When we examine different splits of the sample, such as by age or gender, we find no significant differences in the responses to the treatments in these subsamples. This reinforces our interpretation that the driving force behind the different responses is differences in the motivation to donate blood.

With regard to the possible selection effects of our treatments, we find no evidence that offering the lottery ticket in any way attracts donors who have a higher propensity to generate a rejected donation. There is some evidence that offering the cholesterol test leads to a somewhat higher risk of the donor or donation being rejected, especially in women.

The remainder of this paper is structured as follows: Section II describes the empirical setup of our study, and explains the details of the treatments. It also discusses the behavioral predictions for the different treatments. Section III presents the results. Section IV concludes.

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 then subsequently invited to donate blood again at one of the four donation centers (the center is determined by proximity to the town where the individual lives). Approximately three weeks prior to the appointment, the donors are mailed an invitation giving a specific date, 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.

A. The Treatments

The experimental intervention used the standard procedure for approaching donors. This standard procedure was also applied in summer to a fraction of donors forming the control group.

For the experimental intervention, additional treatment-specific information was included in the regular written invitation. Three distinct treatments were employed to elude specific information in the summer campaign. For each of the three treatments, the information was provided on a card (15 x 21cm). The card's face side showed the image of a flower meadow with the standard text, "This summer, you can make a difference", whereas the reverse side bore a message 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 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. This campaign lasts from June 6 to August 31, 2006."

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. This campaign lasts from June 6 to August 31, 2006."

The retail price of a cholesterol test bought in a pharmacy is approximately CHF 15.- (USD 13) and that of a lottery ticket CHF 5.- (USD 4.30). Irrespective of the invitation, all the donors who showed up were treated equally at the donation center. For instance, even donors, who did

not know that they would get a lottery ticket were offered one. The reason was that there was no practical way of checking which invitation a donor received when he or she showed up at the center. For this reason, the lottery ticket treatment could only be implemented at donation center 4, the second largest donation center. The other three donation centers (centers 1 to 3) offered the donors a free cholesterol test. The sequence of the treatments was randomized over days; that is, only one treatment was mailed out per day. Thus, the treatment was randomized only within weeks and within donation centers. This requires us to control for donation center, week, and weekday in the empirical analysis below.

B. Descriptive Statistics

In total, 12,268 blood donors were mailed an invitation during the experiment. Table 1 gives an overview of the number of subjects in each treatment. We aimed at having one half of the subjects in the cholesterol test / lottery ticket treatment, one quarter in the appeal, and one quarter in the baseline, in each center, respectively. Table 2 shows the mean age, and percentage of female participants per treatment. As is shown, the mean values are very similar, which is not surprising, given the procedure that was applied.

Figure 1 gives an overview of the behavior of the donors in the year prior to the experiment. As can be seen in the figure, between 45 and 50 percent of the individuals who received an invitation showed up to donate blood. Of these, we distinguish two outcomes: usable donations (that is, donations that passed all the medical screens) and rejected donations (that is, donations that failed one or more of the medical screens and could not be used). These tests range from criteria such as “no flu episode over the last three weeks,” to restrictions on travel over the last six months (for example, to malaria regions), blood pressure limits, and HB counts. As can be seen, there is a stable fraction of approximately 5 percentage points of the invitations that result in rejected donations. Thus, rejected donations occur regularly even within our sample of repeat donors. We take the outcome variable, rejected donations, as a proxy for at-risk donations. For given screening procedures and screening criteria, we assume that a higher fraction of rejected donations indicates a higher risk of an undetected infectious blood component.

Figure 2 displays the outcomes during the experimental period. The first bar shows the overall outcomes: On average, an invitation generated a usable donation with probability 0.42, and a rejected donation with frequency 0.05. This is similar to what was observed a year earlier. The figure also shows the experimental outcomes for five subgroups. Based on people’s history of blood donations, we grouped the subjects into five groups in the following way: In response to the last four invitations sent to a subject, how many times did he or she show up to donate blood? In order to define this in a clean way, we exclude all the individuals who received fewer than four invitations before the experiment. This reduces the sample to 10,490 observations.

The bars show that the past frequency of donations is a powerful predictor of current donations. Therefore, we use this variable to group individuals by their preference for blood donations. Figure 2 shows that the absolute frequency of rejected donations increases with the number of donations in response to the previous invitations. Thus, rejected donations are common across the pool of donors, occurring across all donors, not just infrequent or new donors.

C. Behavioral Predictions

In this subsection, we discuss the predictions of several different theories to explain three key outcomes in our experiment: Usable donations, rejected donations and the relative effectiveness of different incentives. We distinguish between usable donations and rejected donations, because this strategy allows us to disentangle the impact of our treatments on donor motivation and the impact of donor motivation on the selection process itself. We first discuss the potential effects of the experimental treatments on usable donations, and then turn to rejected donations.

We focus on two types of theories that propose systematically different models of interaction between material incentives and prosocial motivation. First-generation models of prosocial behavior in economics treat offers of material rewards as selective incentives that generate a standard relative price effect (see, for example, the review of theoretical contributions on prosocial motivation in Andreoni (2007)). Material incentives are thus predicted to have a non-negative effect on the required behavior; that is, on voluntary contributions.

In contrast, second-generation models of prosocial behavior propose a possible interaction between material incentives and prosocial motivation. A first approach assumes a change in preferences due to an external intervention. Research from psychology posits that incentives can undermine the intrinsic motivation for a task (Deci and Ryan, 1985; Lepper and Greene, 1978). The basic idea is that incentives change the way individuals think about an activity: if incentives are used, performing the activity may impair perceptions of self-determination or self-esteem and thus the level of prosocial motivation. Overall, a lower level of activity may result. These mechanisms and potential applications in economics are discussed more extensively in Frey (1997) and Frey and Jegen (2001). A formalization of the approach can be found in Frey and Oberholzer-Gee (1997).

More recent theoretical work takes preferences as fixed and instead posits alternative mechanisms by which incentives are made ineffective. In a recent model by Bénabou and Tirole (2006), individuals differ in how much they prefer to contribute to a prosocial activity and in their marginal utility of money. Individuals also have an image concern: an individual wishes to be seen as someone who likes contributing to a prosocial 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 for this is that a positive response to such incentives may communicate an individual's positive marginal utility of money and thereby convey a selfish signal. Bénabou and Tirole propose two interpretations regarding the image concern in their model. First, individuals may be concerned with what their actions signal to outside observers (for example, neighbors observing that they are going to donate blood). Second, individuals signal to their future selves what their preferences were in the past (see, more extensively, Bénabou and Tirole, 2007). In our experiment, the invitations are mailed to the subjects privately. Thus, there is no scope for public image concerns in our setting, and it should be considered a test of only the self-signaling interpretation of their model.

Two other theoretical arguments that can predict a decrease in prosocial behavior when incentives are introduced warrant discussion. Ellingsen and Johannesson (forthcoming) develop a model in which an individual's willingness to engage in a prosocial activity depends on his or her belief about the counterparty's altruism, which they refer to in their paper as "pride." If the use of incentives signals that the counterparty is selfish, then individuals are less inclined to act in a way that enhances the other party's payoff.⁷ Given that the Red Cross applied the incentives only in order to overcome a shortage, and applied them only temporarily, it is difficult to argue that the use of incentives signals a change in the preferences of the Red Cross. Therefore, our design makes negative effects of incentives due to pride effects unlikely.

