

# Altruism begets altruism\*

Stephanie A. Heger<sup>†</sup>

Department of Economics, The University of Bologna

Robert Slonim

Economics Department, University of Technology Sydney

September 7, 2021

**ABSTRACT:** Guided by Bem’s (1972) self-perception theory, we design an experiment to ask whether morally-motivated behaviour, e.g., charitable giving, is history-dependent. Using a popular policy nudge, the default option, we exogenously vary altruism “now” and show that giving “now” causes a 41 percentage point (or 200%) increase in the probability of giving “later”; that is, altruism begets altruism. We further show that, consistent with self-perception theory, the *choice* to behave altruistically “now”, rather than the nudge itself, is the crucial element in the causal relationship. These findings are consistent with a model of positive path-dependence, which we interpret as moral consistency.

---

\*This project benefitted from many helpful comments and conversations, in particular, we thank Christine Exley, Lata Gangadharan, Jonathan Meer, Tom Wilkening, Nina Xue and seminar and conference participants at the University of Sydney, the University of Melbourne, the Early Career Behavioral Economics Conference (ECBE 2019), the ESA Junior Faculty Seminar Series and the Science of Philanthropy Initiative (SPI). Slonim gratefully acknowledges funding from ARC DP 150101307. This project was approved by the Human Research Ethics Committee at the University of Sydney, Project 2017/529.

<sup>†</sup>[Corresponding author] Email: saheger@gmail.com.

# 1 Introduction

Policy shapes society by encouraging socially desirable behaviour . Thanks in part to the United States tax policy which allows individuals to deduct charitable gifts from their pre-taxed income (Clotfelter, 1980; Meer and Priday, 2019), charitable giving made up 2% of GDP in the USA in 2019 (see Giving USA, 2020). Additionally, charitable giving is associated with increased happiness (Anik et al., 2009) and health (Yörük, 2014). Thus, it is not surprising that there is a large body of research that examines interventions aimed at increasing charitable giving, including the effects of price (Karlan and List, 2007), various fundraising schemes (Huck, Rasul, and Shephard, 2015), efficiency concerns (Gneezy, Keenan, and Gneezy, 2014; Exley, 2015b), social pressure (List and Lucking-Reiley, 2002; Frey and Meier, 2004; Shang and Croson, 2009), and identity (Kessler and Milkman, 2016). And in addition to examining the immediate effects of interventions, there is a growing literature looking at the intertemporal spill-overs of pro-social behaviours generated by interventions at charitable giving(Shang and Croson, 2009; Cairns and Slonim, 2011; Gneezy et al., 2012; Castillo, Petrie, and Samek, 2017). For example, Shang and Croson (2009) find that subjects who experienced social pressure to give to public radio in  $t - 1$  were more likely to give in  $t - 1$  but also more likely to renew their membership in time  $t$  and donate more in time  $t$  relative to those subjects who did not experience social pressure at time  $t - 1$ . Gneezy et al. (2012) show that when subjects are randomly assigned to incur the cost of a charitable donation at  $t - 1$ , they will be more likely to behave honestly in time  $t$ , relative to those randomly assigned to make a costless donation at time  $t - 1$ .<sup>1</sup> In other words, the existing literature has examined how an intervention at time  $t - 1$  affects behaviour in time  $t - 1$  and behaviour in time  $t$ .

In this paper, by contrast, we estimate how *choices* at  $t - 1$  affect *choices* at  $t$ . We hypothesize and show that altruism begets altruism;<sup>2</sup> a nudge induces people to give now, and this increase in *choosing to give* now causes an increase in later giving.<sup>3</sup> We do this by instrumenting for choices at  $t - 1$  with the random assignment to a default option nudge that encourages people to give at time  $t - 1$  and then we also elicit giving decisions at time  $t$ . Our

---

<sup>1</sup>Cassar, d’Adda, and Grosjean (2014); Peysakhovich and Rand (2016); Engl, Riedl, and Weber (2018) also study positive spill-overs in prosociality, but not due to an increase in charitable giving. For example, Peysakhovich and Rand (2016) examine how random exposure to incentive structures for cooperation in infinitely repeated prisoner’s dilemma games affect future cooperation in one-shot coordination games, while Cassar, d’Adda, and Grosjean (2014) study the effect of institutional strength in a market game on subjects’ behaviour in a trust game.

