Altruism and Impatience: The Role of Time Preferences in Donation Choices

July 5, 2020

Abstract

Relying on a quasi-experimental setting and a sample of Italian households, we study the role of time preferences on donation choices. Specifically, we apply the generalized propensity score methodology finding that both the amount and the probability of donating (i.e. altruism) vary non-monotonically with impatience in intertemporal choice, declining at higher levels of impatience. We then provide a theoretical model whose predictions are consistent, under fairly general conditions, with these empirical findings. Consistent with previous experimental evidence, these results support the view that psychological discounting matters for altruistic behaviour.

Keywords: Two-limit Tobit, Generalized Propensity Score, Quasi-experiment, Altruism, Discounting

JEL Classification: C21, D03, D64

Word count: 9090

1 Introduction

Altruism is a trait of mankind that cannot be ignored (Harman 2010; Kirchgässner 2010). Nevertheless, for long it hardly fitted in the *homo œconomicus* scheme. In models populated by selfish individuals, altruistic donations were one of the biggest puzzles in economics (Andreoni 2006). After Becker (1974)'s seminal contribution on social interacting agents, however, the role of altruistic preferences in accounting for individual choices has been more and more recognized by economic theory, and several forms of other-regarding preferences have been put forward to cope with the economic puzzle of philanthropy (Andreoni 2006; Bénabou and Tirole 2006). Typically, altruistic preferences are introduced in the standard model of consumer choice by adding up an altruistic term - weighted by the degree of altruism - to one's own consumption (Shapiro 2010).

However, while the economic literature on the reasons behind altruism is well developed (see, for example, Ottoni-Wilhelm et al. 2017), the literature on the drivers of altruism is still sparse, and mainly focuses on the effect of tax systems on money donations.[1] The role of deep preferences remains much less investigated. Yet, individual parameters and attitudes, generally not controlled for in many studies, seem to have an important role in affecting individual behavior, especially in countries where tax incentives (e.g. tax deductions) are less important (see Cappellari et al. 2011) and social expenditure is low (see Bauer et al. 2013); neglecting them may cause serious concerns in terms of omitted variables (Albanese et al. 2013).

With this paper we aim to contribute to the literature on altruistic behaviour by studying the relationship between donations and a specific preference trait, i.e. impatience (the propensity to postpone gratification in exchange of delayed rewards). To the best of our knowledge, only a few papers have provided so far specific evidence about it.

In two field experiments, Breman (2011) shows that the amount donated increases when donors commit to future donations; nevertheless, as the experiments cannot control for differences in individual time preferences, it is not possible to disentangle whether this effect is due to differences in intertemporal preferences or budget constraints. Andreoni and Serra-Garcia (2016) show that introducing time in the process of giving (in their model individuals derive two distinct utility payoffs when they make a giving decision and when they pay for the gift) can uncover important dynamics and heterogeneity in the motivations for giving. In line with Breman (2011), their experiments find that donations increase when

[1] For example, List (2011) studies international heterogeneity in donations with respect to differences in tax policies and national attitude towards social needs. The analysis shows that the likelihood of donating increases (non-linearly) with household income and education.

individuals are asked to donate later.[2] On a theoretical ground, Dreber et al. (2016) find that an increase in the cost of self-control will increase altruistic behaviour in a model of dual-self where a patient self makes decisions in each period to maximize the discounted sum of utility net of a cost of self-control.

The only article providing empirical evidence on the pure effect of time preferences is Angerer et al. (2014), who study altruism through a donation experiment based on a sample of primary school children. They let children decide in a dictator-game like framework how many experimental tokens they want to keep for themselves and how many they want to donate to a well-known charity. They find that the level of donations decreases non-linearly with impatience.[3] Though remarkable, these results emerge under a very particular setting, based on donations made by primary school children. In this respect, our work complements these findings and provides more external validity.

Against this background, we rely on quasi-experimental analysis adopting the generalized propensity score methodology (GPS) proposed by Hirano and Imbens (2004), and similarly by Imai and Van Dyk (2004), which is an extension of the propensity score methodology for binary treatments, the well-established methodology for reducing bias caused by non-random treatment assignment in observational studies. Similarly to the binary propensity score method, the GPS approach requires that after controlling for observable characteristics, any remaining differences in treatment intensities across individuals is independent of the potential outcome (i.e. the weak unconfoundedness assumption). This implies that the GPS has a balancing property, i.e. for units belonging to the same GPS strata the level of treatment examined can be considered as random. Therefore, in the same way of the standard methodology, the GPS methodology is subject to the major drawback that an unobservable factor may drive the results, hindering any type of causal claims. However, differently from the standard approach, the GPS method allows us estimating the so-called *dose-response* function, that is the (average) response in the probability of donating to changes in the level of a continuous treatment variable (i.e. the level of impatience). This approach has already been successfully used in a variety of recent observational studies (see Fryges and Wagner 2008; Becker et al. 2012; Giannetti 2019).

We apply the GPS to a very large panel dataset of adults from the Italian Survey on Household Income and Wealth (SHIW). This survey contains information about *whether* and *how much* individuals donate to any charity (our measure of altruism) along with direct measures of time (and risk) preferences, which are often unobservable outside the laboratory and hard to proxy in observational studies. On top

[2]Andreoni and Serra-Garcia (2016) also show that individual are dynamically inconsistent (they choose to give when the choice is made in advance but reverse their choices when giving occurs immediately).

[3]They also find that altruism increases with age, and that girls are more altruistic than boys.

of that, our empirical strategy takes advantage of a crucial asset of our dataset, i.e. the presence of a large set of relevant covariates, which are necessary for this estimation technique. Indeed, we can rely on a full set of economic, social and demographic variables. Importantly, we are able to replicate the main results over different waves and pools of subjects, both for the probability and the amount donated.

Our empirical results indicate that the relationship between time preferences and altruism is non-monotonic. In particular, we find that at lower levels of impatience both the probability and the amount donated to charity are rather flat or decrease slowly with impatience, but then decline more sharply as impatience grows at higher levels. Out of the effect of an observable factor that we try to address in a number of different ways in the paper, these results support the hypothesis that – in addition to tax incentives – individual parameters and attitudes do have an important impact on altruistic behaviors.

In addition, we provide a theoretical framework that allows under fairly general conditions non-monotonic effects of time preferences on altruism, consistent with those emerging from the empirical analysis. Though not aimed at rationalizing the only possible mechanism at work, the model shows that non-monotonic effects may arise even in simple discounted utility model with common and general assumptions on utility functional forms, where two typical motives for altruism are taken into account, i.e.: joy of giving (Andreoni 1989, 1990) and the so called "pure altruism" (Becker 1974), combined with the willingness to preserve own consumption above a certain level (Chatterjee and Ravikumar 1999; Alvarez-Pelaez and Diaz 2005). As described in Section 6, the joy of giving utility originates from the very act of donating, while the pure altruism utility is the beneficial reward enjoyed for the effect of donation; therefore, they logically and practically refers to different times reflecting the inherently intertemporal dimension of donation (Angerer et al. 2014). The theoretical model predicts that, at lower levels of impatience, the effect of time preference on donations depends on the distribution of the two motives among the population, thus being essentially an empirical issue. At higher levels of impatience, however, it becomes more and more likely that the effect of time preference on altruism is unambiguously negative.

As mentioned, our results complements previous findings by providing more external validity and improves upon our understanding of the relationship between time preference and donations. Such a topic has relevant implications. First, understanding the effect of time preference on donation is helpful for the design of fund-raising and tax-deduction schemes for charity organizations and policy makers (Cappellari et al. 2011; Andreoni and Serra-Garcia 2016). From a development viewpoint, it can improve our comprehension of the channels through which patience may affect societies in a persistent way (Galor and Özak 2016): one of such channels - beside the well-known ones of human and physical capital accumulation (Knack and Keefer 1997, Glaeser et al. 2002,Barsky et al. 1997) - can be social capital as this concept is strictly connected to altruism (Fehr 2009; Cox 2004).[4] Additionally, it may also prove useful to gauge insights on the broader effects of preference-impacting shocks (Becchetti et al. 2017). More in general, it contributes to the growing empirical literature studying how patience help shaping a large number of lifetime outcomes, such as health and labour earnings (e.g. Wang et al. 2016, Golsteyn et al. 2014, Sutter et al. 2013).

The paper is organized as follows. Section 2 gives details on the data and illustrates the main variables of interest. Section 3 describes the methodology to estimate a dose-response function to evaluate the effect of impatience on altruism, providing details about the common support condition and the balancing of covariates. Section 4 provides the main empirical results from the GPS estimation, while Section 5.1 contains several robustness checks. Section 6 presents our theoretical model. Section 7 concludes.

2 Data Description

We select our data from the 2008, 2010 and 2012 SHIW surveys, the years in which questions about individual discount rates were first introduced. The (biannual) survey is a large representative sample of the Italian population and covers – for all three years – questions on several aspects of the individual's life, such as education, living and working conditions, as well as information on individual's attitudes toward risk and level of impatience. This survey has been extensively used in several studies to identify micro-founded effects on a great variety of topics (see for example, Battistin et al. 2009; Jappelli and Pistaferri 2000; De Blasio and Nuzzo 2010; Bottazzi et al. 2006). The sampling unit is the household, and although the information are mainly available for the head of the household, there are also additional information both at the household level and for each component.[5]

To keep control of the composition of our sample, we restrict our attention to those heads of household who participated in the study all three years (i.e. a balanced sample), i.e. about 40% of all heads of household. However, we do not conduct a full panel data analysis for two reasons. Firstly, it is hard to control for time-varying factors with the GPS methodology and a non-linear model. Secondly, the mea-

[4] Donation has also been extensively used as a measure of social capital: e.g. blood donation in Guiso et al. (2006).

[5] For further information about the survey, see `https://www.bancaditalia.it/statistiche/tematiche/indagini-famiglie-imprese/bilanci-famiglie/index.html?com.dotmarketing.htmlpage.language=1`

surement of discount rates (our measure of impatience) slightly changed across surveys (see below). In order to check the robustness of our results over time and avoid confounding changes in impatience with changes in measurements, we replicate each year the same analysis with the same sample of individuals. Therefore, we first apply the GPS propensity score analysis to the impatience level derived from 2012 survey, using as (lagged) covariates data from 2010 survey. We then check the robustness of the results relying on impatience level from survey 2010 and (lagged) covariates from 2008 survey. Nevertheless, to further test the robustness of our results we will replicate the analysis of year 2012 also for the larger sample of individuals appearing also in 2010 survey (but not 2008).

2.1 Covariates selection

In the following sections, we first describe in detail our outcome variable (i.e. our measures of altruism) along with our measure of impatience (i.e. our "treatment measure"), while Table (1) describes and summarizes our set of control variables. It has been shown that covariates that are influenced by the treatment can cause the ignorability assumption to be violated and lead to larger biases when they are included (Wooldridge 2016; Pearl 2012; Rosenbaum 1984). In other words, they are "bad controls". Wooldridge (2016) also shows that matching on instrumental variables is also a bad idea as it leads to more asymptotic bias than excluding the instrumental variables.

Therefore, we select into our set of covariates only variable that are either unaffected by the treatment (e.g. sex) or have been measured before our variable of interest was determined, i.e. they are predetermined. In our robustness analysis, we however check our results by excluding from the set variables that may appear problematic (e.g education). In particular, we rely on lagged values of our covariates (i.e. year 2010) as these variables are predetermined with respect to the current time errors (i.e. 2012). As Table (1) shows, the data contains a large number of covariates. In particular, we use information that have been identified in previous studies as determinants of donation, such as individual and family income (see for e.g. List 2011). Moreover, we have information on individual characteristics, such as individual sex, education and age, as well as information on individual attributes, which are often unobservable characteristics, such as individual attitude towards risk (see Angerer et al. 2014; Cappellari et al. 2011). Concerning this latter variable, we both have a generic categorical measure of risk-tolerance, which is available for all years, as well as a more sophisticated continuous variable which measures the degree of risk-aversion (ranging from zero to 1), though only available in 2012. In particular, our categorical measure of risk tolerance takes on 4 values according to whether the individual is willing to accept a low-return for no risk, a fair-return for a good protection from risk, a good return for a fair protection from risk, and a high-return for high risk (see Figure 1 and Tab 1). We checked that this measure correlates well with the continous measure of risk. As this data highlights, only a small fraction of individuals (around 20%) are willing to take on risk.