Sliwka (2007) proposes a model in which there are three types of agents: selfish ones, prosocial ones, and conformists. Conformists either behave like selfish types, or like prosocial types, depending on which type they believe is more prevalent in a society. The use of rewards can therefore signal to conformists that there are more selfish people in society, leading them to also behave in a selfish manner. Thus, in this model, the emphasis is not on the effectiveness of the incentives themselves, but rather on the information their use conveys about the behavior of individuals. In our experiment, the appeal treatment conveys exactly the same information about the shortage as the incentive treatments do. Thus, the appeal treatment allows us to examine this motive in isolation, and to compare it to the outcomes when incentives are used.

Usable Donations: The first outcome we look at is the frequency of usable donations; that is, the probability that an invitation led to a donation that passed all medical screens. The main and simple prediction from the first-generation models on prosocial behavior in economics is that offering a material incentive has a non-negative effect on usable donations.

The prediction is less clear if other psychological theories are taken into account and prosocial motivation or perceptions are endogenous to the use of incentives. Both the psychological interpretation of intrinsic motivation (Frey, 1997) and the self-signaling interpretation of Bénabou

⁷More formally, Ellingsen and Johannesson (forthcoming) mostly frame their results in terms of principal-agent models. They construct separating equilibria in which selfish principals use incentives, while non-selfish principals do not use incentives. It is a property of these equilibria that agents exert less effort when incentives are used.

and Tirole (2006) predict not only a relative price effect increasing blood donations but also either a negative effect on prosocial motivation or a lower signalling value. In sum, the sign of the net effect is indeterminate, and a lower level of blood donations is possible. The psychological model of intrinsic motivation makes the further prediction that any negative effect should be largest for those individuals with the strongest intrinsic motivation to donate, which we will measure by the propensity to react to invitations prior to the experiment. As the incentives used in this study are quite small compared with the considerable effort required by the donors to make a donation, the specific environment of our experiment seems highly appropriate for testing the theoretical prediction.

Rejected Donations: The second outcome we examine is rejected donations. Previous studies indicate that, in the few instances in which blood donors are paid for blood donations, there is a significantly higher risk of infection markers in these donations (Eastlund, 1998; van der Poel, Siefried, and Schaasber, 2002). However, in these studies, there is scope for attracting donors with poor characteristics over the long run. In contrast, in our treatment, there is much less scope for attracting new “low-quality” donors, since all our subjects are previous donors, and no incentives were used prior to the experiment. Further, the incentives are much smaller than in the settings analyzed in these other studies. There may be more subtle selection effects due to a differential attractiveness of the lottery ticket offered as the incentive. For example, donors who have are more risk-seeking in many areas of life have a higher willingness to pay for a lottery (Dohmen et al., 2005). Thus, it is possible that offering a lottery ticket also attracts individuals who incurred more health risks in the past.

Similarly, the treatment offering a cholesterol test may be more attractive to individuals with poor health status. High readings of cholesterol are associated with a host of adverse health outcomes and behaviors, such as metabolic syndrome, high blood pressure, and fewer workouts (Hata and Nakajima, 2000; Schmidt et al., 2008). Many of these correlates, in particular, high blood pressure and its treatment, make rejections of donations more likely. Positive or negative selectivity might result from applying the cholesterol test as an incentive, depending on whether individuals with a good or a bad cholesterol reading find the free test more attractive. Previous evidence suggests that among blood donors, individuals with worse health status tend to seek medical tests (Sanchez et al., 2001).

The Relative Effectiveness of Incentives: Previous evidence from survey studies suggests that blood donors prefer being offered a health test rather than a material reward such as a lottery ticket (Glynn et al., 2003). This is in stark contrast to the prediction from any model in economics, that, holding the monetary value constant, individuals will prefer cash rewards over in-kind rewards. In our case, the retail price of the cholesterol test is about three times the price of the lottery ticket. Thus, a prediction of the relative effectiveness based on the economic model is not possible. We therefore contrast our behavioral evidence with the survey evidence

on blood donors’ attitudes towards various kinds of rewards.

D. The Empirical Strategy

As described earlier, we examine two experimental outcomes, usable and rejected donations. We model the probability that a usable donation is generated as

$$\Pr(\textit{use}_{ict}) = \Pr(\gamma \textit{treat}_{ct} + x_i b + \delta_c(t) + \omega_c(t) + e_i > 0), \tag{1}$$

where *use* is a dummy variable equal to 1 if the invitation of individual *i* (invited to center *c* on day *t*) resulted in a usable donation and 0 if it did not. The vector *treat* is a set of dummy variables, equal to 1 if the respective treatment was mailed out in center *c* on date *t*, and 0 otherwise. Thus, the vector γ summarizes the impact of the treatments on the probability of donating blood relative to the baseline invitation. The vector *x* contains demographic controls such as age, gender, and the donation history of the individual prior to the experiment. The variable $\delta_c(t)$ is a week fixed effect of date *t* in center *c*.⁸ We allow these week effects to differ across centers. We also allow separate weekday fixed effects $\omega_c(t)$. There is an immediate reason to do this, as not all centers are open for donations on all weekdays. This affects the attractiveness of other weekdays in a center-specific way. Given the way the treatments were rolled out over time, this specification ensures that, conditional on these variables, the treatment variation is fully randomized.

We assume that e_i is normally distributed with variance 1; thus equation (1) results in a probit model. We always report marginal effects in our tables.⁹ We have also estimated linear probability models based on (1), and the estimated coefficients are very similar.

The second experimental outcome we examine is the donations that had to be rejected. We adopt the same strategy in estimating the impact of the treatments on this outcome.

III. Results

A. The Overall Treatment Effects

We present the results in two steps. We first show a raw comparison of the treatments, normalizing the data only by donation center times donor frequency. With the randomization procedure described in Section II, it is necessary to remove possible center differences. We additionally normalize by past donor frequency, in order to make the treatment effects comparable throughout all graphs. We then present the estimates from the probit models, also controlling

⁸Since *t* indexes the date, the week is a function of the calendar date, which we express as the function $\delta_c(t)$.

⁹Denote the collection of explanatory variables by *Z*, and the coefficient vector by β . Then, $\partial \Pr(\textit{use}_{ict}) / \partial z_i = \phi(\bar{Z}\beta) \beta_i$, where \bar{Z} is the mean of all explanatory variables and $\phi()$ is the p.d.f. of the normal distribution.

for center-specific week effects, and center-specific weekday effects, at which level our treatments are fully randomized.

Usable Donations: Turning to the impact of the treatments, Figure 3 shows the treatment outcomes, normalized by the mean of donation center times donor frequency over all treatments. The figure shows the change in the fraction of usable donations of the three treatments relative to the baseline invitation, along with the standard error of the mean. As the figure shows, there is almost no effect of the appeal alone, and only a weak effect of the free cholesterol test. The lottery ticket, however, has a positive effect and appears to be the most effective of the three treatments. The frequency of donations is about 4 percentage points higher than in the baseline treatment.