<sup>2</sup>Our hypothesis is supported by the review in Gee and Meer (2019), who conclude that while there is some evidence of donor fatigue (Meier, 2007; Damgaard and Gravert, 2018), “the preponderance of evidence finds that gifts today do not cannibalize gifts tomorrow.”

<sup>3</sup>See Thaler and Sunstein (2003) and Sunstein and Thaler (2008) for a review of nudges.

experimental and econometric approach is thus distinct from previous literature that shows a positive correlation in giving over time (Landry et al., 2010; Adena and Huck, 2019);<sup>4</sup> we provide causal evidence of the effect of altruistic choices *now* on altruism *later*.

Our hypothesis comes from self-perception theory (Bem, 1972), which provides a framework for modeling morally-motivated behaviours over time. Self-perception theory posits that individuals use past choices to make inferences about their own identity, which then inform future choices. Bénabou and Tirole (2011) formalize self-perception theory in economics; they model individuals with imperfect memories of their identity who use their past choices to make inferences about their identity, which serves as a guide for current choices. Further, self-perception theory emphasizes the role of *choice* in linking behaviours to identity (Zanna, 1972). Specifically, if an individual is forced to donate then they cannot infer much about their identity as an altruistic person, whereas if they had made an active choice to donate then they can make an inference about their identity as an altruistic person from the past donation. Thus, self-perception theory predicts a path-dependency between *chosen* moral actions over time.<sup>5</sup>

Motivated by the history-dependence in chosen actions modelled by Bem (1972) and Bénabou and Tirole (2011), we model our decision-maker’s utility at time  $t$  as dependent on his current choice of charitable giving and his  $t - 1$  choice of altruism using a habit formation model (Pollak, 1970). Further, his  $t - 1$  choice depends on his previous choices of altruism as well as on whether he is nudged towards altruism or selfishness at  $t - 1$ . We thus model moral consistency in altruism as habit persistent charitable giving, meaning that charitable giving is not just positively correlated over time, but that an increase in past giving *causes* an increase in giving today. This is also consistent with Meer (2013), who finds evidence consistent with habit formation preferences for giving—using the performance of college athletic teams as an instrument for giving when “young”, Meer (2013) shows that giving when young causes an increase in giving 20 years after graduation.

Self-perception theory serves as the guiding light to design an experiment capable of answering whether altruism begets altruism. We conduct an online experiment in which we nudged individuals to either donate to charity or to keep the cash for themselves by setting their default option to “donate” (henceforth: Default Charity) or to “keep” (henceforth: Default Cash), respectively. To avoid donating, subjects in the Default Charity condition

---

<sup>4</sup>Cappelen et al. (2017) use a similar approach in a different context and find that incentivizing subjects to go to the gym increases the likelihood of exercise, which in turn, increases the subjects’ academic performance.

<sup>5</sup>Thus, we also contribute to a relatively recent literature that examines the role of agency in charitable giving (Eckel, Herberich, and Meer, 2018; Kessler, Milkman, and Zhang, 2019), which finds that an increased sense of agency does not affect donation behaviour at the extensive margin, but does have a positive and significant effect on the intensive margin.

must opt-out of giving to charity; by contrast, subjects in the Default Cash condition must opt-in to giving to charity and opt-out of keeping cash (Round 1). We chose a default option nudge for two reasons: (1) to generate exogenous variation in giving “now”; and (2) to maintain the active choice in Round 1 behaviour. Consistent with past research on default option nudges (Benartzi and Thaler, 2007; Choi et al., 2003; Cronqvist and Thaler, 2004; Madrian and Shea, 2001; Kessler and Roth, 2012, 2014), we find that our nudge positively impacts charitable giving behaviour.<sup>6</sup> Specifically, we find that subjects in the Default Charity condition are twice as likely to choose to donate in Round 1 than subjects in the Default Cash condition.