INSERT TABLE (1) HERE

INSERT FIGURE (1) HERE

The richness of the set of covariates makes the application of the GPS methodology to this dataset appropriate. This methodology relies on the weak unconfoundedness assumption, which is not statistically testable, and requires that after controlling for observable characteristics, any remaining difference in treatment intensity across individuals is independent of the potential outcome of interests (see further below).

2.2 Measuring altruism

We develop our measure of altruism relying on the SHIW question on donation.[6] We can observe whether the head (or any another member) of the family has contributed with money to charitable organizations but we cannot observe whether (s)he has also contributed with time. Previous research has analysed the relationship between time and money donations finding evidence that individual unobserved characteristics, such as their altruistic attitude, drive a significant and positive relationship between time and money donation (Cappellari et al. 2011, Bauer et al. 2013). Differently from previous studies, however, we also have information on the amount of donation to charitable contribution. More specifically, the head of the household answered the following question:

"Did you or a member of the household make donations or other contributions (e.g. to non-profit associations, voluntary organizations, charities)? (If "Yes") What was the amount of the payments?

From this question we derive a dummy variable *Donation* equal to 1 if the individual answered yes, and 0 otherwise, which we will use as dependent variable in our analysis of altruism. We also derived a variable *Donation Amount* which is equal to the amount of money individuals gave for donations. The assumption is that this measure is similar to that derived in dictator games in economic experiments

[6] For a discussion about altruism as charity, see for example Khalil (2004).

(i.e. the size of donation sent as a first mover). As shown in Table (1) and Figure (2), almost one fifth of respondents has on average made a donation in 2012 (the share was basically equal in 2010). This is in line both with the share of 17% documented by Bauer et al. (2013) using international survey data and also with the average share of money contributors reported in Cappellari et al. (2011) using Italian data.[7] Conditional on having given, the amount donated is rarely above 500 Euro (see Figure 2), with the average amount donated being 322 Euros, which corresponds to 1.2% of the contributors' average income (1.7% of the overall average income).[8]

INSERT FIGURE (2) HERE

2.3 Measuring impatience

As stated above, an important feature of our study is having a direct measure of time preferences. Specifically, the level of discount rates to measure individual impatience were elicited relying to following question asked to the head of the family:

"You have won the lottery and will receive a sum equal to your household's net yearly revenue. You will receive the money in a year's time. However, if you give up part of the sum you can collect the rest of your win immediately. To obtain the money immediately would you give up 10% of your win?"

A series of alternative questions were asked depending on the answer given to that question (see Figure 3). For example, if the respondent's answer to 10% discount was "yes", the interviewer asked for a discount of 20%, whereas if the answer was "no" the interviewer asked for a discount of 4%. In this latter case, if the answer was again "no" the interviewer also asked whether the head of the family was willing to give up the money for a discount of 2%, while if the answer was "yes" the interviewer asks for a discount of 7%.[9] To determine the level of impatience in each year (i.e. the implied discount rate), we use the midpoint of the range of two discount values. Thus, for example, if a respondent was willing to accept a reduction of 10% but did not accept a discount of 20%, the midpoint is 15%. Moreover, if the respondent provided the same reply in the survey, i.e. was always willing to accept a discount or to reject one, we set the discount rate at the two limits, i.e. 30% or 0%. These boundary cases are

[7] Cappellari et al. (2011) using the year 2000 wave of the *Indagine Multiscopo sulle famiglie - Aspetti della vita quotidiana* by the Italian National Statistical Office (ISTAT) report an average of 19.3% for women and 21.7% for men.

[8] For the US, Andreoni (2006) documents an average share ranging between 1.5% and 2.1% between 1968 and 2001.

[9] It must be noticed that time preference is measured at the individual level, while questions about donations are asked at the household level. This could potentially create a problem if time preference of singles are systematically different from individual living in a household. We check the for this issue in Section 5.1.

specifically accounted for in our model specification of the generalized propensity score (i.e. we rely on a two-limit Tobit). In 2010 this question was slightly different, allowing for a bit less of variability in discount rates,i.e. there were only 5 categories instead of 8 (with a range between zero up to 20%), while in 2008 this question was also administered to a random subset of households.

INSERT FIGURE (3) and (4) HERE

Overall, the distributions of time preferences across years are quite similar (see Figure 4). In 2012 about 35% of the respondents had a discount rate above 8.5%, and about 7% had a discount rate equal and above 30% (i.e. the upper limit). A large fraction of respondents (about 35%) would not be willing to accept any discount rate. In 2010 about 30% of the respondents had a discount rate above 7.5%, and about 20% had a discount rate equal and above 20% (i.e. the upper limit), and about 30% of the participants would not be willing to accept any discount rate. As highlighted above, to avoid to account for additional time-varying confounding factors inborn in panel-data analysis, as well as to avoid changes in measurement in discount rates, in the following we focus on data derived from surveys 2012, although checking the robustness of our results in year 2010.

Several techniques are available to elicit individual discount rates and there is no consensus on best practices and methodologies (Hardisty et al. 2013, Cohen et al. 2020). Eliciting individual discount rates as in the SHIW surveys (i.e. through multiple staircase choice-method) has advantages with respect to previous methods because it avoids answer inconsistencies and appears easier for participants to understand (Hardisty et al. 2013). Differently from economic experiments, individual choices are not incentivized. While at first glance, this may appear a drawback, this type of measure is less subject to the concern that respondent's trust in future payments will affect the response. Nevertheless, Falk et al. (2014) have shown that these measures correlate well with experimental measures when the number of survey items increase (see also Cohen et al. 2020). Finally, in our case there is no need to infer ex-post the discount rates (and thus imposing a functional form for preferences) as participants are directly asked the value of their discount rates (i.e. the percentage).

3 Model Specification

In the absence of experimental data, matching methods provide an appealing alternative. The main feature that makes matching methods such an attractive empirical tool is the possibility of mimicking an experiment *ex-post*. In particular, Rosenbaum and Rubin (1983) showed that conditioning on the propensity score is sufficient to balance treatment and comparison groups. Subsequently, the literature has extended the propensity score methods to the cases of multivalued treatments (Imbens 2000, Lechner 2001) and, more recently, to continuous treatments (Hirano and Imbens 2004 ;Imai and Van Dyk 2004).

The approach developed by Hirano and Imbens (2004) is particularly suited for our paper. They proposed estimating an entire "*dose-response function*" of a continuous treatment, i.e. a relationship between the exposure to a continuous treatment and an outcome variable. In the following we briefly recall this approach. Readers already familiar with this procedure, may directly move to the next section.

We define a set of potential outcome $\{Y_i(t)\}$ for $t \epsilon T$, where T represents the continuous set of potential treatments (in this paper, the level of impatience) defined over the interval $[t_0, t_1]$, and $Y_i(t)$ is referred to as the unit-level dose-response function (in the paper, the probability of donating). For each individual $i = 1, ..., N$, we observe a $k \times 1$ vector of covariates, X_i; the level of treatment delivered, T_i; the corresponding outcome $Y_i = Y_i(T_i)$.

Hirano and Imbens (2004) define *weak unconfoundedness* for continuous treatments as

$$Y(t) \perp T \,|X \quad for\, all\, T \tag{1}$$

which is a generalization of the concept of unconfoundedness for binary treatments. That is, individuals are different in terms of their characteristics X: some are more/less likely than others to have higher discount rates (i.e. level of impatience). Weak unconfoundedness implies that, once it has been controlled for observable characteristics X, any remaining difference in the level of impatience (i.e. T) across individuals is independent of the potential donation (i.e. $Y(t)$).

The generalized propensity score is defined as

$$R = r(T\,|X) \tag{2}$$

where $r(t, x) = f_{T|X}(t\,|x)$ is the conditional density of the treatment given the covariates. The generalized propensity score is assumed to have a balancing property (likewise the propensity score for binary treatments), i.e. the probability that $T = t$ does not depend on the value of X within strata $r(t, x)$. In other words, if we look at two individuals with the same *ex-ante* probability of having a particular level of impatience (i.e. discount rate), once we condition on observable characteristics X, their actual level of impatience is independent of X. The propensity score summarizes thus all the information in a multi-dimensional vector X so that

$$X \perp 1\{T = t\} \,|r(t, X) \tag{3}$$

The balancing property implies, along with the weak unconfoundedness, that assignment to treatment is *weakly uncounfounded given the generalized propensity score.* Then, for every t

$$f_T(t|r(t,X), Y(t)) = f_T(t|r(t,x)) \tag{4}$$

The GPS can thus be used to eliminate any bias associated with differences in the covariates by first estimating the conditional expectation of the outcome (i.e. in our case the probability of donating) as a function of the treatment level T (i.e. the discount rates) and the generalized propensity score R (i.e. $\beta(t,r) = E[y\,|T = t,\ R = r]$). Then, the dose-response function is estimated by averaging – at a particular level of the treatment intensity – the estimated conditional expectation over the generalized propensity score (i.e. $\mu(t) = E[\beta(t, r(t, X))]$).

3.1 Practical implementation

1) Estimation of the propensity score. Hirano and Imbens (2004) specified the treatment intensity relying on a normal distribution. However, in this context, our treatment variable T, i.e. the level of impatience, is a fractional variable bounded between [0, 0.3]. Therefore, we cannot specify it as a normal distribution. Moreover, we cannot even resort to a fractional logit regression as in Papke and Wooldridge (1996). By doing so, we would not be able in the subsequent steps to fully specify the density function of the treatment estimates. Likewise, by having several observations at limits, we cannot assume that the treatment intensity follows a Beta distribution as in Bia et al. (2014). We therefore specify for the treatment intensity the following two-limit Tobit Model:

$$T_i|X_i \sim \Phi(L_1 - x_i\beta/\sigma)^{d_0} \cdot 1/\sigma\phi(y_i - x_i\beta/\sigma)^{d_1} \cdot \Phi(L_2 - x_i\beta/\sigma)^{d_2}$$

where $\Phi(.)$ is the standard normal cumulative distribution function, $\phi(.)$ is the standard normal probability density function, σ is the standard deviation, x_i is a row vector of covariates and β_1 a column vector. L_1 and L_2 are the lower and upper limits of the censored distribution (in our case 0 and 0.3 respectively). The observed values within the limits are denoted with y. For each observation, depending upon whether the observed value is either equal to or within the two limits, only one of the exponents d_j $(j = 0, 1, 2)$ assumes the value of one.[10]

2) Common support condition and balancing of covariate. We test the common support condition as follows. First, we split the sample into three groups $j = 1, 2, 3$, which are defined according to the distribution of the impatience level (i.e. discount ratio). We then calculate the median treatment intensity T_{Mj} for each treatment group j, and evaluate the GPS for the whole sample at median treatment intensities using the estimates for β and σ derived from the estimation of the propensity score. For each group and each observation $i = 1, .., N$ we calculate $\hat{R}(T_{Mj}, X_i)$. We then divide into three blocks the GPS we obtained. We test the common-support condition by plotting the GPS values $\hat{R}(T_{Mj}, X_i)$ for each block against the distribution of the GPS (i.e. $\hat{R}(T_{Mj}, X_i)$) for the rest of the sample. We then drop those observations that will lye outside the common support.

As in Hirano and Imbens (2004), we also test the balancing property applying the blocking approach As above, we again divide the sample into three groups according to the distribution of the impatience level (i.e. discount ratio). Then, we evaluate the GPS at the median values of the treatment variable (i.e. discount ratio) within each group. We then split each group into five blocks according to the quintiles of the GPS evaluated at the median level.[11] We compare the mean difference of each covariates within each of these blocks, with respect to individuals who have a GPS such that they belong to that block (i.e. the same *predicted* treatment intensity) with those who are in the same block, but have a different actual treatment intensity (i.e. groups). In other words, we assign each individual to the respective GPS block - as evaluated at the median level - and we compare the means of covariates of those individuals who belong to a different treatment level but have similar GPS (i.e. in the control group).

3) Estimate the conditional expectation of the outcome. Once we obtain the GPS estimates of the first stage $\hat{R}$, we estimate the conditional expectation of the outcome Y_i as a flexible function:

$$E\{(Y_i \,|\, T_i,\ \hat{R}_i)\} = h(\alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \alpha_3 T_i^3 + \alpha_4 \hat{R}_i + \alpha_5 \hat{R}_i^2 + \alpha_6 \hat{R}_i^3 + \alpha_7 T_i \hat{R}_i)$$

of the GPS terms (which control for selection into treatment intensities) and the observed treatment intensities T_i. We estimate these parameters relying on a Probit model. It is important to stress that the estimated coefficients have not direct causal interpretation but only suggest whether the covariates introduce any bias (Hirano and Imbens 2004). If the estimated coefficients of the GPS terms are equal to zero the GPS is not relevant to reduce any bias. On the contrary, if the GPS terms are (jointly) significant,

[10]For a similar issue and specification, see Giannetti (2019).