Table 3 presents the marginal effects from probit estimates. The first column shows the estimates using the entire sample, and controlling for center-specific week effects and center-specific weekday effects. The second column adds gender and age as control variables. The third column restricts the sample to donors who have received at least four invitations prior to the experiment. In this column, we also add fixed effects for the number of times an individual donated blood in response to the last four invitations. The third column is our preferred specification, as it controls most tightly for differences between donors. The estimates confirm the qualitative picture from Figure 3. The only treatment with an economically significant impact on donations is the lottery ticket: The point estimate is always around 5 percentage points, irrespective of the specification, and it is statistically significant. Turning to our preferred specification (column 3), we also see that the lottery ticket is more effective than the appeal ($p = 0.049$). This shows that the lottery ticket increased participation over and above what could be attributed to the information contained in the extra postcard enclosed with the invitation.

Thus, the overall positive treatment effect of the lottery ticket on usable donations is in line with the standard prediction for selective incentives in economics. While previous donors in surveys rate free cholesterol testing as a valuable reward, surprisingly the health test is ineffective in increasing blood donations. In fact, the lottery ticket is significantly more effective ($p = 0.09$). Our results show no evidence of either of the two material incentives reducing blood donations, as derived as a potential outcome from psychological theories on the interaction between incentives and prosocial motivation.

Rejected Donations: We now turn to the second experimental outcome, which is the frequency with which an invitation resulted in a donor being rejected or a blood donation being discarded. The second panel in Figure 3 shows our first cut at the treatment effects, using only the crude normalizations to approximate randomization. There is no strong indication that any of the treatments had a clear selection effect. The point estimate for the cholesterol test is largest, but also has a substantial standard error. The results in columns (4) to (6) of Table 3 show the marginal effects from probit estimations. They estimate the impact of the cholesterol

test to be somewhat larger, and marginally significant, at least in our preferred specification with the tightest controls in place. Thus, if anything, the cholesterol test treatment increased the fraction of rejected donations, but not the fraction of usable donations.

B. Preferences for Blood Donations and the Response to the Experiment

As a next step, we examine how the response to the treatments differs by the preference to donate blood as measured by the frequency of donations prior to the experiment. We define two subsamples: Those individuals who responded only twice or less often to the last four invitations (infrequent donors), and those who donated three or four times (frequent donors).

Usable Donations: We proceed in the same way as before: Figure 4 shows the raw response to the treatments for the two subsamples. The first panel shows the response of the infrequent donors to the experiment. The infrequent donors respond to the lottery ticket most strongly. There is also an indication that the other two treatments (appeal and cholesterol test) have a small but positive effect. Turning to the frequent donors, the picture is qualitatively very different. In short, for these donors, no intervention seems capable of raising contributions above the level achieved with the baseline invitation.

We turn to a formal statistical test of these features in Table 4. In order to have a clean comparison between the two types of donors, we estimate separate probits for them. This way, we allow the control variables to have potentially different coefficients in the two subsamples. This avoids identification of the treatment effects by imposing identical coefficients on the control variables.

Columns (1) and (2) display the estimates for the infrequent donors, columns (3) and (4) for the frequent donors. The respective second specification is the one that parallels our preferred specification from Table 3. Studying the infrequent donors, the marginal effects from the probit model reflect the qualitative picture from Figure 4. They show that the lottery ticket has a strongly positive effect, estimated at about a 9 percentage point increase in useable donations. This effect is quantitatively important, and statistically significant. In comparing the effect of the lottery ticket, we again see that it is more effective than the appeal ($p = 0.07$), even though, for this group of donors, the appeal alone also significantly raises donations by about 3 percentage points. The lottery ticket, again, also works better than the cholesterol test ($p = 0.025$), which has essentially no effect on donations. The estimates also confirm the impression that no treatment raises usable donations for the frequent donors: In fact, the appeal lowers usable donations to a sizeable extent. The point estimates on the lottery ticket and the cholesterol test are actually negative, but not statistically significant. The effect of the lottery ticket is also not statistically significantly different from the effect of the appeal treatment, or from that of the cholesterol test.

It is also interesting to compare the estimates across donor groups. We clearly reject the hypothesis that the overall response to the treatments is the same in the two groups ($p < 0.01$). Some of the difference is driven by the negative response of frequent donors to being mailed the appeal. However, the lottery ticket also works significantly better for infrequent donors than for frequent donors ($p = 0.05$).

Rejected Donations: Turning to the impact of the treatments with regard to rejected donations in the two subgroups, Figure 5 gives a first impression of the results. Again, there are no large effects visible. The formal estimates are presented in Table 5. In both subsamples, point estimates for the selection effect are largest for the cholesterol test. However, the effect estimates, based on fewer observations than in the full sample, are not statistically significant. Overall, the results show no statistically significant differences in the responses between the two subsamples.

Interpretation: To summarize, the evidence presented in this section shows that the previous history of blood donation modulates the response to the treatments: Individuals who donated infrequently in the past respond to the treatments more positively. In particular, using a lottery ticket as a selective incentive has a large and positive effect on their donations: Usable donations increase by approximately 9 percentage points—a large increase over the baseline of about 25 percent of usable donations in this group. This response is significantly different from the effect of the lottery ticket on frequent donors. In fact, we observe a small, albeit statistically insignificant, decline in usable donations from frequent donors when they are treated with the lottery ticket. Overall, these results are consistent with models in which heterogeneity between donors stems from differences in their motivation to donate blood. For instance, in the model of Bénabou and Tirole (2006), obviously individuals with a low motivation to donate blood do so at a lower rate. Moreover, their marginal costs are lower, and thus it is easier to induce them to donate more by increasing the rewards somewhat. The result that the response to the lottery ticket is significantly stronger for the infrequent donors could also be interpreted as consistent with a motivation crowding effect (though the smoking gun evidence would have been a drop in donations by frequent donors in the treatments with material incentives). For the less motivated donors, the net effect of the positive relative price effect and the negative but small motivation crowding effect is larger than for the motivated donors, who experience a larger motivation crowding effect.

In contrast to evidence from surveys (Glynn et al., 2003), we do not find that the cholesterol test works well. Donors who are motivated to a greater or lesser degree do not provide an increase in usable donations in response to the cholesterol test. Our evidence shows that, on the contrary, the lottery ticket works significantly better than the cholesterol test in the overall sample. This effect stems mainly from the infrequent donors, for whom there is a large difference, while for frequent donors, the difference in incentives seems to matter much less. If anything, our evidence

indicates that offering a cholesterol test increases the undesirable outcome of rejected donations. Thus, our results provide an example showing that it is important to examine behavior, not attitudes as measured in a survey.

C. Responses in other Subsamples

In this subsection, we examine the implications of alternative interpretations regarding our results. In particular, we examine whether any of the differences in the response to the treatments that we found by cutting the sample by the frequency of donations are due to some other characteristics related to the frequency of donations. For example, gender is highly correlated with donations. Per invitation, women in our sample generate a usable donation with a 5-percentage point lower probability (see Table 3). Moreover, a previous study (Mellström and Johannesson, forthcoming) has found that women respond less well than men to incentives. We also examine whether the differences in treatment effects stem from differences in age. Older donors respond more reliably to invitations: For every year of age, usable donations increase by 0.5 percentage points. Finally, we examine whether the difference in treatment effects should be interpreted as a difference in commitment to a regular donation schedule rather than as a difference in the preference for donating blood.