The critical part of the design is that at a later point in the experiment, “later”, we ask subjects to make another donation to a different charity to test whether initial altruistic behaviour increases altruism in the future (Round 2). Directly motivated by our model and experimental design, we estimate a local average treatment effect using our randomly assigned nudge as an instrument for Round 1 giving and find that the nudge-induced increase in choosing to give in Round 1 causes giving in Round 2 to increase by 200% or 40 percentage points. To address the question of external validity of our main result, we significantly increase the time lapse between the Round 1 and Round 2 decision to one week in a separate experiment.<sup>7</sup> Overall, our experiments show that the nudge-induced choice to be altruistic in “now” begets more altruism “later”, thus generating a virtuous cycle of altruism.

To provide direct support for the validity of our identification strategy and to demonstrate the central causal role that *choice* in Round 1 has on Round 2 donation choices, we ran additional treatments in which we randomly assigned subjects to a default position in Round 1, but do not give them the choice to opt-out of their default position; that is, they are forced to make a donation or are forced to keep the money in Round 1 (henceforth: No Choice Treatments). The results from the No Choice Treatments support our main finding. In particular, participants in Default Charity (No Choice) and Default Cash (No Choice) donate at equal rates in Round 2, emphasizing the criticality of *choice* in Round 1 in linking behaviour inter-temporally.<sup>8</sup> In fact, we run three sets of No Choice Treatments, described in

---

<sup>6</sup>Conceptually, setting a default option works by decreasing the marginal psychological costs of choosing the desired behaviour.

<sup>7</sup>We followed Andreoni and Serra-Garcia (2021) in choosing one week as the time between “now” and “later”. We show in Section 3.2.2 that moral consistency persists even with the one week delay between Round 1 and Round 2.

<sup>8</sup>Gneezy et al. (2012) reports results from an experiment in which subjects who are randomly assigned to make a costly donation are more likely to behave honestly in a subsequent period than subjects who are randomly assigned to make a costless donation. Importantly, particularly in relation to our study, subjects in both the costly and costless treatment were forced to donate rather than having to choose whether to behave altruistically. Thus, while a direct impact through salience is possible, Gneezy et al. (2012) prevents an indirect channel predicted by self-perception theory that we will explore here.

detail in Section 2.1.2, in which we vary the degree to which subjects reflect on the choice of donating or keeping money for themselves in Round 1 to ensure that the key aspect driving the path-dependence in the Choice Treatments is, indeed, the active choice in Round 1.

Our contribution is thus twofold. First, we provide evidence in favor of moral consistency; that is, we show that choosing to donate “now” causes an increase in altruism “later”. To claim this causal relationship, we show that the exclusion restriction assumption holds and that the nudge itself is not responsible for the inter-temporal effect, but rather the choice to act altruistically that the nudge induces *now* causes the increase *later*. Therein lies our second contribution—we use experimental treatments to directly test that the theoretical assumptions behind our empirical test hold.

Last, we explore heterogeneous effects of moral consistency. Motivated by Bem (1972)’s self-perception theory and the model of Bénabou and Tirole (2011), we explore the role of identity and moral consistency, contributing to the growing literature on the role of identity in economics (Akerlof and Kranton, 2005) and, more specifically, in charitable giving (Benjamin, Choi, and Fisher, 2010; Kessler and Milkman, 2016). Consistent with both our model and Benabou and Tirole’s (2011) model,<sup>9</sup> individuals whose past limited donations suggest that altruism is a weakly held facet of their identity behave in a significantly more morally consistent manner. For these individuals, choosing to behave altruistically in Round 1 causes an 83 percentage point (or 492%) increase in altruism in Round 2. Self-perception theory suggests that the altruistic behaviour induced by the nudge is more informative for weak altruists than for strong altruists because weak altruists have a much sparser history that will make the Round 1 *choice* more salient.

## 2 Design & Procedures

In this section, we describe our experimental design and the data generated by the experiment. We also present a model of consumption choice and motivate our hypotheses in Section 2.4.