[11]We groups are defined according to the following cutpoints: 0-0.035;0.035-0.125; 0.125-0.3. Choosing a finer or coarser specification does not change significantly the results.

their introduction significantly reduces the bias of the estimated response of the probability of donating to changes in individual level of impatience.

4) Obtain the dose-response function. Finally, we average the estimated regression function over the score function by evaluating it at the desired level of the treatment. Given the estimated parameters $\hat{\alpha}$, the observed level of impatience T_i and the estimated GPS $\hat{R}$, the *average* potential outcome (i.e. the average probability) is obtained as

$$E\{Y\hat{(t)}\} = \frac{1}{N}\sum_{i=1}^{N} h(t, \hat{R}_i^t, \hat{\alpha}) = \frac{1}{N}\sum_{i=1}^{N} h(\hat{\alpha}_0 + \hat{\alpha}_1 T_i + \hat{\alpha}_2 T_i^2 + \hat{\alpha}_3 T_i^3 + \hat{\alpha}_4 \hat{r(t, X_i)} + \hat{\alpha_5} \hat{r(t, X_i)}^2 + \hat{\alpha}_6 \hat{r(t, X_i)}^3 + \hat{\alpha}_7 T_i \hat{r(t, X_i)})$$

We obtain the entire dose-response function by estimating for each level of the treatment this average potential outcome. As result, for each individual we evaluate the GPS for each level of the treatment, thus having as many propensity scores as there are levels of treatment. We then obtain the average response by averaging over all the individual responses at each level of the treatment. To take into account estimation of the GPS and the $\hat{\alpha}$-parameters, we use bootstrap methods to obtain the standard errors. We finally display the derivative of the dose–response function with respect to impatience level —which is commonly referred to as the treatment-effect function. This latter is the object of our interest as it has a causal interpretation.

4 Results

Before presenting the main results of the GPS methodology, we first explore the relationship between the individual parameters and the probability of donating in year 2012 by using a probit model. In Table (2), we start with a basic model including only impatience, to then include their squared and cubic interactions. We report regression results for both coefficients and marginal effects.

INSERT TABLE (2) HERE

INSERT FIGURE (5) HERE

In column *a*, we observe that impatience significantly decreases the probability of donating (-24%). In column *b*, we include the squared, while in column *c* the cubic term for impatience. However, we refrain from interpreting the coefficient on the interaction terms as tests about partial effects and interaction terms are not necessarily informative in non-linear model, and it can also be economically misleading (see Greene 2010). We therefore directly report the overall marginal effects for each variable. The results now suggest that impatience positively (although not significantly) affects the probability of donating (4.6%). In columns *c*, results are substantially analogous (24%). While – at a first look – they might appear contradictory, these results highlight non-monotonic relationships between impatience and the probability of donating, as highlighted in Figure (5). We therefore turn to a more useful analysis of the dose-response function.

In Table (3) we report the results of step 1 that estimates the propensity score relying on the Tobit specification. These results suggest that better educated, risk tolerant, and richer individuals are less impatient. They also suggest that older and employed individuals are less patient.

INSERT TABLE (3) HERE

In Figure (6) and Table (4) we report the results related to step 2 testing the common support and balancing property. For example, in Figure (6) - panel *a* - we plot the distribution of the GPS for group 1 (see the black bars) against the distribution of the GPS for the rest of the sample, i.e. group 2 and 3 (see the white bars). Similarly for group 2 and 3 (see Figure 6, panel *b* and *c*). By inspecting the overlap of these distributions, we find that there are 27 participants whose GPS is not among the common regions of the three groups. We thus impose the common support by dropping those participants (less than 2% of our sample), for a total of 2948 observations in 2012.

INSERT FIGURE (6) HERE

INSERT TABLE (4) HERE

Table (5) illustrates the group and block structure. For instance, we compare the covariates of 292 observations in group 1/block 1 to the observations in control 1/ block 1. Taking the sum over all blocks and adding the respective control groups yields the total number of observations (i.e. 2955) in the common support regions. If adjustment for the GPS properly balances the covariates, we would expect all differences not to be statistically significant. Table (4) reports the mean *t-statistics* for each group across all covariates. There is evidence that the balancing property is satisfied, with only 1 out 33 t-values significant after controlling for the GPS. We thus conclude that the estimated generalized propensity score perform well in reducing potential treatment bias.

INSERT TABLE (5) HERE

In Table (6) we report estimation of the dose response function from step 3. As highlighted above, the estimated coefficients have no direct causal interpretation. However, we note that the coefficients of the GPS are highly significant and different from zero, suggesting that the GPS procedure allows to remove potential bias introduced by the covariates.

INSERT TABLE (6) HERE

Our main results from step 4 are presented in Figure (7), where the left panel indicates that there is a non-monotonic relationship between the level of impatience and the probability of donating. In particular, the probability of donating is initially rather flat or mildly decreasing, to then decline more sharply at higher levels of impatience. In other terms, the marginal increase in the level of impatience significantly affects the probability of donating only at medium-high level of impatience. This can be seen from the derivative of the dose-response function with respect to impatience level in the right panel of Figure (7). The 90% confidence band of the treatment effect function always excludes the zero for level of the discount ratio between 19% and 27%. Within this range, an increase in the individual impatience will significantly reduce the probability of donating (i.e. the variation in the probability of donating is negative) around 2%. This effect is also economically significant, if we consider that the sample probability of donating is 19% (see Table 1), corresponding to a reduction of about 10%. For values below this range the effects are any longer significant and unambiguous.[12]

INSERT FIGURE (7) HERE

The dose-response function for the amount donated in 2012 (see Figure 8) exhibits a similar shape as for the probability of donating. Importantly, consistent with the analysis of the probability of donating, the marginal increase in the level of impatience (i.e. the pair-effect function depicted in the treatment-effect function) significantly affects the amount donated only at medium-high level of impatience, approximately between 22%-27%. Within this range, the variation in the amount donated is negative. That is, an increase in the level of impatience will reduce the amount donated up to 25 Euro. This effect is economically significant, considering that the sample average (conditional) amount of a donation is

[12] At the very extreme values of the (empirically) observed range of impatience the effects are somewhat irregular, mostly likely because at these corner values we have very different type of observations.

about 322 Euro (see Table 1), resulting in a reduction of about 8%. Below this range the effects of higher impatience tend to be small and scarcely insignificant.

INSERT FIGURE (8) HERE

5 Robustness checks

It is fair to recall that the weak uncounfoundness assumption cannot be tested. As stated above, the major drawback of our empirical strategy is the existence of an observable factor that may drive both the probability to donate as well as individual impatience. To generate a bias, however, such a factor should not only affect donations but should also vary with the intensity of impatience differently in the treatment and in the control groups. Although it would not be possible to completely rule the existence of this factor, in the following we will try to assess the robustness of our results to a number of model specifications, sample selection, as well as measurement issue. Taking into account this empirical evidence, we then move in Section 6 to the theoretical part of our analysis, where we develop a simple model that could generate a non-monotonic relationship between impatience and donation.

5.1 Unobservable factors and individual trust

We now check whether the results appear robust over different years and over a different sample of individuals.

To begin with, we replicate all the same steps of the previous analysis but using, as outcome variables, the probability of donating in 2010 (but the same pool of subjects, see Section 2 for a description of the sample selection). For the sake of brevity, we now only report the main results from step 4 and for the probability of donating in Figure (9). As Figure (9) highlights, even though in 2010 the question about discount rates does not allow for the same level of variability (see Section 2.2), the results are consistent with the previous analysis: the 90% confidence band of the treatment effect function always excludes the zero for medium-high level of discount ratio, i.e. between 14% and 19%. Within this range, a marginal increase in the individual impatience will significantly reduce the probability of donating. Below this range, the effects are any longer significant and unambiguous.

INSERT FIGURE (9) HERE

Although reassuring, these results however cannot tell us anything about the role of an unobservable factor, as we are relying on the same sample of subjects. For instance, time preference may reflect "trust" rather than pure time preference. As stated above (see Section 2.2), our measures of time-preferences are hypothetical and therefore less subject to the coufonding effect of trust in future payments. However, as any other studies of altruism based on donation to charitable organizations, we indeed suffer from the possibility that individuals may differently trusts the organizations. In particular, individuals with very high rates of impatience and lower donation rate may have a greater distrust than individuals with lower levels of impatience, thus resulting in a lower probability to donate.

Therefore, we replicate the same analysis but this time relying on the entire pool of subjects available only in 2012 and 2010 (i.e. thus including also those individuals we were not able to track back in time in 2008). Even this case (about 4600 individuals), we find exactly the same type of relationship between individual impatience and time discounting (results available upon request). Although this result is not a direct test of the existence of an observable, it is good to find it out that a different pool of subjects (with possibly different degree of variation in trust) exactly exhibit the same shape of the dose-response function over all levels of impatience. To further account for other confounders like trust, we additionally include among the set of regressors a set of regional dummies (results available upon request and at author's webpage). Even in this case, the results are identical. Nevertheless, we acknowledge that we cannot entirely rule out the existence of such an unobservable factor.

5.2 Measurement issues and bad controls

As stated above, time preferences are measured at the individual level, while the questions about donations are asked at the household level (i.e. "Did you or a member of the household .."). This measurement issue could potentially create a problem if time preference of singles are systematically different from individuals living in a household. The first thing to notice is that the propensity score already accounts for it, as within our set of covariates the family status is included among the set of covariates. Nevertheless, to further check for this issue, we first perform a Kolmogorov test: it rejects the hypothesis that both time preferences and donations are different by family status. Furthermore, we repeat the above analysis by restricting our sample to family with only one member as income holder. The results are even this case substantially identical (available upon request). On top of that, we repeat the same analysis but using different dependent variables: the amount of the donation per income holder, as well as the amount of the donation as an income share (see Meer and Priday 2020). In both cases, results are economically analogous (available upon request and and at author's webpage), although we lose a bit of statistical significance in the case of income share.

Another issue is related to some of our controls, such as education, income and risk preferences, as they might be simultaneously affected by individual time preferences. For example, several studies have shown how impatience can be an important determinant of education (see for example Sutter et al. 2013). Therefore, even though these controls are predetermined in our regressions (as they are measured in 2010), we check the robustness of our results by repeating the entire analysis and excluding each of them. Even in the case, results are analogous (available upon request and at author's webpage).

6 The theoretical setting

In order to assess the relationship between time preference and donation, we develop an intertemporal model with two periods. The model shows how a non-monotonic relationship between time preference and donations may arise in a relatively simple discounted utility model featuring two forms of altruism extensively considered in the literature and the willingness to preserve own consumption above a certain level.

Namely, we consider that altruistic behaviour can increase the donor's utility through two channels, well-known in literature (Becker 1974; Andreoni 1990): the "good feeling" coming from the very act of donating (the so-called "joy of giving" or "warm-glow" channel), and the reward associated to the beneficial effect that the donation will bring to the receiver (the so-called "pure altruism" channel).

We assume that the former brings utility as soon as donation is made (i.e.: at the same time as current consumption), because it is inherently connected with the act of giving. The latter channel is instead assumed to unfold its effect later, in the second period of our model, together with future consumption; this is consistent with channels of altruistic behaviour operating through affinity and identification (Wade-Benzoni and Tost 2009; Cialdini et al. 1997; Batson 1995; Aron et al. 1991). Since the pure altruism effect is delayed, it is related to the concept of patience (mentioned in the introduction) as the propensity to postpone gratification in exchange of delayed rewards. This intertemporal intrinsic nature of altruism cannot be neglected as brain regions associated with the ability of projecting oneself into the future are also connected with the attitude of projecting oneself into the perspective of others (Yi et al. 2011; Buckner and Carroll 2007). From a modelling viewpoint, assuming that each motive is associated only to one period simplifies the algebra but it does not drive per se the qualitative results, which would remain valid even if both motives' effects occurs in both periods as long as warm glow is relatively more relevant in the present and pure altruism is relatively more relevant in the future.[13]

We also assume that a minimum level of own consumption (either present or future) has to be preserved (Chatterjee and Ravikumar 1999; Alvarez-Pelaez and Diaz 2005; Rosenzweig and Wolpin 1993). In our model minimum consumption has not to be strictly interpreted as a subsistence level but it broadly refers to the willingness to keep own consumption above a minimum target, which may vary according to different socio-economic contexts. This implies that donating is possible only provided that at least a certain amount is left for consumption.[14]

We formalize what described above with the agent's objective function:

$$U = (1-\alpha)u(c) + \alpha\gamma j(a) + \frac{1}{1+\rho}[(1-\alpha)u(d) + \alpha\phi v(a)] \tag{5}$$

where c is consumption in the first period, a is the donation amount, $\rho \geq 0$ is the intertemporal discount rate,[15] d is consumption in the second period. The utility from consumption, $u(\cdot)$, is strictly increasing and concave in its argument.