Usable Donations: Table 6 presents the results. As it is important to control for all possible characteristics that may influence blood donations, in this exercise, we present only the strictest specifications. In columns (1) and (2), we cut the sample by gender. As the table shows, there is no overall difference in the response to the treatments. Men and women display the same pattern in response to the treatments. In particular, there is no difference in the response to the lottery ticket. This is in contrast to (Mellström and Johannesson, forthcoming), who find that women respond less well than men to incentives. In columns (3) and (4), we cut the sample by age. We use 45, the median age in our sample, as the age cutoff. Again, we find overall no statistically significant difference in the response to the treatments. The point estimate of the lottery ticket is somewhat higher for young donors than for old donors. However, the difference is not statistically significant at conventional levels ($p = 0.18$), and the difference between subsamples is more pronounced and of statistical significance when we cut the sample by the frequency of donations. Thus, we view age as a proxy variable for the frequency with which people respond to an invitation, not the other way around. Finally, we examine whether the observed differences are not due to differences in the preference for donations, but rather to differences in the regularity of schedules for donations. Maybe the more frequent donors have a fixed schedule that they like, and they simply do not deviate from this schedule. This could explain why infrequent donors are more responsive to our interventions. We can test this interpretation in the following way: The blood donation service offers the possibility of

indicating a choice of interval for being invited to donate blood. If nothing is indicated, the invitations are at least four months apart, but random. Donors can indicate any preference; that is, an invitation every four, six, or twelve months. This allows us to distinguish strength of prosocial preference from commitment to a regular donation schedule. A little more than a third of the individuals indicate such a preference for regular invitations. Columns (5) and (6) present the results. As can be seen, there is no meaningful difference in the point estimates of the treatment effects, and a formal statistical test fails to reject equality of the treatment effects across the two samples.

Rejected Donations: The exercise with subsamples is also conducted for the outcome of rejected donations. The results are presented in Table 7. Surprisingly, we find some indications of differences, at least with respect to the cholesterol test. There are significant effects of the cholesterol test on rejected donations for women, young donors, and donors who did not choose a regular donation schedule. However, we caution against interpreting these results very strongly for two reasons. First, even with respect to the cholesterol test, the differences between subsamples are only weakly statistically significant in the case of men versus women, and not statistically significant in the other two cases. Second, no specific hypothesis leads us to look for these differences. Therefore, the p -values on these three tests would have to be adjusted for multiple hypothesis tests. Once one corrects for this (using the Holm, 1979, procedure), the only significant effect at conventional levels is the effect of the cholesterol test on women’s rejected donations.

In sum, none of the alternative interpretations can explain the results we established in the previous section: There are no gender differences, age differences, or differences in commitment with regard to donation schedules that could explain our results. These findings further corroborate our interpretation that the differences in response to the treatments are related to preferences concerning blood donations.

IV. Summary and Conclusions

The use of selective incentives may reduce prosocial behavior rather than increase it. There is a longstanding concern, at least since Titmuss’s (1971) provocative work, that such a mechanism plays an important role in blood donations, which rely heavily on the prosocial motivations of the donors. In this paper, we test empirically how incentives affect blood donations in a large-scale field experiment over three months with more than 10,000 blood donors. This paper provides the first comprehensive empirical test of the Titmuss hypothesis, and studies different kinds of incentives. In one treatment, we apply a lottery ticket as an incentive to donate blood, in another a free cholesterol test. Both selective incentives have quite a low value for the donor so that the crowding-out effects observed in other domains might plausibly apply here.

Our results indicate no negative incentive effects on usable blood donations. In fact, the lottery ticket proves an effective incentive for raising the number of blood donations. In a second control treatment, filtering out the possible effects of the standard donor appeal, we still find a significant increase in usable donations when donors are offered a lottery ticket. On a closer examination of how the treatment effect varies with blood donation preferences, we find that the option of a lottery ticket is most effective with donors showing a historically weak preference for blood donation, while it is completely ineffective with conscientious donors—the statistical results being significantly different for the two groups. This pattern is broadly consistent with a model positing the heterogeneity of blood donation preferences. We also find that the lottery ticket is far more effective than the free cholesterol test, contrary to the findings of survey studies asking donors how appealing they find different types of incentives (Glynn et al., 2003).

This raises the question as to why the lottery ticket is so effective in raising donations, when monetary incentives have been shown to interact adversely with prosocial motivations in other experiments (Ariely, Bracha, and Meier, forthcoming; Gneezy and Rustichini, 2000a,b). There are two possible explanations for this difference. First, an important difference between these experiments and ours is that our treatments were all privately administered, and there was no obvious way for others to observe that an individual was promised an incentive to donate blood. By contrast, in the other experiments, a potential public image concern was always present. Thus, the absence of public image concerns may explain why the lottery ticket worked well in our field experiment.

Second, one could argue that the lottery ticket was not really perceived as a payment, but rather as a signal of goodwill on the part of the blood donation service. This was what the individuals were responding to, not the financial value of the lottery ticket. However, the same should be true for the cholesterol test. Yet, we do not find that the latter raises donations. Thus, this explanation does not account for the differences between the two kinds of incentives that we find in our data. Somewhat related, one could argue that it was not the money, but rather the gambling aspect that dominated in the perception of the lottery ticket. Had we used the equivalent of money, it would have had the negative consequences observed in other studies. This conjecture could be easily tested in a lab setting. However, even if it were true, our results suggest an easy way for organizations to circumvent it, by using lottery tickets instead of money.

We also examine the outcome of rejected donations; that is, donations that had to be discarded because they failed a medical test, or donors who showed up but had to be turned away in the initial screening at the blood donation center. We take rejected donations as a proxy for the selection of at-risk donors. Ten percent of all donations have to be rejected even among previous donors; thus it is important to know whether the use of incentives changes this outcome and indicates negative selection effects. We find no effect of the lottery ticket on rejected donations. This may not be very surprising, as the stakes are rather low, and the donor pool was fixed.

More negative consequences could be expected if higher payments were used on a permanent basis because this may attract new donors with worse characteristics. Still, we find some evidence that offering the health test leads to a slight increase in rejected donations. Apparently, the offer of the cholesterol test is particularly attractive to donors with temporarily impaired health, making use of the health test as an incentive even less desirable.

Overall, our results suggest that selective incentives and prosocial motivations may coexist even in domains that heavily rely on people's intrinsic motivation. Thus, in light of the recurring seasonal shortages and a steady tightening of donor criteria, incentives may prove useful to motivate previous blood donors to donate more. 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, there was essentially no public image concern, so an important channel by which incentives may become ineffective, as in Bénabou and Tirole (2006), was shut out. Furthermore, a switch to permanent incentives 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 (forthcoming). Thus, 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 incentive schemes affects the prosocial motivations of blood donors.