---

<sup>9</sup>In Benabou and Tirole’s (2011) model, individuals for whom altruism is a weakly held facet of their identity are predicted to behave in a more morally consistent manner. On the other hand, their model also predicts that challenges to strongly-held aspects of identity “today” are met with contradictory responses “tomorrow”. Thus, depending on whether an individual has a weakly or strongly held conviction towards altruism, Bénabou and Tirole (2011) predicts either moral consistency (Nisan, 1985; Nisan and Horenczyk, 1990) or moral licensing (Khan and Dhar, 2006; Monin and Miller, 2001; Ploner and Regner, 2013; Sachdeva, Iliev, and Medin, 2009) (also see Blanken, van de Ven, and Zeelenberg (2015) and Mullen and Monin (2016) for a review of this literature in psychology).

## 2.1 Experimental Design

### 2.1.1 Calibrating Preferences

We ran a pre-experimental calibration exercise to gauge the amount that must be donated to the chosen charity for the average subject to be indifferent to giving up \$1. The calibration exercise is important to set the default options such that some subjects will prefer to donate, while other subjects will prefer to keep cash for themselves. By finding a the median point of indifference between donating to charity and keeping cash for self, we can be confident that the nudge towards charity or the nudge towards keeping cash will be on the appropriate margins.

To do the calibration, we used the same charity, CARE, that we will use in the Round 1 decision of the experiment. This exercise follows the calibration exercise in Exley (2015a) and presents subjects with a multiple price list. On each line, they are asked whether they prefer to keep a \$1 and give \$0 to the charity or keep \$0 and give \$x to the charity, where  $x \in \{\$0, \$0.1, \dots, \$3\}$ . While Exley (2015a) uses a within-subject calibration, our calibration is taken as the median point of indifference across subjects, which was \$1 to self was utility-equivalent to \$1.50 to charity. This is how we chose the values in Round 1: subjects in the Default Cash condition were endowed with \$1 to keep for themselves and subjects in the Default Charity condition were endowed with making a \$1.50 donation to the charity. Subjects in this calibration exercise were excluded from participating in any of the experimental conditions that follow.

### 2.1.2 Main Choice Treatments

The main experiment consists of two decision Rounds. In Round 1, subjects were randomly endowed with \$1 cash (Default Cash) or endowed with a \$1.50 donation to the charity CARE (Default Charity). Figures A1a and A1b display what the subjects saw if they were assigned to the Default Cash and Default Charity treatments, respectively. After providing their endowment, we took two additional steps to facilitate a sense of ownership among subjects of their default position. First, we asked subjects in the Default Charity condition to list three ways the charity CARE might spend this money and we asked subjects in the Default Cash condition to list three ways they might spend their cash endowment. Second, we asked subjects to complete a set of unrelated filler questions. These filler questions created a period over which the subject had ownership of their default position. Having subjects write about their endowment is a common technique in the psychology literature to increase the sense of ownership (Shu and Peck, 2011) and elongating the time of having ownership of one's

endowment has been shown to increase the endowment effect (Strahilevitz and Loewenstein, 1998). Moreover, while completing the filler questions,<sup>10</sup> we reminded subjects of their default position by showing an image of their endowment to further reinforce the ownership of the default option they were given.

After completing the filler tasks, we asked subjects whether they would like to swap their position. Subjects assigned to the Default Cash treatment were asked if they wanted to give back their \$1 to make a \$1.50 donation to CARE while subjects assigned to the default donation treatment were asked if they wanted to not make the \$1.50 donation to get \$1 in cash. Figures A1c and A1d display the decisions faced by the subjects from the Default Cash and Default Charity treatments, respectively. When subjects made their Round 1 choice, they were unaware that there would be a Round 2 choice and we expect that their choices in Round 1 may have differed if they anticipated a Round 2 donation solicitation (Adena and Huck, 2019).

Next, in Round 2, we presented subjects with a multiple price list in which they had to choose one of 11 options. For each item, they could choose to add  $\$X = (0, 0.10, 0.20 \dots 1.00)$  to their bonus and donate  $\$2 \times (1-X)$  to Save the Children (see Table A1). For example, in the first option, subjects could choose to add \$1 to their bonus and donate \$0 to Save the Children, while in the last option, subjects could choose to add \$0 to their bonus and donate \$2 to Save the Children. Subjects had to select one option from the list. We chose a new charity for the Round 2 decision to avoid a potential charity-specific wealth effect; that is, if some subjects donated to CARE in Round 1 (and others did not), then the marginal utilities of donating to CARE in Round 2 could differ by treatment assignment.