The function $j(a)$ refers to the "joy of giving" effect; it is associated with the parameter $\gamma \in (0,1)$ capturing the weight of this altruistic motive. The function $v(a)$ is the "pure altruism" effect and its weight is captured by the parameter $\phi \in (0,1)$. The functions $j(\cdot)$ and $v(\cdot)$ are assumed strictly increasing and concave in their arguments: this warrants that altruism cannot be detrimental but also cannot improve utility at increasing rates. We also assume that $j(\cdot)$ and $v(\cdot)$ are well defined in zero, allowing for the possibility of no donation.[16] As implicitly shown in Eq. (5), we assume that utility from own consumption and donation is additively separable.[17]

The parameter $\alpha \in [0,1/2]$ is associated to the importance of altruism relative to selfishness. It can

[13]In other words, the model is not suited only for the extreme case where all the pure altruism takes place instantaneously, and in such cases other theoretical frameworks should be considered. However, such situations are quite extreme, and the model remains valid for the vast majority of cases.

[14]It is important to notice that in our empirical analysis we do not include the consumption level but we are able to include family income among our set of regressors in the generalized propensity score.

[15]We model the intertemporal utility as in the standard discounted utility model *à* la Samuelson (1937). For a critical review on models of time discounting see Frederick et al. (2002). We assume the same discount factor for both private consumption and donations. Dreber et al. (2016) assume that individuals are more impatient towards donations than towards private consumption, while Ebert and van de Kuilen (2015) suggest that discount rates can potentially vary by domain. In our model assuming higher discount rate on donation would have results similar to a decrease in the parameter ϕ. The general results would remain valid, but the parameter range where the relationship between discounting and donation is negative is reduced.

[16]We rule out Inada type conditions $\lim_{a\to 0} v'(a) = +\infty$ that would prevent a solution with no donation. A simple example of a functional form satisfying our assumption is $j(x) = \ln(1+x)$

[17]In order to keep things simple, we abstract from idiosyncratic shocks and uncertainty; nonetheless, introducing them in presence of risk averse individuals would reinforce our qualitative result making more tightening the constraint effect on donation.

be broadly interpreted as an inverse measure of interpersonal distance, related to the concept of social discounting (Rachlin and Jones 2006; Wade-Benzoni 2008): basically, if others are less important, it is as if altruism towards them is more heavily discounted and hence valued less (α goes down) in the utility function. The parameter α concerns both types of altruism, as also the joy of giving is higher when the receiver is perceived as closer. We constrain α to be less than a half to exclude the unlikely case that donation matters more than own consumption; however, such a restriction can be removed without affecting the results.

The maximization of Eq. (5) occurs under constraints: the first one is the budget constrain stating that the income $y > 0$, which the individual is endowed with in the first period, has to be shared among consumption c, donation a and savings s. The savings, increased by the exogenous interest rate $r \geq 0$, are used to fund the second period consumption d:

$$y = c + a + s \tag{6}$$

$$d = s(1+r) \tag{7}$$

Moreover, a non negativity constraint holds for donation, i.e.: $a \geq 0$. As far as consumption is concerned, we assume that a stricter requirement holds: a minimum level of consumption $\underline{k} > 0$ has to be satisfied, i.e.: $c \geq \underline{k}$ and $d \geq \underline{k}$.[18]

We focus on how the maximization problem's solution (c^*, d^*, a^*) varies with ρ. First, we consider the solution when the consumption constraint is not binding: it turns out that this is the case for a "sufficiently" low level of ρ : $(\rho \leq \bar{\rho})$. Then we consider the maximizing solution under the case $\rho > \bar{\rho}$ and assess the role of increasing impatience in this case as well.

Case a: minimum consumption constraint is not binding

The optimization of Eq. (5) at interior solutions has to satisfy the first order conditions:

$$\frac{\partial U}{\partial c} \quad : \quad u'(c) = \frac{1}{1+\rho} u'(d)(1+r) \tag{8}$$

$$\frac{\partial U}{\partial a} \quad : \quad \alpha\gamma j'(a) + \frac{1}{1+\rho} \alpha\phi v'(a) = \frac{1}{1+\rho}(1-\alpha) u'(d)(1+r) \tag{9}$$

where we have taken into account the budget constraint. The first order conditions hold with equality

[18]We assume that income is high enough to ensure that both minimum consumption requirements can be met, i.e.: $y > \underline{k}(2+r)/(1+r)$, otherwise the problem would't be worth studying.

only provided that $c > \underline{k}$ and $d > \underline{k}$. In order to gain in algebraical tractability, we assume that $j(\cdot) = v(\cdot)$. Hence, we let the "joy of giving" and "pure altruism" effects differ only with respect to their intertemporal implications and their weights (γ and ϕ) but not to the functional form: this allows to make results unconfounded by any difference in the functional forms.

Concerning the role of impatience, as usual (and formally shown in Appendix), current consumption c^* turns out to be increasing in the discount rate ρ whilst the opposite holds for future consumption d^*. As regards donation, we have the following Lemma:

Lemma 1. *When $c^* > \underline{k}$ and $d^* > \underline{k}$, the effect of the discount rate ρ on donation amount a^* is ambiguous and its sign depends on a comparison between the weights of the altruistic motives, as follows:*

$$\frac{\partial a^*}{\partial \rho} \leq 0 \qquad \Leftrightarrow \qquad \gamma \leq \Theta\phi \tag{10}$$

with $\Theta \equiv \left(\frac{1+r}{1+\rho}\right)^2 \frac{u''(d^*)}{u''(c^*)} > 0$.
In case of CRRA-type utility function, i.e.: $u(x) = \frac{x^{1-\theta}}{1-\theta}$ with $\theta > 0$, the condition becomes

$$\frac{\partial a^*}{\partial \rho} \leq 0 \qquad \Leftrightarrow \qquad \gamma \leq \left(\frac{1+r}{1+\rho}\right)^{\frac{\theta-1}{\theta}} \phi \tag{11}$$

when the utility function is logarithmic, i.e.: $u(x) = \ln(x)$, $\Theta = 1$ and the condition boils down into

$$\frac{\partial a^*}{\partial \rho} \leq 0 \qquad \Leftrightarrow \qquad \gamma \leq \phi \tag{12}$$

Proof. See Appendix A.1 □

Lemma 1 expresses the effect of impatience on donation as a relationship on the relative importance of the two altruistic motives: if pure altruism is sufficiently important compared to joy of giving, then donation decreases with impatience. The intuition behind the result is as follows: a lower impatience implies that the future is discounted less, so it also implies that a higher weight is attached to the second period in the intertemporal utility. The second period is when the pure altruism effect occurs: if it is sufficiently high compared to the effect of the joy of giving, then giving more importance to the second period implies more donations. Conversely, if it is joy of giving that matters more than pure altruism, then a lower impatience, by increasing the weight of the future, attenuates the effect of the most important channel of altruism for utility (i.e.: joy of giving in this case) and so it leads to lower donations.

Conditions (10) and (11) show that under pretty general assumptions the comparison of γ and ϕ is weighted by a factor Θ, that depends on the level of time preference ρ with respect to real interest rate r and on the curvature of consumption utility, which can be related to preferences for consumption smoothing and consumption intertemporal elasticity of substitution. These parameters affects quantitatively the threshold for the comparison between γ and ϕ but, unless under a very limit case, do not rule out the qualitative result.

In particular, under the CRRA case, the parameter $\theta > 0$ represents the preference for consumption smoothing and the inverse of consumption intertemporal elasticity of substitution. When θ is very high, e.g.: $\theta \to +\infty$ in the limit case of a Leontief function with zero intertemporal elasticity of substitution, we have that $\Theta \to \frac{1+r}{1+\rho}$, which is above or below 1 according to whether ρ is below or above r. However for realistic values of both of them this implies just shifts of Θ around 1; when $\rho \approx r$, then $\Theta \approx 1$ irrespective of θ. Allowing for a higher intertemporal elasticity of substitution and letting θ decrease toward 1, Θ becomes closer to unity (from above or below according to whether ρ is below or above r) and it eventually equals 1 when $\theta \to 1$, which coincides with the case of logarithmic utility (Condition 12). With even higher values of intertemporal elasticity of substitution ($0 < \theta < 1$), the preference for consumption smoothing is lower, the curvature of the utility function less pronounced and the agents are more reactive to differences in ρ and r: Θ is more sensitive further changes in θ, however unless the level of θ is extremely low, the shift in Θ at most changes the threshold for the comparison between the two altruistic motives (and the less so as long as ρ and r are similar) but do not rule out the main result of Lemma 1, i.e.: the possibility of divergent effects of time preferences on donation according to the relative importance of the altruistic motive. Only approaching the limit case of a linear utility function ($\theta \to 0$), the intertemporal elasticity of substitution is infinite and $\Theta \to -\infty$ or $+\infty$ according to $\rho < r$ or $\rho > r$, making Condition (11) trivial, i.e. never or always verified.

A meta-analysis by Havranek et al. (2015) of cross country intertemporal elasticity of substitution based on 169 published studies suggests - though with sensible variation cross countries - a mean value of intertemporal elasticity of substitution of 0.5, which would correspond to $\theta = 2$ and $\Theta = \sqrt{\frac{1+r}{1+\rho}}$. For Italy Havranek et al. (2015) estimates a lower mean intertemporal elasticity of substitution, at 0.29, corresponding to $\theta \approx 3.4$. With such a value of θ even a ρ to r ratio of 10 or $1/10$ would imply a shift of Θ below or above 1 by less than 0.07.

Hence, to summarize, whenever the minimum consumption level requirement is not binding, i.e. whenever $\rho < \bar{\rho}$, the effect of time preference on altruism is ambiguous: it is basically an empirical issue depending on the distribution of the altruistic motives among the population.

Case b: minimum consumption constraint is binding

As impatience goes up, consumption in the second period goes down monotonically (see Appendix A.1). Hence, for sufficiently high $\rho > \bar{\rho}$,[19] the second period consumption reaches the minimum level requirement: $d^* = \underline{k}$.

The objective function becomes:

$$(1-\alpha)u(c) + \alpha\gamma v'(a) + \frac{1}{1+\rho}\left[(1-\alpha)u(\underline{k} + \alpha\phi v(a))\right]$$

By taking into account the budget constraint: $c = y - a - \underline{k}/(1+r)$, the first order condition can be derived only with respect to a. We get:

$$-(1-\alpha)u'(c) + \alpha v'(a)\left[\frac{\gamma(1+\rho)+\phi}{1+\rho}\right] = 0 \tag{13}$$

The following Lemma holds:

Lemma 2. *When $\rho > \bar{\rho}$ so that $d = \underline{k}$, then donation is unambiguously (weakly) decreasing in impatience: $\partial a^*/\partial\rho \leq 0$.*

Proof. See Appendix A.2 □

Intuitively, since agents want to ensure themselves a sufficient level of consumption in the second period, when they are more impatient they want to increase current consumption but cannot cut too much on second future consumption, so they cut on altruism and donation. Hence, when impatience is high, its effect on donation turns out to be negative.

The overall effect of time preference on donation

Summing up the results in Lemma 1 and Lemma 2, we have the following Proposition:

Proposition 1. *The effect of time preference on donation can be non-monotonic. For $\rho \leq \bar{\rho}$ the effect is ambiguous as it may be positive, negative or nihil depending on the relative importance of altruistic motives (joy of giving, pure altruism) and possibly other preference parameters (i.e.: consumption intertemporal elasticity of substitution). For higher impatience, i.e.: $\rho > \bar{\rho}$, the effect of time preference on donation is negative.*

[19] In Appendix A.2 we characterize the threshold level $\bar{\rho}$ and show how it varies with model parameters.

$$\frac{\partial a^*}{\partial \rho} = \begin{cases} \text{for } \rho \leq \bar{\rho}: & \begin{cases} \frac{\partial a^*}{\partial \rho} \leq 0 & \text{if } \gamma \leq \Theta\phi \\ \frac{\partial a^*}{\partial \rho} > 0 & \text{if } \gamma > \Theta\phi \end{cases} \quad \text{with } \Theta > 0 \text{ as shown in Lemma 1} \\ \text{for } \rho > \bar{\rho}: & \frac{\partial a^*}{\partial \rho} \leq 0 \end{cases}$$

Proof. Proposition 1 follows from joining results from Lemma 1 and Lemma 2. □

Proposition 1 states that while the relationship between time preference and donation below a certain threshold of the discount rate depends on the comparison between the joy of giving and the pure altruism motives, for sufficiently high discount rates (i.e.: above that threshold) time preference and donations are unambiguously in a negative relationship.