References

- Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100:464–477.
- . 2007. "Philanthropy." In Serge-Christophe Kolm and Jean Mercier Ythier, eds. *Handbook of the Economics of Giving, Reciprocity and Altruism*. Holland: North-Holland, vol. 2 of *Handbooks in Economics*.
- Ariely, Dan, Anat Bracha, and Stephan Meier. Forthcoming. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* (forthcoming).
- Bénabou, Roland, and Jean Tirole. 2007. "Identity, Dignity and Taboos: Beliefs as Assets." Working paper, Princeton University.
- . 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96:1652–1678.
- Croson, Rachel, and Jen Shang. 2008. "The Impact of Downward Social Information on Contribution Decisions." Unpublished, Mimeo, University of Texas at Dallas.
- Davey, Richard J. 2004. "Recruiting Blood Donors: Challenges and Opportunities." *Transfusion* 44:597–600.
- Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum Press.
- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Juergen Schupp, and Gert Wagner. 2005. "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey." IZA Discussion Paper No. 1730, IZA, Bonn.
- Eastlund, Ted. 1998. "Monetary Blood Donation Incentives and the Risk of Transfusion-Transmitted Infection." *Transfusion* 38:874–882.
- Ellingsen, Tore, and Magnus Johannesson. Forthcoming. "Pride and Prejudice: The Human Side of Incentive Theory." *American Economic Review* (forthcoming).
- Fernández Montoya, A., J. de Dios Luna del Castillo, A. López Berrio, and A. Rodríguez Fernández. 1996. "Attitudes, beliefs, and motivations in blood donors and non-donors." *Sangre* 41:427–440.
- Ferrari, Joseph R., Russell C. Barone, Leonard A. Jason, and Thomas Rose. 2001. "The Use of Incentives to Increase Blood Donations." *Journal of Social Psychology* 125:791–793.

- Frey, Bruno S. 1997. *Not Just for the Money: An Economic Theory of Personal Motivation*. Cheltenham: Edward Elgar Publishing.
- Frey, Bruno S., and Reto Jegen. 2001. "Motivation Crowding Theory: A Survey of the Empirical Evidence." *Journal of Economic Surveys* 15:589–611.
- Frey, Bruno S., and Stephan Meier. 2004. "Social Comparison and Pro-Social Behavior: Testing conditional Cooperation in a Field Experiment." *American Economic Review* 94:1717–1722.
- Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out." *The American Economic Review* 87:746–755.
- Gilcher, R. O., and S. McCombs. 2005. "Seasonal Blood Shortages Can Be Eliminated." *Current Opinion in Hematology* 12:503–508.
- Glynn, Simone A., Alan E. Williams, Catharie C. Nass, James Bethel, Debra Kessler, Edward P. Scott, Joy Fridey, Steven H. Kleinman, and George B. Schreiber. 2003. "Attitudes toward Blood Donation Incentives in the United States: Implications for Donor Recruitment." *Transfusion* 43:7–16.
- Gneezy, Uri, and Aldo Rustichini. 2000a. "A Fine is a Price." *Journal of Legal Studies* 29:1–18.
- . 2000b. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics*, pp. 791–810.
- Hata, Y., and K. Nakajima. 2000. "Life-Style and Serum Lipids." *Journal of Artherosclerosis and Thrombosis* 7:177–197.
- Higgins, C. 1994. "Blood transfusion: risks and benefits." *British Journal of Nursing* 3:986–991.
- Holm, Sven. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6:65–70.
- Karlan, Dean, and John A. List. 2007. "Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment." *American Economic Review* 97:1774–93.
- Landry, Craig E., Andreas Lange, John A. List, Michael K. Price, and Nicholas G. Rupp. 2006. "Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment." *Quarterly Journal of Economics* 121:747–782.
- Lepper, Mark R., and David Greene. 1978. *The Hidden Costs of Reward: New Perspectives on the Psychology of Human Motivation*. Hillsdale, N.J.: Erlbaum.
- Levitt, Steven, and John A. List. 2007. "What do Laboratory Experiments Measuring Social Preferences tell us about the Real World?" *Journal of Economic Perspectives* 21(2):153–174.

- Meier, Stephan. 2007. "A Survey of Economic Theories and Field Evidence on Pro-Social Behavior." In Bruno S. Frey and Alois Stutzer, eds. *Economics and Psychology: A Promising New Field*. Cambridge, MA: MIT Press.
- Mellström, Carl, and Magnus Johannesson. Forthcoming. "Crowding Out in Blood Donations: Was Titmuss Right?" *Journal of the European Economic Association* (forthcoming).
- Piliavin, J. A. 1990. "Why Do They Give the Gift of Life - a Review of Research on Blood-Donors since 1977." *Transfusion* 30:444–459.
- Reich, Pascale, Paula Roberts, Nancy Laabs, Artina Chinn, Patrick McEvoy, Nora Hirschler, and Edward L. Murphy. 2006. "A Randomized Trial of Blood Donor Recruitment Strategies." *Transfusion* 46:1090–1096.
- Sanchez, Ana M., Dannie I. Ameti, George B. Schreiber, Ruth A. Thomson, Annie Lo, James Bethel, and Alan E. Williams. 2001. "The Potential Impact of Incentives on Future Blood Donation Behavior." *Transfusion* 41:172–178.
- Schmidt, M., V. Cleland, R. Thomson, T. Dwyer, and A. Venn. 2008. "A Comparison of Subjective and Objective Measures of Physical Activity and Fitness in Identifying Associations with Cardiometabolic Risk Factors." *Annals of Epidemiology* 18:378–386.
- Sliwka, Dirk. 2007. "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes." *American Economic Review* 97:999–1012.
- Stutzer, Alois, Lorenz Goette, and Michael Zehnder. 2007. "Active Decisions and Blood Donations: Evidence from a Field Experiment." Working Paper Series No. 07-13, Federal Reserve Bank of Boston.
- Titmuss, Richard Morris. 1971. *The Gift Relationship: From Human Blood to Social Policy*. New York: Random House.
- Trimmel, Michael, Helene Lattacher, and Monika Janda. 2005. "Voluntary Whole-Blood Donors, and Compensated Platelet Donors and Plasma Donors: Motivation to Donate, Altruism and Aggression." *Transfusion and Apheresis Science* 33:147–155.
- van der Poel, C. L., E. Siefried, and W. P. Schaasber. 2002. "Paying for Blood Donations: Still a Risk?" *Vox Sanguinis* 83:285–293.
- World Health Organization, . 2004. *Global Database on Blood Safety. Report 2001-2002*. Geneva: World Health Organization.

V. Figures and Tables

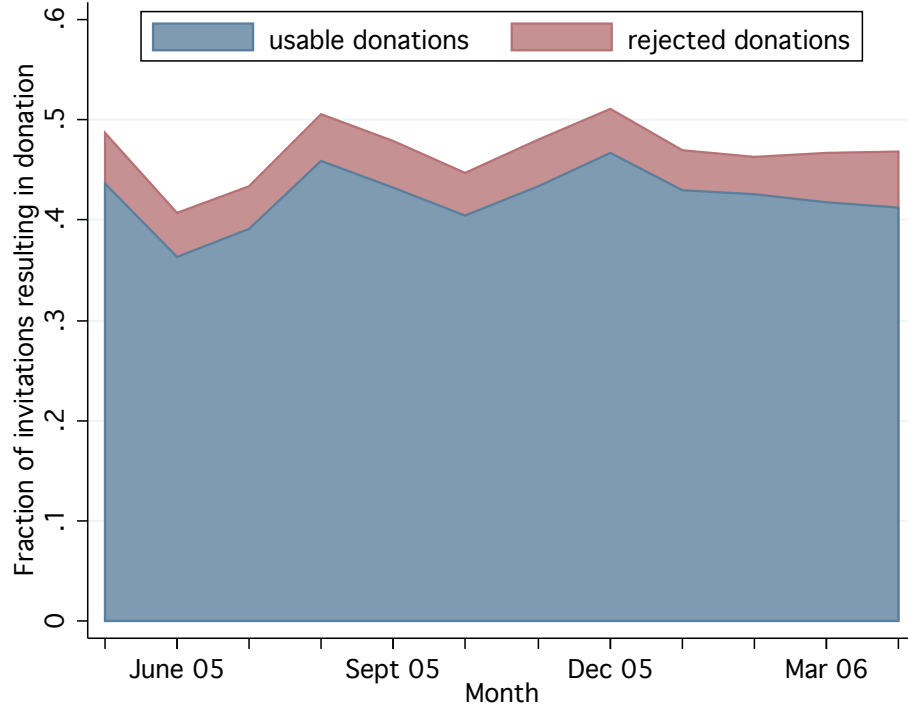


Figure 1: The Propensity to Donate Blood 12 Months Prior to the Experiment.