After completing the two rounds of decisions, we asked a brief series of demographic questions as well as questions about their past charitable giving behaviour. We summarize and discuss these statistics below in Table 1.

## 2.2 Delay Treatments

We also conducted a follow-up experiment in which we increased the length of time between the Round 1 and Round 2 decisions to one week, henceforth: *the Delay Treatment*. Participants were aware that this was a two-part study when they agreed to participate. The Delay Treatment is *identical* to the main Choice Treatment except that the experiment is divided into two parts that were separated by one week, where the Round 1 decision took place in Part 1 and the Round 2 decision took place during Part 2 of the study. The procedures for Part 1 in the delay experiment are identical to the original experiment without

---

<sup>10</sup>Please see the full experimental protocol here to see the filler tasks the subjects performed.

the delay. Specifically, participants were first informed of their endowment (\$1.00 cash in the Default Cash Delay or that a \$1.50 donation to CARE would be made on their behalf in the Default Charity Delay), then the same questions about how the participant believed they would use their endowment or how CARE would use the donation (depending on the treatment assignment), followed by the same filler tasks and finally the choice to either keep their endowment or switch (i.e., the Round 1 decision). Part 1 concluded with participants being asked to return one week later to complete the experiment. When participants return for Part 2, they are reminded of the choice they made in Round 1 and then asked to make the Round 2 donation decision from the same multiple price list as the Choice Treatment. We then collected the demographic information.

## 2.3 Data

The data from Main Choice treatments was run in May and November 2018 using Amazon Mechanical Turk (Mturk) workers from the United States who have HIT approval rate greater than 99% and have had more than 10000 HITs approved. The Delay Treatment was completed in June 2021 by 673 Mturk workers with the same criteria as the primary experiment.<sup>11</sup>

Table 1 displays the summary statistics for the subjects in our primary experiment.<sup>12</sup> Approximately half of the subjects are female, the majority work full-time and 77% have donated to charity at least once in the last year. To proxy for each subject’s identity towards altruism, we asked whether they had donated money to a charity 0, 1, 2, 3 or 4 or more times in the past year (not including the donation made during the experiment). Using this variable, we classify subjects as having a strong conviction towards altruism if they indicated that they have given 4 or more times in the past year. Otherwise, a subject is classified as having a weak identity towards altruism.

## 2.4 Model, Hypotheses and Empirical Strategy

Next, we turn to modeling the choice to donate at  $t$ , given previous donation choices, to examine whether altruism begets altruism. To formalize this question, we allow for habit

---

<sup>11</sup>With the growing use of online subject pools in social sciences research, the question of how these online pools differ from our standard university-based subject pools, is an important question. (Snowberg and Yariv, 2021) compare elicited choices across several important dimensions from an Mturk subject pool, a university-based subject pool and a representative sample and find a high level of correlation across all three subject pools. However, they also find that there is more noise in the data generated by the Mturk subject pools than in student-based subject pool (see also Gupta, Rigott, and Wilson (2021)).

<sup>12</sup>See Table A2 for descriptive statistics collected in the Delay Treatment.

TABLE 1: SUMMARY STATISTICS

	Treatment Conditions	
	Default Charity	Default Cash
Altruism Strongly Held Value	.38 (.47)	.31 (.49)
Female .52	.54 (.50)	.50 (.50)
Age	40.51 (11.82)	38.38 (11.56)
Unemployed	.05 (.22)	.09 (.29)
Employed full-time	.64 (.48)	.58 (.49)
Employed part-time	.14 (.35)	.18 (.38)
Retired	.05 (.21)	.04 (.20)
Income < \$10,000	.07 (.25)	.05 (.23)
Income > \$150,000	.04 (.19)	.02 (.15)
Observations	191	224

Means reported with standard deviations in parentheses.

persistent or path-dependent preferences (Pollak, 1970). These preferences allow for changes in past consumption to directly affect an individual’s optimal consumption profile today. This model is specifically well-suited to our question because our experimental design allows us to directly estimate the “persistence” parameter for which we obtain a closed-form solution from the first-order conditions of the model described below. In this section, we will derive the relevant first-order condition of our habit persistence model and show that a positive persistence parameter is consistent with the idea that “altruism begets altruism.”