Bringing the model to data, the implication of Proposition 1 is that when discounting is not too high its effect on donation may be ambiguous as it depends on the distribution of preferences among the population. As impatience is sufficiently high and grows further, however, it can be expected that its effect on donation is most likely negative because of the greater likelihood that for someone it is above the threshold, with the associated negative effect stated in Proposition 1. These implications are in line with findings from the empirical part which shows a clear negative relationship for sufficiently high discount rates.

It is fair to say that other theoretical channels may be give rise to the same empirical pattern, hence this model is not claimed to provide the only possible mechanism. Rather, it shows how a rationale for non-monotonic effects of time preference on donations may emerge even in a standard discounted utility model under reasonable assumptions.

7 Discussion and Conclusions

In this paper we contribute to the literature on the determinants of altruism by studying the role that individual impatience plays in individual choice to donate (if and how much) to charitable organizations.

Specifically, we relied on a quasi-experimental setting to estimate the (average) response in the probability of donating to changes in the level of individual impatience. Specifically, we relied on the generalized propensity score methodology to estimate a continuous dose-response function (Hirano and Imbens 2004; Imai and Van Dyk 2004). To perform our analysis we relied on the SHIW panel dataset of Italian households in 2008, 2010 and 2012. The use of this dataset is appropriate because it contains information on impatience, which is often an unobservable characteristic outside the laboratory, along with a full set of economic, social and demographic variables, which are necessary to plausibly implement this method. To account for measurement issues and to check the robustness of our results, we repeat the analysis separately over different years without conducting a full panel analysis.

Our first result suggests that below a certain level of impatience (less than 19%) the *probability* of donating is mildly decreasing or rather flat as impatience grows; instead, at higher level of impatience (between 19%-27%), an increase in the levels of impatience significantly and more sharply decreases the probability of donating. Moreover, our second result highlights that below a certain level of impatience (less than 22%) the responsiveness of the *amount* donated to an increase in the level of impatience is rather flat, but at higher level of impatience (22%-27%) an increase in the level of impatience reduces the amount donated. In this latter case, the (average) amount donated can decrease of about 8% even for a smaller increase in the level of impatience. Both these results are robust when we repeat the analysis relying on different years and pools of subjects.

We then derived a theoretical framework assessing how individual time preferences may affect donation. We build a simple intertemporal model where we have considered that donating may increase utility both through the concurrent channel coming from the joy of giving and through the delayed channel arising from the pure altruism; in addition, we have taken into account the willingness to preserve own consumption above a certain level. Thus, in this model what matter is the temporal distance between the benefits that accrue to the receiver, and the benefits accruing to the donor at the time of the donation. The model results suggest that at lower levels of impatience the effect of time preference on donation may be ambiguous, but then it is negative at higher discount rates. These results would remain valid even if we had assumed that the effects of both motives would occur in both periods (as long as the joy of giving motive is relatively more relevant in the present), and different discount rates apply to future consumption (Dreber et al. 2016; Ebert and van de Kuilen 2015).

While the empirical results are robust across different years and pool of individuals, a causal interpretation of the results need caution as the existence of an unobservable factor might render spurious the effect of impatience. For example, it may be the case that individuals with very high discount rates may also be the one with lower trust on charitable organizations. Although it is not possible to completely rule out this unobservable factor, we believe that our results are nevertheless informative for two reasons. On the one hand, our results remain always stable when changing a variety of parameters, i.e. the year of the analysis, the pool of subjects as well as the dependent variable. In addition, our empirical investigation controlled for many relevant individual characteristics, some of which – due to data constraints – are often unobservable outside the laboratory (e.g. income and risk aversion). On top of that, our results are also consistent with previous experimental research highlighting a non-monotonic relationship between the level of impatience and the probability of donating (e.g. Angerer et al. 2014).

Nevertheless, we need to acknowledge that an unobservable factor may indeed play a role. Despite so, we believe that the robustness of our findings highlight that individual parameters and attitudes have an important impact, in addition to tax incentives, on altruistic behaviors. In particular, our main result highlights that impatience plays an important role in decreasing the probability of donating, thus confirming and complementing previous experimental results. Our analysis has focused on time preference, but the dataset and the methodology that we have used is suited to explore in further research the role of other individual preference parameters, such as risk aversion, thus enriching the empirical comprehension of donation drivers.

References

Albanese, G., de Blasio, G., Sestito, P., 2013. Trust and preferences: evidence from survey data. Temi di discussione (Economic working papers) 911. Bank of Italy.

Alvarez-Pelaez, M.J., Diaz, A., 2005. Minimum consumption and transitional dynamics in wealth distribution. Journal of Monetary Economics 52, 633–667.

Andreoni, J., 1989. Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. Journal of Political Economy 97, 1447–1458.

Andreoni, J., 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving? Economic Journal 100, 464–477.

Andreoni, J., 2006. Philanthropy. Handbook of the economics of giving, altruism and reciprocity 2, 1201–1269.

Andreoni, J., Serra-Garcia, M., 2016. Time-Inconsistent Charitable Giving. NBER Working Papers 22824. National Bureau of Economic Research, Inc.

Angerer, S., Glätzle-Rützler, D., Lergetporer, P., Sutter, M., 2014. Donations, risk attitudes and time preferences: A study on altruism in primary school children. Journal of Economic Behavior & Organization .

Aron, A., Aron, E.N., Tudor, M., Nelson, G., 1991. Close relationships as including other in the self. Journal of Personality and Social Psychology , 241–253.

Barsky, R.B., Juster, F.T., Kimball, M.S., Shapiro, M.D., 1997. Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. The Quarterly Journal of Economics 112, 537–579.

Batson, C.D., 1995. Prosocial motivation: Why do we help others?, in: Tesser, A. (Ed.), Advanced social psychology. McGraw-Hill, pp. 333–381.

Battistin, E., Brugiavini, A., Rettore, E., Weber, G., 2009. The retirement consumption puzzle: evidence from a regression discontinuity approach. The American Economic Review 99, 2209–2226.

Bauer, T.K., Bredtmann, J., Schmidt, C.M., 2013. Time vs. money: The supply of voluntary labor and charitable donations across europe. European Journal of Political Economy 32, 80–94.

Becchetti, L., Castriota, S., Conzo, P., 2017. Disaster, Aid, and Preferences: The Long-run Impact of the Tsunami on Giving in Sri Lanka. World Development 94, 157–173.

Becker, G.S., 1974. A Theory of Social Interactions. Journal of Political Economy 82, 1063–1093.

Becker, S.O., Egger, P.H., von Ehrlich, M., 2012. Too much of a good thing? on the growth effects of the eu's regional policy. European Economic Review 56, 648 – 668.

Bénabou, R., Tirole, J., 2006. Incentives and prosocial behavior. American Economic Review 96, 1652–1678.

Bia, M., Flores, C.A., Flores-Lagunes, A., Mattei, A., 2014. A Stata package for the application of semi-parametric estimators of dose?response functions. Stata Journal 14, 580–604.

Bottazzi, R., Jappelli, T., Padula, M., 2006. Retirement expectations, pension reforms, and their impact on private wealth accumulation. Journal of Public Economics 90, 2187–2212.

Breman, A., 2011. Give more tomorrow: Two field experiments on altruism and intertemporal choice. Journal of Public Economics 95, 1349–1357.

Buckner, R., Carroll, D., 2007. Self-projection and the brain. Trends in Cognitive Science 11, 49–57.

Cappellari, L., Ghinetti, P., Turati, G., 2011. On time and money donations. The Journal of Socio-Economics 40, 853–867.

Chatterjee, S., Ravikumar, B., 1999. Minimum Consumption Requirements: Theoretical And Quantitative Implications For Growth And Distribution. Macroeconomic Dynamics 3, 482–505.

Cialdini, R., Brown, S., Lewis, B., Luce, C., Neuberg, S., 1997. Reinterpreting the empathy-altruism relationship: When one into one equals oneness. Journal of Personality and Social Psychology 73, 481–494.

Cohen, J.D., Ericson, K.M., Laibson, D., White, J.M., 2020. Measuring time preferences. Journal of Economic Literature .

Cox, J.C., 2004. How to identify trust and reciprocity. Games and Economic Behavior 46, 260 – 281.

De Blasio, G., Nuzzo, G., 2010. Individual determinants of social behavior. The Journal of Socio-Economics 39, 466–473.

Dreber, A., Fudenberg, D., Levine, D.K., Rand, D.G., 2016. Self-control, social preferences and the effect of delayed payments .

Ebert, S., van de Kuilen, G., 2015. Measuring multivariate risk preferences .

Falk, A., Becker, A., Dohmen, T., Huffman, D., Sunde, U., 2014. An Experimentally validated preference survey module. Technical Report. Mimeo, Universität Bonn.

Fehr, E., 2009. On The Economics and Biology of Trust. Journal of the European Economic Association 7, 235–266.

Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time Discounting and Time Preference: A Critical Review. Journal of Economic Literature 40, 351–401.

Fryges, H., Wagner, J., 2008. Exports and productivity growth: First evidence from a continuous treatment approach. Review of World Economics 144, 695–722.

Galor, O., Özak, O., 2016. The Agricultural Origins of Time Preference. American Economic Review 106, 3064–3103.

Giannetti, C., 2019. Debt specialization and performance of european firms. Journal of Empirical Finance 53, 257 – 271.

Glaeser, E.L., Laibson, D., Sacerdote, B., 2002. An Economic Approach to Social Capital. Economic Journal 112, 437–458.

Golsteyn, B.H., Grönqvist, H., Lindahl, L., 2014. Adolescent time preferences predict lifetime outcomes. Economic Journal 124, F739–F761.

Greene, W., 2010. Testing hypotheses about interaction terms in nonlinear models. Economics Letters 107, 291–296.

Guiso, L., Sapienza, P., Zingales, L., 2006. Does Culture Affect Economic Outcomes? Journal of Economic Perspectives 20, 23–48.

Hardisty, D., Thompson, K., Krantz, D., Weber, E., 2013. How to measure time preferences? an experimental comparison of three methods. Judgment and Decision Making 8, 236–249.

Harman, O., 2010. The price of altruism: George Price and the search for the origins of kindness. Random House.

Havranek, T., Horvath, R., Irsova, Z., Rusnak, M., 2015. Cross-country heterogeneity in intertemporal substitution. Journal of International Economics 96, 100–118.

Hirano, K., Imbens, G.W., 2004. The propensity score with continuous treatments. Applied Bayesian modeling and causal inference from incomplete-data perspectives 226164, 73–84.

Imai, K., Van Dyk, D.A., 2004. Causal inference with general treatment regimes. Journal of the American Statistical Association 99.

Imbens, G.W., 2000. The role of the propensity score in estimating dose-response functions. Biometrika 87, 706–710.

Jappelli, T., Pistaferri, L., 2000. Using subjective income expectations to test for excess sensitivity of consumption to predicted income growth. European Economic Review 44, 337–358.

Khalil, E.L., 2004. What is altruism? Journal of economic psychology 25, 97–123.

Kirchgässner, G., 2010. On minimal morals. European Journal of Political Economy 26, 330–339.

Knack, S., Keefer, P., 1997. Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. The Quarterly Journal of Economics 112, 1251–1288.

List, J.A., 2011. The market for charitable giving. The Journal of Economic Perspectives , 157–180.

Meer, J., Priday, B.A., 2020. Generosity Across the Income and Wealth Distributions. Technical Report. National Bureau of Economic Research.

Ottoni-Wilhelm, M., Vesterlund, L., Xie, H., 2017. Why do people give? testing pure and impure altruism. American Economic Review 107, 3617–33.

Papke, L.E., Wooldridge, J.M., 1996. Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates. Journal of Applied Econometrics 11, 619–32.

Pearl, J., 2012. On a class of bias-amplifying variables that endanger effect estimates. arXiv preprint arXiv:1203.3503 .