Notes: A usable donation leads to a blood product. A rejected donation is one where either the donor is rejected in the initial screening or the donation fails one of the screening tests after the donation.

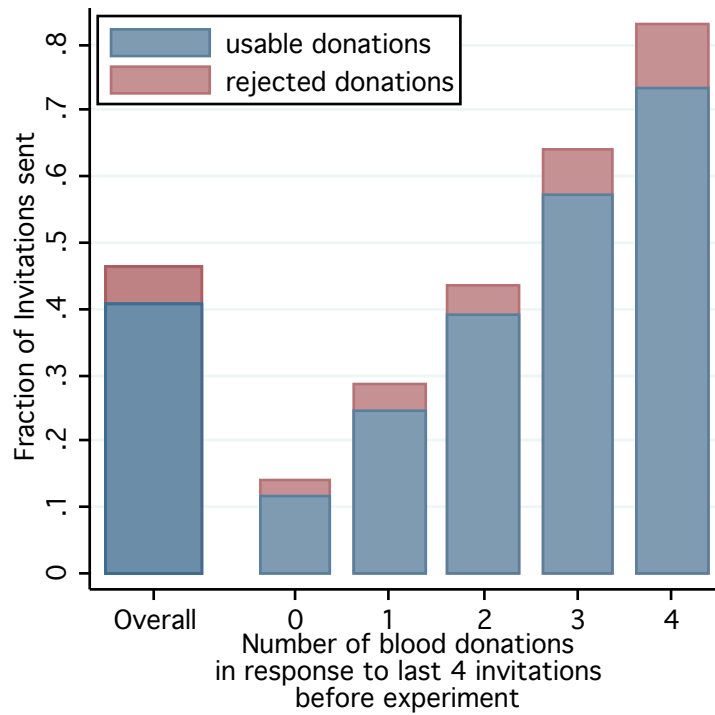


Figure 2: The Outcomes during the Experiment

Notes: A usable donation leads to a blood product. A rejected donation is one where either the donor is rejected in the initial screening or the donation fails one of the screening tests after the donation.

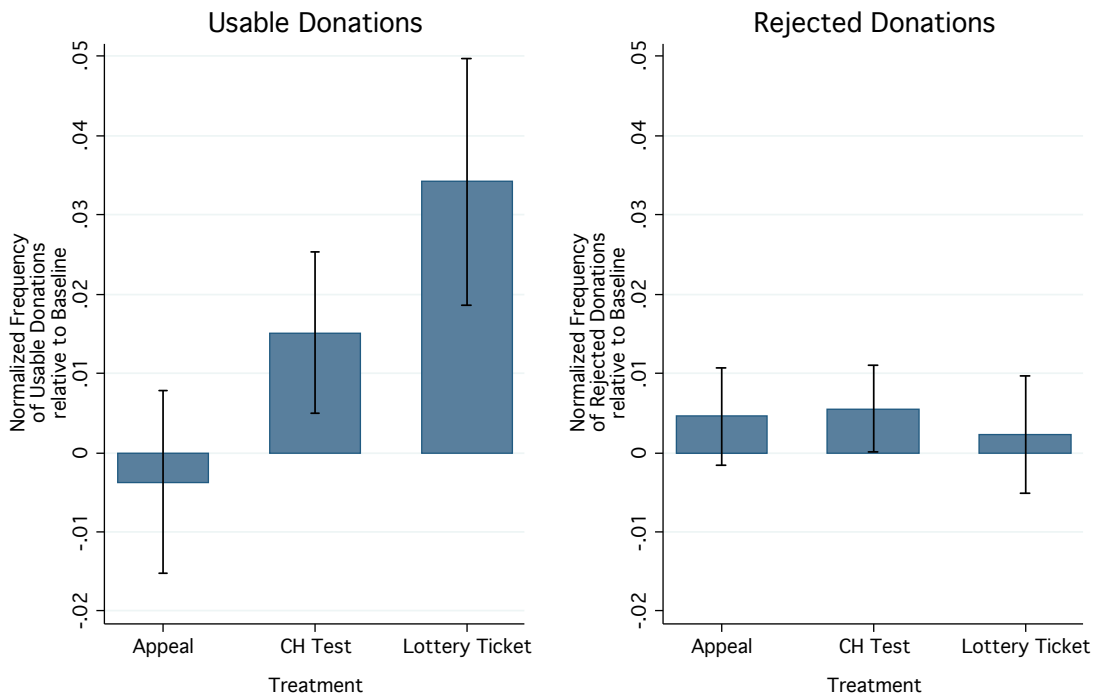


Figure 3: The Impact of the Treatments on Outcomes

Notes: The two outcome variables are normalized by donation center times donor type mean values, relative to the baseline treatment. The vertical lines indicate the standard error of the mean. A usable donation leads to a blood product. A rejected donation is one where either the donor is rejected in the initial screening or the donation fails one of the screening tests after the donation. CH test = cholesterol test.

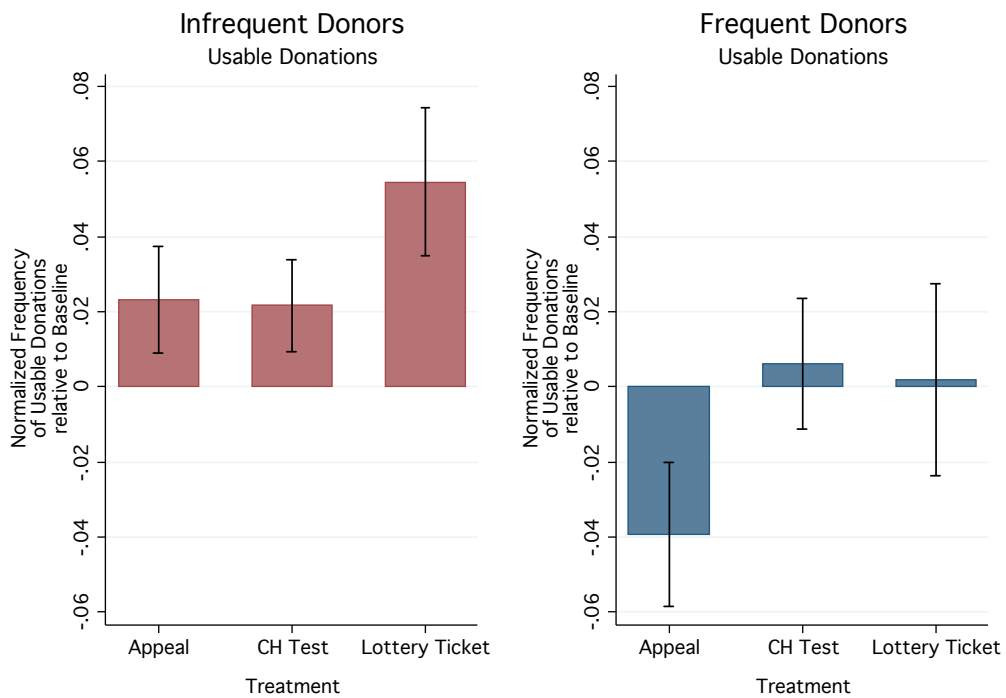


Figure 4: The Treatment Effect by Frequency of Prior Donations: Usable Donations

Notes: The outcome variable is normalized by donation center times donor type mean values, relative to baseline treatment. The vertical lines indicate the standard error of the mean. “Infrequent donors” indicates individuals who only responded twice or less to the last four invitations. “Frequent donors” indicates individuals who donated three or four times. A usable donation leads to a blood product. CH test = cholesterol test.