We consider an individual who has preferences over two goods at time  $t$ , private consumption ( $c_t$ ) and charitable giving ( $A_t$ ). The individual’s preferences can be represented by a utility function with the following form,

$$U(c, A) = u(c_t, c_{t-1}) + \alpha v(A_t, A_{t-1}) \tag{1}$$

where  $\alpha \in [0, 1]$  governs the intensity of the individual’s preference for altruism and warm glow. Thus, today’s utility depends on the choices the individual makes today as well as the choices made in  $t - 1$ . The functions  $u(\cdot)$  and  $v(\cdot)$  are concave in consumption and donations to charity, respectively. At time  $t$ , a subject chooses  $(c_t, A_t)$  taking their previous choices,  $(c_{t-1}, A_{t-1})$ , as given.

$$\max_{c_t, A_t} U(c_t, A_t \mid \bar{c}, \bar{A}) = \max_{c_t, A_t} u(c_t - \gamma_c c_{t-1}) + \alpha v(A_t - \gamma_A A_{t-1}) \text{ subject to } I = c_t + p \times A_t \tag{2}$$

where the parameter  $\gamma_c$  and  $\gamma_A \in \mathbb{R}$  represent the intensity of the past consumption choices  $(c_{t-1}, A_{t-1})$  on today's utility and will pin down whether there is negative, positive or no path-dependence.  $I$  is income and  $p$  is the relative price of making a donation. We want to compare the optimal choices at time  $t$  of individuals nudged towards altruism versus subjects nudged towards selfishness at time  $t - 1$ . Let  $A_t(Z)$  and  $A_{t-1}(Z)$  represent the choices at  $t$  and  $t - 1$ , respectively, for an individual who receives nudge  $Z \in 0, 1$ , where  $Z = 1$  indicates the subject was nudged towards altruism and  $Z = 0$  indicates the individual was nudged towards selfishness. From the first order conditions we find that

$$A_t(Z = 0) - \gamma_A A_{t-1}(Z = 0) = A_t(Z = 1) - \gamma_A A_{t-1}(Z = 1) \quad (3)$$

Rearranging and taking expectations of equation 9, we obtain

$$\frac{E[A_t | Z = 1] - E[A_t | Z = 0]}{E[A_{t-1} | Z = 1] - E[A_{t-1} | Z = 0]} = \gamma_A \quad (4)$$

The left-hand-side of equation 10 tells whether donation choices at  $t - 1$  positive or negatively affect the optimal donation choice at time  $t$ . Further, equation 10 is the instrumental variable estimand,  $\beta^{IV}$ . Thus, we propose to test for positive or negative spillovers by estimating the local average treatment effect (Imbens and Angrist, 1994) using instrumental variables (Angrist, Imbens, and Rubin, 1996).

Our main hypothesis relies of three assumptions. First, the instrument,  $Z$ , is randomly assigned. We satisfy this assumption in our experimental design. Second, the effect of the instrument,  $Z$ , must be monotonic in that a subject in the Default Charity condition must be at least as likely to donate in Round 1 than he would have been had he been assigned to the Default Cash condition. The monotonicity assumption is related to the denominator of equation 10, which is the first stage of our IV estimate. Further, there is ample evidence that default option nudges are effective (Benartzi and Thaler, 2007; Choi et al., 2003; Cronqvist and Thaler, 2004; Madrian and Shea, 2001; Kessler and Roth, 2012, 2014). Thus, we hypothesize that, on average, subjects in the Default Charity condition will be more likely to donate in Round 1 then subjects in the Default Cash condition; that is,  $E[A_{t-1} | Z = 1] - E[A_{t-1} | Z = 0] > 0$ . We test and provide support for this hypothesis in Section 3.1.