Rachlin, H., Jones, B.A., 2006. Social discounting. Psychological Science 17, 283–286.

Rosenbaum, P.R., 1984. The consquences of adjustment for a concomitant variable that has been affected by the treatment. Journal of the Royal Statistical Society. Series A (General) , 656–666.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.

Rosenzweig, M.R., Wolpin, K.I., 1993. Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India. Journal of Political Economy 101, 223–244.

Samuelson, P.A., 1937. A Note on Measurement of Utility. Review of Economic Studies 4, 155–161.

Shapiro, J., 2010. Discounting for you, me, and we: Time preference in groups and pairs. Technical Report. mimeo.

Sutter, M., Kocher, M.G., Glätzle-Rützler, D., Trautmann, S.T., 2013. Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. The American Economic Review 103, 510–531.

Wade-Benzoni, K.A., 2008. Maple trees and weeping willows: The role of time, uncertainty, and affinity in intergenerational decisions. Negotiation and Conflict Management Research 1, 220–245.

Wade-Benzoni, K.A., Tost, L.P., 2009. The egoism and altruism of intergenerational behavior. Personality and Social Psychology Review 13, 165–193.

Wang, M., Rieger, M.O., Hens, T., 2016. How time preferences differ: Evidence from 53 countries. Journal of Economic Psychology 52, 115–135.

Wooldridge, J.M., 2016. Should instrumental variables be used as matching variables? Research in Economics 70, 232–237.

Yi, R., Charlton, S., Porter, C., Carter, A.E., Bickel, W.K., 2011. Future altruism: Social discounting of delayed rewards. Behavioural processes 86, 160–163.

A Proofs

A.1 Proof of Lemma 1

Proof. We rewrite the two first order conditions with respect to c and a, where we have incorporated the budget constraint: $d = (1+r)(y-c-a)$.

$$\Phi_c \equiv \frac{\partial U}{\partial c} = u'(c) - \frac{1}{1+\rho}u'(d)(1+r) = 0 \tag{14}$$

$$\Phi_a \equiv \frac{\partial U}{\partial a} = \alpha\left(\gamma + \frac{\phi}{1+\rho}\right)v'(a) - (1-\alpha)\frac{1+r}{1+\rho}u'(d) = 0 \tag{15}$$

By using (14), we can rewrite Eq. (15) as:

$$\alpha\left(\gamma + \frac{\phi}{1+\rho}\right)v'(a) - (1-\alpha)u'(c) = 0 \tag{16}$$

These conditions implicitly define the solution c^*, a^* at interior solution, i.e. when $a^* > 0$ and $d^* > \underline{k}$. In order to determine the effect of ρ on $a*$, we use the implicit function theorem in two dimensions:

$$\begin{pmatrix} \frac{dc^*}{d\rho} \\ \frac{da^*}{d\rho} \end{pmatrix} = -H^{*(-1)} \begin{pmatrix} \frac{\partial \Phi_c(c^*,a^*)}{\partial \rho} \\ \frac{\partial \Phi_a(c^*,a^*)}{\partial \rho} \end{pmatrix}$$

where the matrix H is defined as:

$$H \equiv \begin{pmatrix} \frac{\partial \Phi_c(c^*,a^*)}{\partial c} & \frac{\partial \Phi_c(c^*,a^*)}{\partial a} \\ \frac{\partial \Phi_a(c^*,a^*)}{\partial c} & \frac{\partial \Phi_a(c^*,a^*)}{\partial a} \end{pmatrix}$$

We compute the following derivatives of Φ_c and Φ_a with respect to a and c, assessing their sign based on assumptions on $u(\cdot)$, $v(\cdot)$ and on parameters.

$$\frac{\partial \Phi_c}{\partial c} = (1-\alpha)u''(c) + (1-\alpha)\frac{(1+r)^2}{1+\rho}u''(d) \quad \Rightarrow \quad < 0 \tag{17}$$

$$\frac{\partial \Phi_c}{\partial a} = (1-\alpha)\frac{(1+r)^2}{1+\rho}u''(d) \quad \Rightarrow \quad < 0 \tag{18}$$

$$\frac{\partial \Phi_a}{\partial c} = -(1-\alpha)u''(c) \quad \Rightarrow \quad > 0 \tag{19}$$

$$\frac{\partial \Phi_a}{\partial a} = \alpha\left(\gamma + \frac{\phi}{1+\rho}\right)v''(a) \quad \Rightarrow \quad < 0 \tag{20}$$

We also compute the derivative of Φ_c and Φ_a with respect to ρ

$$\frac{\partial \Phi_c}{\partial \rho} = \frac{1+r}{(1+\rho)^2}(1-\alpha)u'(d) \qquad \Rightarrow \quad > 0 \tag{21}$$

$$\frac{\partial \Phi_a}{\partial \rho} = -\alpha \frac{\phi}{(1+\rho)^2} v'(a) \qquad \Rightarrow \quad < 0 \tag{22}$$

Since we need to evaluate derivatives in $\{c^*, a^*\}$, we can use Eq. (15) to rewrite (21) in interior solution as:

$$\frac{\partial \Phi_c(c^*, a^*)}{\partial \rho} = \frac{\alpha}{(1+\rho)} \left(\gamma + \frac{\phi}{1+\rho} \right) v'(a^*) \qquad \Rightarrow \quad > 0 \tag{23}$$

The inverse of matrix H is hence given by:

$$H^{-1} = \frac{1}{\det H} \begin{pmatrix} \frac{\partial \Phi_a}{\partial a} & -\frac{\partial \Phi_c}{\partial a} \\ -\frac{\partial \Phi_a}{\partial c} & \frac{\partial \Phi_c}{\partial c} \end{pmatrix}$$

From the system (17)-(20), it can be seen that:

$$\det H^* = \frac{\partial \Phi_c}{\partial c} \frac{\partial \Phi_a}{\partial a} - \frac{\partial \Phi_c}{\partial a} \frac{\partial \Phi_a}{\partial c} > 0$$

Strictly concavity of $u(\cdot)$ ensures that H is invertible and, together with the strictly concavity of $v(\cdot)$, $\det H > 0$.

We can then obtain that $\partial c^* / \partial \rho > 0$ since:

$$\frac{\partial c^*}{\partial \rho} = -\frac{1}{\det H} \left(\frac{\partial \Phi_a}{\partial a} \frac{\partial \Phi_c}{\partial \rho} - \frac{\partial \Phi_c}{\partial a} \frac{\partial \Phi_a}{\partial \rho} \right) \qquad > 0 \tag{24}$$

The sign of $\partial a^* / \partial \rho$ is ambiguous, since:

$$\frac{\partial a^*}{\partial \rho} = -\frac{1}{\det H} \left(-\frac{\partial \Phi_a}{\partial c} \frac{\partial \Phi_c}{\partial \rho} + \frac{\partial \Phi_c}{\partial c} \frac{\partial \Phi_a}{\partial \rho} \right) \qquad \lesseqqgtr 0 \tag{25}$$

It holds that

$$\frac{\partial a^*}{\partial \rho} < 0 \qquad \text{iff} \qquad -\frac{\partial \Phi_a}{\partial c} \frac{\partial \Phi_c}{\partial \rho} + \frac{\partial \Phi_c}{\partial c} \frac{\partial \Phi_a}{\partial \rho} > 0 \tag{26}$$

Plugging (19), (23), (17), and (22) into (26), after some algebraical computations and simplifications we get Condition (10):

$$\frac{\partial a^*}{\partial \rho} \leq 0 \text{ iff } \frac{\gamma}{\phi} \leq \left(\frac{1+r}{1+\rho} \right)^2 \frac{u''(d^*)}{u''(c^*)} \tag{27}$$

Under CRRA utility function: $u(x) = x^{1-\theta}/(1-\theta)$ with $\theta > 0$, it holds that:

$$\left.\frac{u''(d)}{u''(c)}\right|_{c^*,d^*} = \left(\frac{c^*}{d^*}\right)^{1+\theta} \tag{28}$$

$$\left.\frac{c}{d}\right|_{c^*,d^*} = \left(\frac{1+\rho}{1+r}\right)^{1/\theta} \tag{29}$$

where (29) follows from the first order condition (8).

By using (28) and (29) in Condition (27) we get condition (11).

The logarithmic case, i.e.: $u(x) = \ln(x)$ can be obtained analogously or by taking the limit for $\theta \to 1$ in Condition (11). In particular, since in the logarithmic case $u''(d)/u''(c) = [(1+\rho)/(1+r)]^2$, the Condition becomes $\gamma < \phi$, as in Condition (12).

It can be also proved that d^* is unambiguously decreasing in ρ. From the budget constraint $d^* = (1+r)(y - c^* - a^*)$we have

$$\frac{\partial d^*}{\partial \rho} = -\left(\frac{\partial c^*}{\partial \rho} + \frac{\partial a^*}{\partial \rho}\right) = \frac{1}{\det H}\left(\frac{\partial \Phi_a}{\partial a}\frac{\partial \Phi_c}{\partial \rho} - \frac{\partial \Phi_c}{\partial a}\frac{\partial \Phi_a}{\partial \rho} - \frac{\partial \Phi_a}{\partial c}\frac{\partial \Phi_c}{\partial \rho} + \frac{\partial \Phi_c}{\partial c}\frac{\partial \Phi_a}{\partial \rho}\right)$$

After some algebraical computations and simplifications we get

$$\frac{\partial d^*}{\partial \rho} = -\left(\frac{\partial c^*}{\partial \rho} + \frac{\partial a^*}{\partial \rho}\right) = \frac{1}{\det H}\left(\alpha v''(a)(\gamma\rho + \gamma + \phi)^2 + (1-\alpha)\gamma(1+\rho)^2 u''(c)\right] \qquad < 0 \tag{30}$$

□

A.2 Proof of Lemma 2

Proof. Since d^* is decreasing in ρ, for ρ sufficiently high ($\rho > \bar{\rho}$), the minimum consumption requirement $\underline{k}$ is met, i.e.: $d = \underline{k}$. The objective function becomes:

$$(1-\alpha)u(c) + \alpha\gamma v'(a) + \frac{1}{1+\rho}\left[(1-\alpha)u(\underline{k} + \alpha\phi v(a)\right]$$

By replacing c from the budget constraint: $c = y - a - \underline{k}/(1+r)$, the objective function can be derived only with respect to a. We get the first order condition:

$$\Psi \equiv -(1-\alpha)u'(c) + \alpha v'(a)\left[\frac{\gamma(1+\rho)+\phi}{1+\rho}\right] = 0 \tag{31}$$

By applying the implicit function theorem in one dimension we have that:

$$\frac{da^*}{d\rho} = -\frac{\frac{\partial \Psi}{\partial \rho}}{\frac{\partial \Psi}{\partial a}} \tag{32}$$

It holds:

$$\frac{\partial \Psi}{\partial a} = (1-\alpha)u''(c) + \alpha v''(a)\left[\frac{\gamma(1+\rho)+\phi}{1+\rho}\right] \qquad < 0 \tag{33}$$

$$\frac{\partial \Psi}{\partial \rho} = -\frac{\alpha\phi}{(1+\rho)^2}v'(a) < 0 \tag{34}$$

Plugging (33) and (34) into (32), it follows that $\frac{da^*}{d\rho} < 0$.

The threshold level $\bar{\rho}$ is the level of ρ that satisfies Eq. (15) when $d = \underline{k}$ and $a_{\underline{k}} \equiv y - \frac{k}{1+r} - u'^{-1}\left[\frac{1+r}{1+\rho}u'(\underline{k})\right]$:

$$\bar{\rho} = \frac{1-\alpha}{\gamma\alpha}\frac{u'(\underline{k})}{v'(a_{\underline{k}})} - \frac{\phi+\gamma}{\gamma} \tag{35}$$

The threshold $\bar{\rho}$ is increasing in y and r and is decreasing in the altruism parameters α, γ, ϕ and in $\underline{k}$. The economic intuition are the following: $\bar{\rho}$ is higher if income relative to the minimum consumption level is higher because it is easier to afford a consumption above the minimum level. Similarly, $\bar{\rho}$ is higher if the interest rate is higher so that less savings are necessary to afford the same level of consumption in the future. As far as the altruistic parameters are concerned, an increase in their value implies a lower $\bar{\rho}$ because an increase in altruistic parameters implies a shift toward donation in the trade-off with own consumption, but this reduction in consumption makes the minimum consumption constraint more likely to be binding.

As $a_{\underline{k}}$ decreases with ρ, for high ρ, it reaches 0 and any further increase in ρ has no longer effect given the non-negativity constraint.