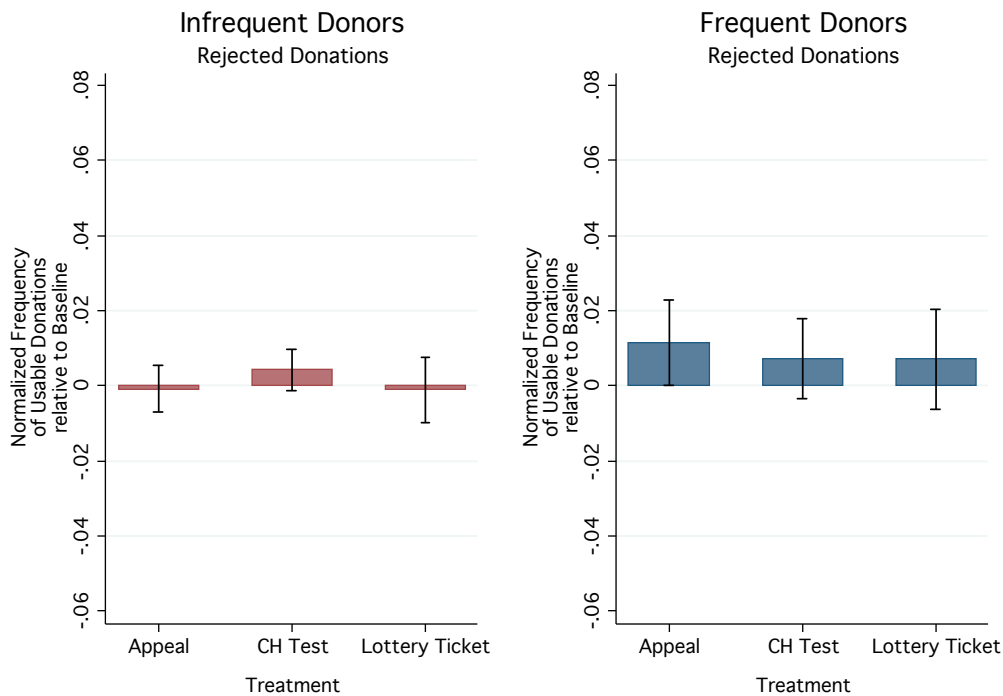


Figure 5: The Treatment Effects by Frequency of Prior Donations: Rejected Donations

Notes: The outcome variable is normalized by donation center times donor type mean values, relative to baseline treatment. The vertical lines indicate the standard error of the mean. “Infrequent donors” indicates individuals who only responded twice or less to the last four invitations. “Frequent donors” indicates individuals who donated three or four times. A rejected donation is one where either the donor is rejected in the initial screening or the donation fails one of the screening tests after the donation. CH test = cholesterol test.

Table 1: The Distribution of Treatments
(Number of Subjects)

<i>Treatment</i>	<i>Donation Center</i>				Total
	1	2	3	4	
Baseline	1,987	512	393	790	3,682
Appeal	1,523	498	284	645	2,950
CH test	3,014	815	602		4,431
Lottery ticket				1,205	1,205
Total	6,524	1,825	1,279	2,640	12,268

Notes: CH test = cholesterol test.

Table 2: Descriptive Statistics

	Age	Gender (female = 1)	<i>N</i>
<i>Overall</i>	43.35 (14.16)	0.41	12,268
<i>By treatment</i>			
Baseline	43.59	0.42	3,682
Appeal	43.91	0.41	2,950
CH test	42.95	0.39	4,431
Lottery ticket	42.69	0.40	1,205

Notes: CH test = cholesterol test.

Table 3: The Experimental Outcomes
Marginal Effects from Probit Model

	Solicitation resulted in usable donation (=1)			Solicitation resulted in rejected donation (=1)		
Appeal	0.020 (0.014)	0.015 (0.014)	-0.006 (0.016)	0.005 (0.006)	0.005 (0.006)	0.004 (0.006)
CH test	0.000 (0.014)	0.002 (0.014)	-0.001 (0.016)	0.010 (0.007)	0.009 (0.006)	0.011* (0.007)
Lottery ticket	0.049** (0.025)	0.054** (0.026)	0.053* (0.030)	-0.006 (0.011)	-0.005 (0.011)	-0.002 (0.011)
<i>Control variables</i>						
Age		0.010*** (0.000)	0.005*** (0.000)		0.001*** (0.000)	0.000** (0.000)
Gender (female = 1)		-0.058*** (0.010)	-0.047*** (0.012)		0.019*** (0.004)	0.019*** (0.005)
Number of previous invitations followed						
1 out of 4			0.214*** (0.019)			0.032*** (0.011)
2 out of 4			0.357*** (0.017)			0.042*** (0.011)
3 out of 4			0.487*** (0.014)			0.073*** (0.012)
4 out of 4			0.602*** (0.012)			0.095*** (0.012)
<i>Differences between treatments</i>						
Lottery ticket vs. ... (<i>p</i> -values)						
Appeal	0.251	0.137	0.049	0.349	0.377	0.615
CH test	0.069	0.061	0.088	0.217	0.241	0.286
Other controls?	No	Yes	Yes	No	Yes	Yes
Received at least 4 prev. invitations	No	No	Yes	No	No	Yes
Pseudo- R^2	0.018	0.090	0.219	0.015	0.033	0.060
N	12267	12267	10490	12225	12225	10453

Notes: All specifications include a full set of dummy variables for donation center times week and donation center times weekday. Other controls also include a full set of dummy variables indicating the minimum time difference between invitations set by the donation center. Baseline treatment is the standard invitation. “Appeal” means a special card was added to the invitation, calling subjects up to donate. “Lottery ticket” means that on this card, the subjects were offered a lottery ticket if they showed up. ”CH test” means that on the card, the subjects were offered a free cholesterol test if they showed up.