**Hypothesis 1. *Default Option Hypothesis:*** *Participants who are defaulted into making a donation are more likely to donate in Round 1 than participants who are defaulted into keeping cash.*

Third, given the large experimental literature that find *positive* spill overs in pro-social

behaviour (Shang and Croson, 2009; Cairns and Slonim, 2011; Gneezy et al., 2012; Cassar, d’Adda, and Grosjean, 2014; Peysakhovich and Rand, 2016; Castillo, Petrie, and Samek, 2017; Engl, Riedl, and Weber, 2018), we hypothesize that subjects in the Default Charity condition will *also* be more likely to donate in Round 2 than subjects in the Default Cash condition; that is,  $E[A_t | Z = 1] - E[A_t | Z = 0] > 0$

**Hypothesis 2. *Positive Spill Over Hypothesis:*** *Subjects in the Default Charity treatment,  $Z = 1$ , will be more likely to donate in Round 2 than subjects in the Default Cash condition,  $Z = 0$ .*

Finally, we turn to the main hypothesis about the sign of  $\gamma_A$ . As a direct consequence of hypotheses 1 and 2, we hypothesize that  $\gamma_A > 0$ , implying positive path-dependence. We interpret a positive path-dependence as moral consistency since  $\gamma_A > 0$  implies that the Default Charity condition exogenously increases altruism in Round 1 and that this nudge-induced increase in altruism in Round 1 causes an increase in altruism in Round 2.<sup>13</sup>

**Hypothesis 3. *Moral Consistency Hypothesis:***  $\beta^{IV} > 0$ , *implying that  $\gamma_A > 0$  which means that an increased propensity to choose to donate in Round 1 will:*

- (i) *increase the propensity to donate in Round 2;*
- (ii) *increase the amount donated in Round 2.*

To analyze the experimental data, we estimate a two-stage least squares instrumental variable regression in which standard errors are adjusted to account for the two-stage estimation procedure.<sup>14</sup> In the first stage, we estimate the effect of the treatment assignment,  $Z_i$ , on Round 1 donation behaviour,  $A_{i,t-1}$ .

$$A_{i,t-1} = \rho_0 + \rho_1 \mathbf{1}[Z_i = 1] + \rho_2 \mathbf{X} + \nu_i, \quad (5)$$

where  $\mathbf{X}$  is a vector of demographic controls, including gender, employment status, and income. In the second stage, we use the predicted values of Round 1 donation behaviour,  $\widehat{A_{i,t-1}}$ , to obtain the causal effect of donating in Round 1 on donating in Round 2,  $Y_i$ , and adjust the standard errors to account for the stage-stage estimation procedure. The interpretation of the coefficient,  $\beta^{IV}$ , is the change in Round 2 donation rates that are caused by the *treatment-induced* donation choice in Round 1.

$$A_{i,t} = \beta_0 + \beta^{IV} \widehat{A_{i,t-1}} + \beta_2 \mathbf{X} + \varepsilon_i, \quad (6)$$

<sup>13</sup>In Appendix B we consider immoral consistency, that is whether keeping the money in Round 1 causes people to behave more selfishly in Round 2.

<sup>14</sup>Because our model is just-identified, the two stage least square estimator and the limited-information maximum likelihood estimator are equivalent. We thus only report the two-stage least squares model.

We also explore heterogenous effects stemming from identity towards altruism, the details of which are found in Section 3.2.1 and Appendix B. We continue to draw from self-perception theory by positing that identity is inferred from past choices. Thus, the composite of past charitable giving,  $\Theta_A$ , serves as a proxy for the facet of identity related to altruism. In our model, it is straightforward to show that  $\frac{\partial \gamma_A}{\partial \Theta_A} <$ , thus moral consistency is *decreasing* in the strength of the individual’s altruistic identity ( $\Theta_A$ ).

### 3 Results

Our results are consistent with self-perception theory and moral consistency; that is, exogenously-induced altruism in one period causes an increase in altruism in a subsequent period. In other words, altruism begets altruism. To establish our main result, we show that Hypotheses 1 and 2 hold, result 1 and result 2, respectively. We then establish our main result, that altruism begets altruism, result 3, in Section 3.2. We then show heterogeneous effects of our result in Section ?? and that moral consistency is robust to a longer delay between Round 1 and Round 2 in Section 3.2.2. Finally, using the data collected from the No Choice Treatments, we provide important evidence regarding the exclusion restriction to support our identification strategy in Section 3.3. For each result, we begin with a statement of the result, followed by the evidence to support it.