Hence to summarize when $\rho > \bar{\rho}$, we have $da^*/d\rho \leq 0$. □

B Tables and figures

Table 1: Summary statistics and Variable Description

Variable	Description	Mean	Std. Dev.	Min.	Max.	N
Donate_{2012}	This is a dummy variable equal to 1 if the respondent donated in 2012, and 0 otherwise.	0.190	0.393	0	1	2975
$\text{Amount Donated}_{2012}$	This variable contains the amount donated by the respondent in 2012.	61.313	323.609	0	11000	2975
$\text{Discount rates}_{2012}$	This is the treatment variable and measures the impatience level of the respondent. It is explained in detail in section 2.2.	0.082	0.094	0	0.3	2975
$\text{Risk tolerance}_{2010}$	This is categorical variable measuring the (financial) risk tolerance of the respondent (i.e. 1 if he prefers low risk for low returns, 4 if he prefers high risk for high return).	1.741	0.789	1	4	2975
Education_{2010}	This is categorical variable measuring the education level of the respondent.	3.698	1.692	1	8	2975
$\text{Family Size}_{2010}$	This variable measures the total number of family component.	2.45	1.239	1	12	2975
Age_{2010}	This variable measures the age of the respondent (i.e. head of the household).	0.408	0.492	0	1	2975
$\text{Family Income Holders}_{2010}$	This variable measures the total number of family component earning a salary.	60.714	14.372	20	95	2975
$\text{House Property}_{2010}$	This is a dummy variable equal to one if the household own the house is living in, and 0 otherwise.	1.667	0.73	1	6	2975
Single_{2010}	This is a dummy variable equal to one if the head of household is single, and 0 otherwise.	0.764	0.424	0	1	2975
Divorced_{2010}	This is a dummy variable equal to one if the head of household is divorced, and 0 otherwise.	0.1	0.3	0	1	2975
Widow_{2010}	This is a dummy variable equal to one if the head of household is widow, and 0 otherwise.	0.079	0.269	0	1	2975
Log Income_{2010}	This variable measures the logarithm of the income of the head of household.	0.2	0.4	0	1	2975
Employed_{2010}	This is a dummy variable equal to one if the head of household is employed, and 0 otherwise.	9.832	0.916	1.142	12.472	2975

Table 2: BASIC PROBIT MODEL

The dependent variable is a dummy variable (i.e. *Donation*) equal to 1 if the individual donated in year 2012, and zero otherwise.

	Coeff	*Margins*	*Coeff*	*Margins*	*Coeff*	*Margins*
	a	*a*	*b*	*b*	*c*	*c*
Impatience	-1.516***	-0.233***	2.722*	0.046	5.825	0.242
	(0.502)	(0.077)	(1.617)	(0.121)	(3.659)	(0.235)
Impatience^2			-16.631***		-51.151	
			(6.114)		(36.584)	
Impatience^3					84.993	
					(87.901)	
Constant	-1.331***		-1.426***		-1.451***	
	(0.060)		(0.070)		(0.076)	
Log-likelihood	-1443	-1443	-1440	-1440	-1439	-1439
Observations	2975	2975	2975	2975	2975	2975

Notes: * denotes $p < 0.10$,** $p < 0.05$ and ***$p < 0.01$

Table 3: Estimation of the Generalized Propensity Score (GPS)

This table reports the marginal effects of a two-limit Tobit model. The dependent variable is the discount rates in 2012. Descriptions of the regressors are available in Table(1).

Variable	*Coef.*	*Std. err*
Risk Tolerance$_{2010}$	-0.007*	(0.004)
Education$_{2010}$	-0.004*	(0.002)
Family Size$_{2010}$	0.016***	(0.004)
Female	-0.013*	(0.007)
Age$_{2010}$	0.001***	(0.000)
Family Income Holders$_{2010}$	-0.021***	(0.005)
House Property$_{2010}$	-0.014**	(0.007)
Single$_{2010}$	0.021*	(0.011)
Divorced$_{2010}$	0.023*	(0.012)
Widow$_{2010}$	0.043***	(0.010)
Log Income$_{2010}$	-0.023***	(0.004)
Employed$_{2010}$	0.017*	(0.009)
Constant	0.234***	(0.040)
Sigma	0.142***	(0.003)
Log-likelihood	-188	
Observations	2975	

Notes: * denotes $p < 0.10$,** $p < 0.05$ and ***$p < 0.01$

Table 4: Balance of covariates for the GPS

This table illustrates the results for test of the balancing properties. Observations are first divided into three "treatment" groups according to the actual level of the discount rates: [0, 0.035], [0.035, 0.125] and [0.125, 0.3]. In addition, within each group, observations are divided into five blocks according to the estimated GPS (see Table (5)). For each variable, we then compare the equality of covariates between units who belong to the the treatment interval of the HHI, and units that are in the same GPS interval but belong to another treatment interval. The balancing property has been tested using a standard two-sided *t-test*. There is strong evidence against the balancing property when $1.96 < t < 2.576$.

	Group 1: [0,0.035]			*Group 2: [0.035,0.125]*			*Group 3: [0.125,0.3]*		
	Mean	**Std**	**t-test**	**Mean**	**Std**	**t-test**	**Mean**	**Std**	**t-test**
Risk Tolerance$_{2010}$	0.047	0.028	1.658	-0.055	0.031	**-1.767**	0.018	0.039	0.457
Education$_{2010}$	-0.023	0.053	-0.432	0.059	0.066	0.892	0.064	0.074	0.857
Family Size$_{2010}$	0.064	0.046	1.400	-0.059	0.048	-1.227	0.001	0.058	0.009
Age$_{2010}$	0.108	0.511	0.212	-0.658	0.559	-1.177	0.227	0.673	0.337
Family Income Holders$_{2010}$	0.021	0.024	0.869	0.005	0.029	0.190	-0.002	0.034	-0.072
House Property$_{2010}$	-0.004	0.015	-0.273	0.005	0.017	0.295	0.013	0.019	0.702
Single$_{2010}$	0.003	0.011	0.249	0.003	0.012	0.249	-0.004	0.015	-0.275
Divorced$_{2010}$	-0.004	0.010	-0.399	0.002	0.011	0.150	0.000	0.013	0.004
Widow$_{2010}$	-0.004	0.013	-0.314	-0.023	0.015	-1.537	0.023	0.014	1.601
Log Income$_{2010}$	-0.039	0.027	-1.446	0.030	0.033	0.919	0.096	0.035	**2.719**
Employed$_{2010}$	-0.008	0.018	-0.436	0.006	0.019	0.336	0.020	0.024	0.846

Table 5: Cell size for mean comparison of treat and control units

Block	*Group 1*	*Control 1*	*Group 2*	*Control 2*	*Group 3*	*Control 3*
1	290	543	191	440	109	885
2	290	256	191	414	109	573
3	290	289	191	381	109	458
4	290	211	191	430	109	300
5	290	200	191	328	108	188
Total	1449	1499	955	1993	544	2404

Groups are generated according to three cutpoints of discount ratio in 2012 (i.e. 0.035, 0.125 and 0.3), whereas blocks are generated according to the quintiles of the GPS evaluated at the median treatment intensity for each group. The sum of observations over blocks in a group yields the total number of observations in that group. The sum of observations in a group with observations from the respective control group yield the total number of observations in the common support region.

Table 6: Estimation of the dose-response function for the probability of donating

The dependent variable is a dummy variable (i.e. *Donation*) equal to 1 if the individual donated in 2012 and zero otherwise.

Variable	Coeff	Std Error
Impatience$_{2012}$	14.779	(10.322)
Impatience$^2{}_{2012}$	-192.272***	(74.364)
Impatience$^3{}_{2012}$	387.144**	(158.668)
Gps	8.628***	(1.704)
Gps^2	-4.973***	(1.182)
Gps^3	0.785***	(0.233)
Gps·Impatience$_{2012}$	0.493	(2.042)
Constant	-4.114***	(0.517)
Log-likelihood	-1396	
Observations	2948	