Table 4: Frequent vs. Infrequent Donors
 Dependent Variable: Solicitation resulted in usable donation (=1)
 Marginal Effects from Probit Model

	Infrequent donors		Frequent donors	
Appeal	0.035** (0.017)	0.032* (0.017)	-0.053** (0.023)	-0.056** (0.024)
CH test	0.016 (0.017)	0.014 (0.016)	-0.025 (0.024)	-0.020 (0.024)
Lottery ticket	0.087*** (0.032)	0.090*** (0.033)	-0.045 (0.045)	-0.021 (0.045)
<i>Control variables</i>				
Age		0.004*** (0.000)		0.005*** (0.001)
Gender (female =1)		-0.030*** (0.011)		-0.051*** (0.017)
Number of previous invitations followed				
1 out of 4		0.168*** (0.017)		
2 out of 4		0.296*** (0.017)		
4 out of 4				0.133*** (0.015)
<i>Differences between treatments</i>				
Lottery ticket vs. ... (<i>p</i> -values)				
Appeal	0.108	0.070	0.869	0.444
CH test	0.037	0.025	0.682	0.992
Different response to treatments?	<i>p</i> < 0.01 ^a	<i>p</i> < 0.01 ^b		
Other controls?	No	Yes	No	Yes
Pseudo- <i>R</i> ²	0.024	0.109	0.030	0.072
<i>N</i>	6142	6142	4340	4340

Notes: All specifications include a full set of dummy variables for donation center times week and donation center times weekday. Other controls also include a full set of dummy variables indicating the minimum time difference between invitations set by the donation center. ^a $\chi^2(3)$ -test for equality of treatment effects in columns (1) and (3). ^b $\chi^2(3)$ -test for equality of treatment effects in columns (2) and (4). Baseline treatment is the standard invitation. “Appeal” means a special card was added to the invitation, calling them up to donate. “Lottery ticket” means that on this card, the subjects were offered a lottery ticket if they showed up. “CH test” means that on the card, the subjects were offered free a cholesterol test if they showed up.

Table 5: Frequent vs. Infrequent Donors
 Dependent Variable: Solicitation resulted in rejected donation (=1)
 Marginal Effects from Probit Model

	Infrequent donors		Frequent donors	
Appeal	-0.000 (0.007)	0.000 (0.006)	0.007 (0.013)	0.008 (0.013)
CH test	0.008 (0.007)	0.008 (0.007)	0.022 (0.014)	0.021 (0.014)
Lottery ticket	-0.008 (0.009)	-0.006 (0.009)	0.010 (0.031)	0.012 (0.031)
<i>Control variables</i>				
Age		0.000*** (0.000)		0.000 (0.000)
Gender (female = 1)		0.012*** (0.005)		0.029*** (0.010)
Number of previous invitations followed				
1 out of 4		0.018*** (0.007)		
2 out of 4		0.023*** (0.007)		
4 out of 4				0.022*** (0.008)
<i>Differences between treatments</i>				
Lottery ticket vs. ... (<i>p</i> -values)				
Appeal	0.474	0.503	0.919	0.916
CH test	0.188	0.195	0.699	0.755
Different response to treatments?	$p = 0.89^a$	$p = 0.84^b$		
Other controls?	No	Yes	No	Yes
Pseudo- R^2	0.034	0.062	0.035	0.044
N	5932	5932	4336	4336

Notes: All specifications include a full set of dummy variables for donation center times week and donation center times weekday. Other controls also include a full set of dummy variables indicating the minimum time difference between invitations set by the donation center. ^a $\chi^2(3)$ -test for equality of treatment effects in columns (1) and (3). ^b $\chi^2(3)$ -test for equality of treatment effects in columns (2) and (4). Baseline treatment is the standard invitation. “Appeal” means a special card was added to the invitation, calling subjects up to donate. “Lottery ticket” means that on this card, the subjects were offered a lottery ticket if they showed up. “CH test” means that on the card, the subjects were offered free cholesterol test if they showed up.

Table 6: The Treatment Effect in Different Subsamples
 Dependent Variable: Solicitation resulted in usable donation (=1)
 Marginal Effects from Probit Models

	Gender		Age		Commitment	
	Female	Male	Young	Old	Low	High
Appeal	0.010 (0.024)	-0.020 (0.021)	0.014 (0.020)	-0.024 (0.023)	-0.017 (0.020)	0.012 (0.027)
CH test	-0.009 (0.024)	-0.003 (0.022)	0.021 (0.020)	-0.026 (0.024)	0.009 (0.020)	-0.034 (0.030)
Lottery ticket	0.052 (0.046)	0.060 (0.038)	0.078** (0.038)	0.024 (0.042)	0.061 (0.051)	0.052 (0.037)
<i>Control variables</i>						
Age	0.006*** (0.001)	0.005*** (0.001)	0.006*** (0.001)	0.003*** (0.001)	0.005*** (0.001)	0.006*** (0.001)
Gender (female = 1)			-0.048*** (0.014)	-0.028* (0.017)	-0.066*** (0.014)	-0.031 (0.019)
Number of previous invitations followed						
1 out of 4	0.200*** (0.030)	0.228*** (0.024)	0.191*** (0.023)	0.201*** (0.028)	0.214*** (0.024)	0.210*** (0.032)
2 out of 4	0.329*** (0.028)	0.374*** (0.021)	0.352*** (0.024)	0.319*** (0.022)	0.362*** (0.021)	0.341*** (0.029)
4 out of 4	0.585*** (0.022)	0.618*** (0.015)	0.592*** (0.020)	0.586*** (0.017)	0.620*** (0.015)	0.576*** (0.021)
Different response to treatments?		$p = 0.70$		$p = 0.33$		$p = 0.26$
Other controls?	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.216	0.219	0.193	0.178	0.225	0.217
N	4086	6401	5177	5310	6868	3617

Notes: Age cutoff is 45. All specifications include a full set of dummy variables for donation center times week and donation center times weekday. See Table 3 for further notes.

Table 7: The Treatment Effect in Different Subsamples
 Dependent Variable: Solicitation resulted in rejected donation (=1)
 Marginal Effects from Probit Models

	Gender		Age		Commitment	
	Female	Male	Young	Old	Low	High
Appeal	0.003 (0.011)	0.006 (0.008)	-0.002 (0.007)	0.007 (0.010)	0.009 (0.008)	-0.011 (0.010)
CH test	0.026** (0.012)	0.002 (0.008)	0.014* (0.008)	0.008 (0.010)	0.014* (0.008)	0.005 (0.012)
Lottery ticket	-0.007 (0.018)	0.002 (0.014)	-0.015 (0.009)	0.013 (0.022)	-0.009 (0.016)	-0.002 (0.015)
<i>Control variables</i>						
Age	-0.000 (0.000)	0.001*** (0.000)	-0.000 (0.000)	0.001** (0.000)	0.000** (0.000)	0.000 (0.000)
Gender (female = 1)			0.028*** (0.006)	0.006 (0.007)	0.023*** (0.006)	0.017*** (0.007)
Number of previous invitations followed						
1 out of 4	0.035** (0.016)	0.027* (0.014)	0.027** (0.011)	0.027 (0.019)	0.032** (0.013)	0.030* (0.017)
2 out of 4	0.040** (0.017)	0.042*** (0.015)	0.022** (0.011)	0.050** (0.020)	0.044*** (0.014)	0.039** (0.019)
4 out of 4	0.116*** (0.022)	0.077*** (0.014)	0.121*** (0.021)	0.074*** (0.016)	0.100*** (0.015)	0.077*** (0.020)
Different response to treatments?		$p = 0.70$		$p = 0.33$		$p = 0.26$
Other controls?	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.216	0.219	0.193	0.178	0.225	0.217
N	4086	6401	5177	5310	6868	3617

Notes: Age cutoff is 45. All specifications include a full set of dummy variables for donation center times week and donation center times weekday. See Table 3 for further notes.