Notes: * denotes $p < 0.10$,** $p < 0.05$ and ***$p < 0.01$

Figure 1: RISK TOLERANCE

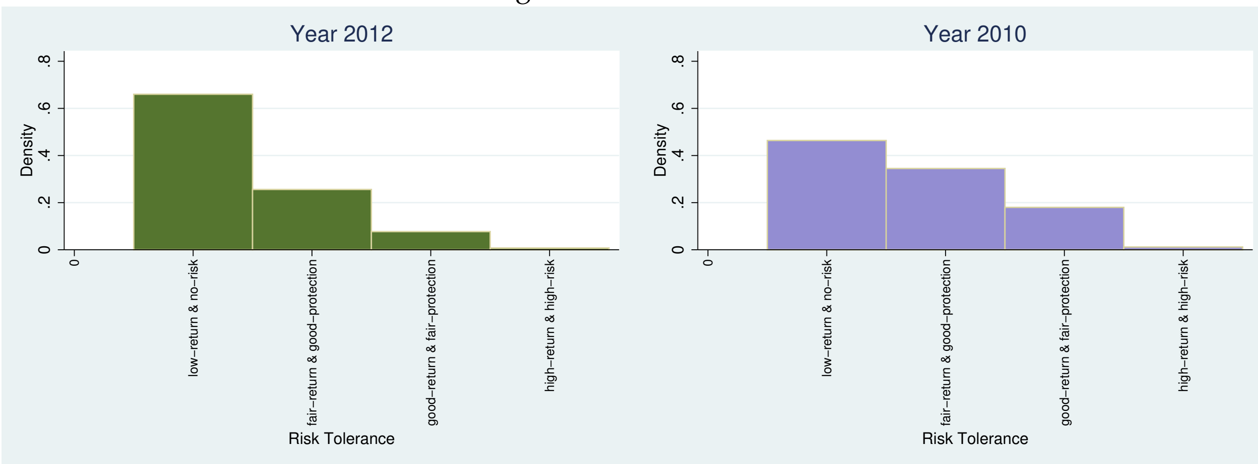


Figure 2: Donation class

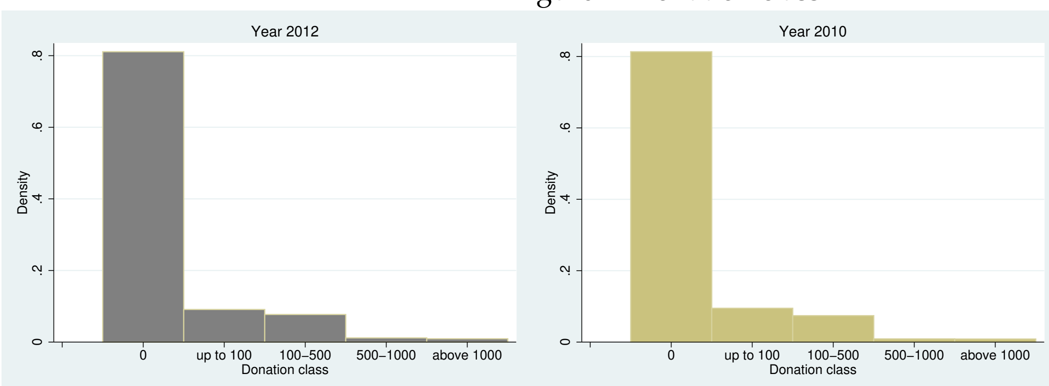


Figure 3: Elicitation of Discount rates

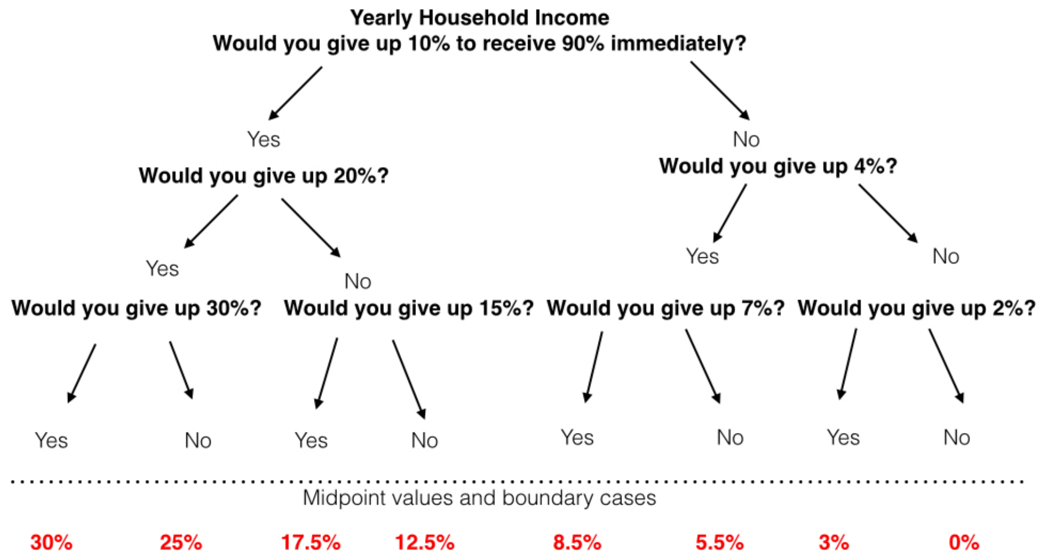


Figure 4: Distribution of answers on discount rates across years

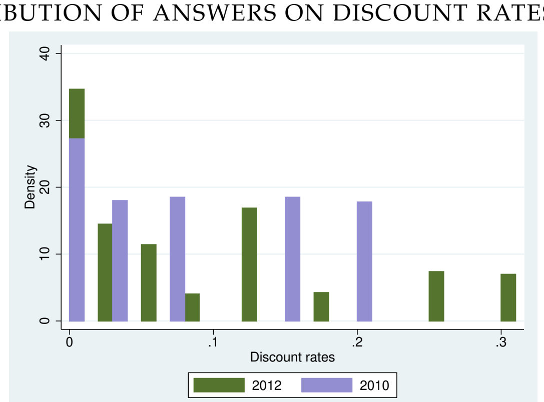


Figure 5: PROBIT MODEL MARGINS

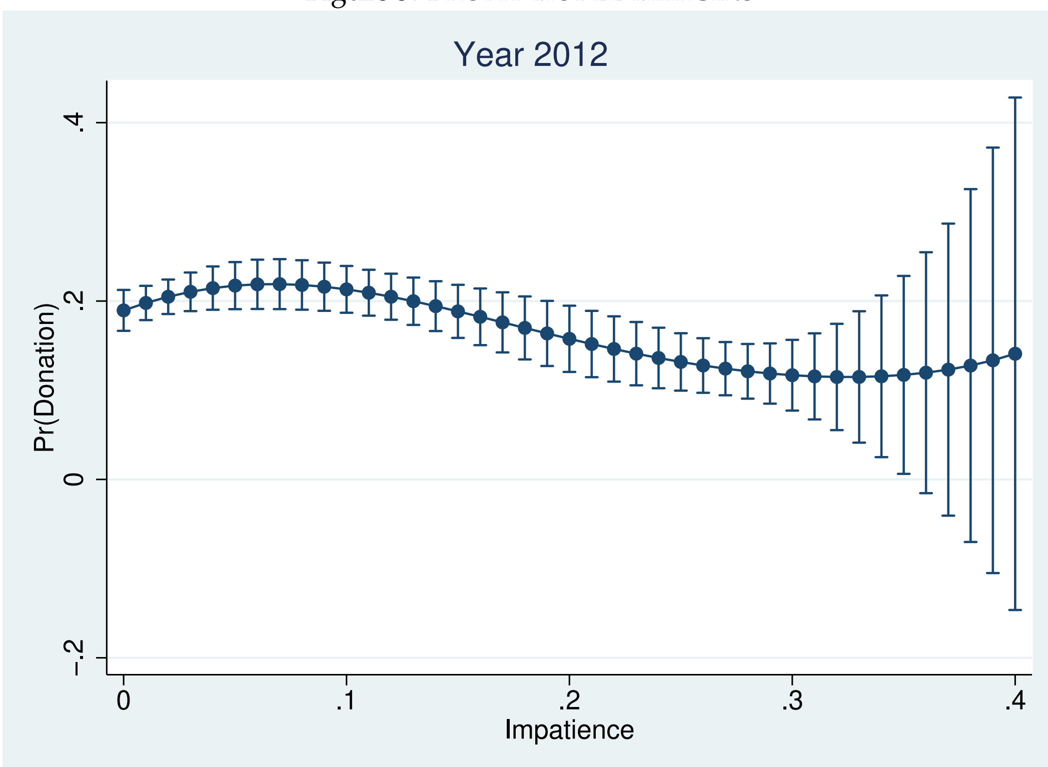
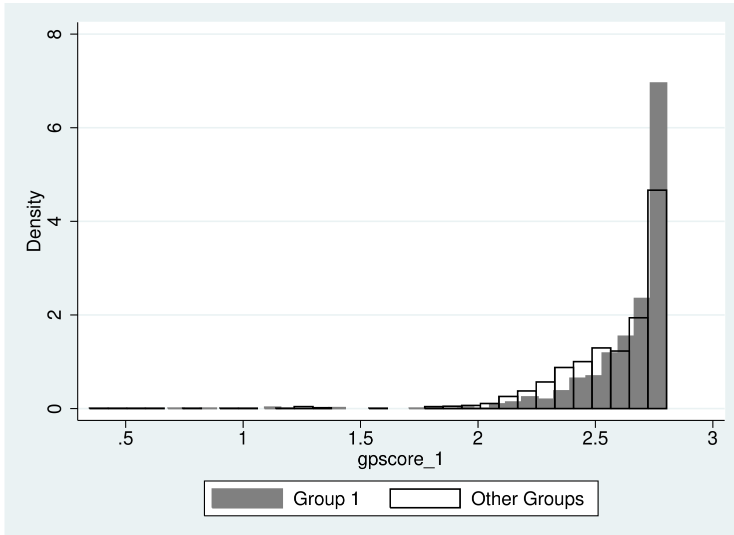
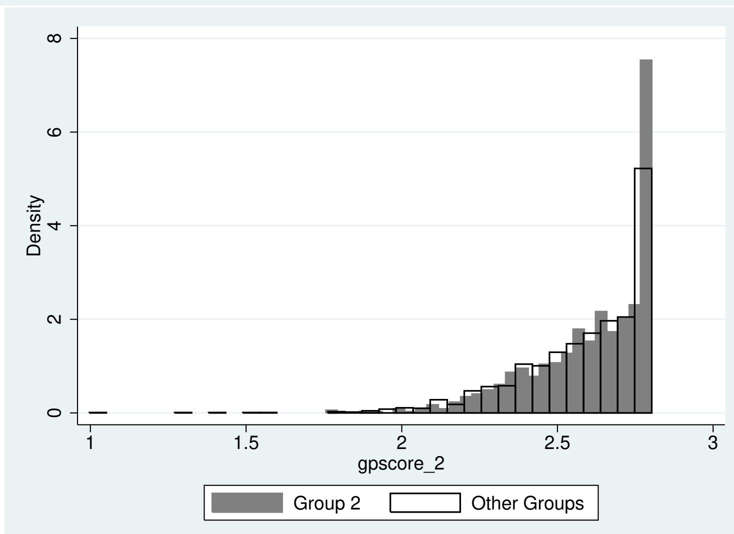


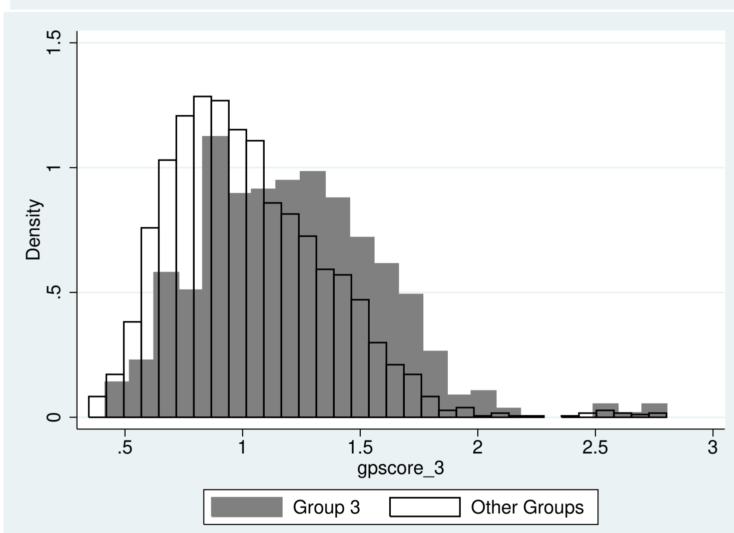
Figure 6: COMMON SUPPORT OF THE GENERALIZED PROPENSITY SCORE



a.



b.



c.

Figure 7: DOSE-RESPONSE AND TREATMENT-EFFECT FUNCTION FOR DONATION (PROBABILITY)

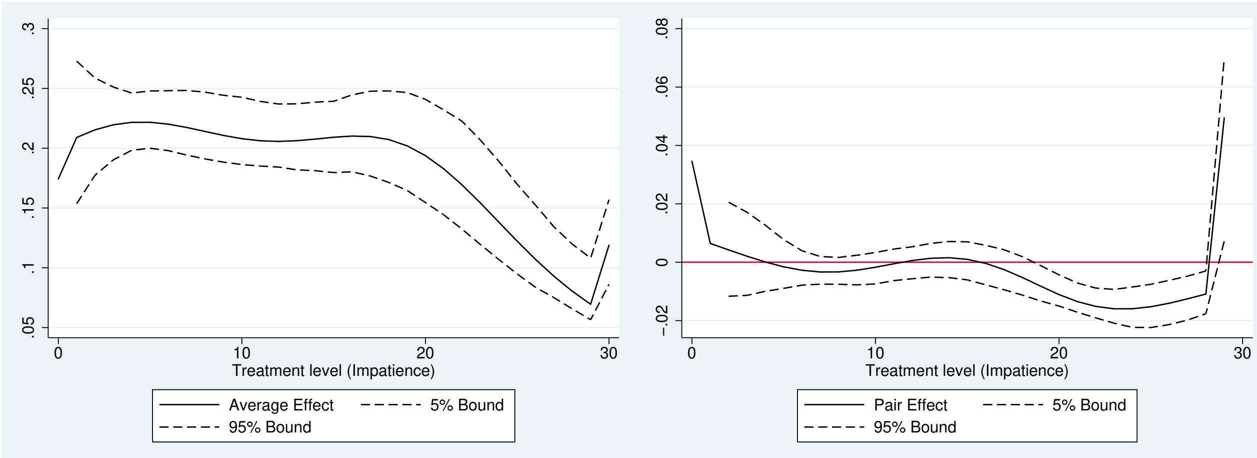


Figure 8: DOSE-RESPONSE AND TREATMENT-EFFECT FUNCTION FOR DONATION AMOUNT

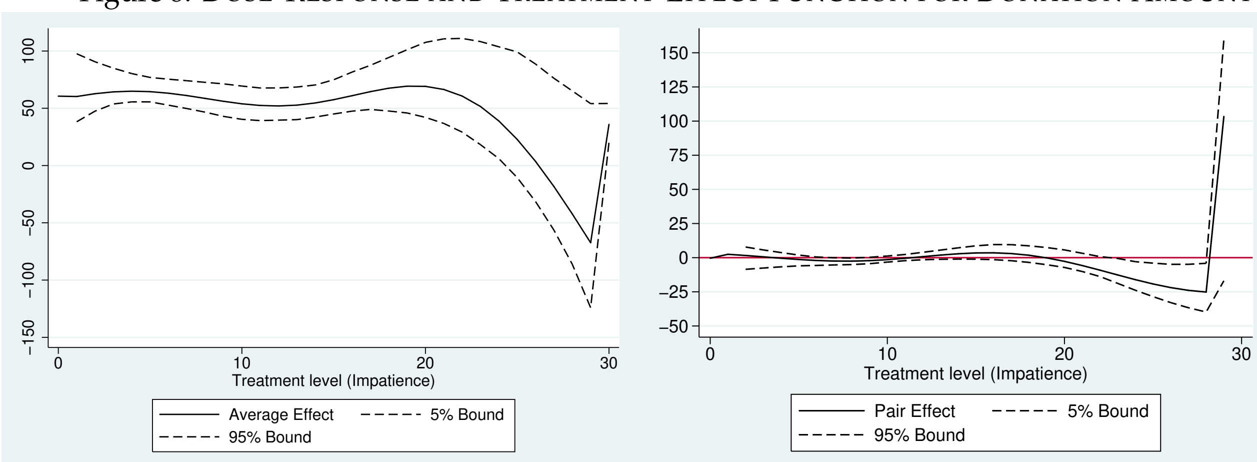


Figure 9: DOSE-RESPONSE AND TREATMENT-EFFECT FUNCTION FOR DONATION (2010)