#### 3.1 Support for Hypotheses 1 and 2

**Result 1.** *Our nudge significantly affects Round 1 donation choices. Subjects who are defaulted into giving to charity are significantly more likely to choose to donate in Round 1 than subjects who are defaulted to keep a utility-equivalent amount of cash.*

In Table 2, we show that the Default Charity treatment increases the propensity to give in Round 1 by 19 percentage points compared to the Default Cash treatment. Importantly, Table 2 also provides evidence of a strong first stage—the  $F$ -statistic in our first stage is 18.96, which exceeds the “rule of thumb” for a relevant instrument proposed by Staiger and Stock (1997).<sup>15</sup>

**Result 2.** *Consistent with hypothesis 2, subjects in the Default Charity condition are significantly more likely to give in Round 2 than subjects in the Default Cash condition.*

---

<sup>15</sup>In column (2), the  $F$ -statistic drops due to the inclusion of additional control variables that have little predictive power in the Round 1 donation behaviour.

TABLE 2: RELEVANCE OF INSTRUMENT: EFFECT OF DEFAULT TREATMENT ON ROUND 1 DONATION RATES

	(1)	(2)
Default Charity	0.19*** (0.04)	0.18*** (0.04)
Constant	0.17*** (0.02)	0.13** (0.06)
Observations	415	415
$R^2$	0.05	0.05
$F$ statistic	18.96	3.57
Demographic Controls	No	Yes

OLS regression estimates. Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for the effect of default charity nudge, the test of the hypothesis on the coefficient for Default Charity are one-tailed.

TABLE 3: ROUND 2 DONATION RATES & AMOUNTS

	Propensity to Donate		Donation Amount	
	(1)	(2)	(3)	(4)
Default Charity	0.08* (0.05)	0.07* (0.05)	0.11* (0.07)	0.09* (0.07)
Constant	0.32*** (0.03)	0.05 (0.1)	0.37*** (0.04)	0.005 (0.15)
Observations	415	415	415	415
$R^2$	0.006	0.04	0.006	0.04
Demographic Controls	No	Yes	No	Yes

OLS regression estimates. Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Table B4 presents the analogous results for immoral consistency. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for Hypothesis 4, the estimates on the coefficient for Default Charity are one-tailed.

## 3.2 Main Results: Self-Perception & Moral Consistency

**Result 3.** *Consistent with Hypothesis 3, we find evidence of moral consistency, i.e.,  $\beta^{IV} > 0$ , implying that  $\gamma_A > 0$ ; that is, choosing to give in Round 1 increases giving in Round 2. In particular, choosing to give in Round 1 increases the probability of giving in Round 2 by 200% (40 percentage points) and increases the amount given in Round 2 by \$0.59.*

In Table 4, we estimate the effect that the nudge-induced increase in choosing to give in Round 1 has on charitable giving in Round 2 using the instrumental variable approach discussed in Section 2.4. Column (1) and column (2) indicates that giving in Round 1 causes a 41 percentage point and 44 percentage point (200% increase above the baseline) increase in the propensity to give in Round 2, respectively. In columns (3) and (4) the dependent variable is donation amount in Round 2. Column (3) and (4) indicate that giving in Round 1 causes subjects to increase their giving by \$0.59 (200%) in Round 2. In sum, altruism begets altruism.

TABLE 4: MORAL CONSISTENCY: ROUND 2 DONATION RATES & AMOUNTS

	Propensity to Donate		Donation Amount	
	(1)	(2)	(3)	(4)
$\widehat{A}_{i,1}$	0.41** (0.23)	0.44** (0.23)	0.59** (0.32)	0.62** (0.33)
Constant	0.25*** (0.06)	0.28*** (0.08)	0.27*** (0.08)	0.31*** (0.11)
Observations	415	415	415	415
$R^2$	0.2	0.21	0.21	0.22
Demographic Controls	No	Yes	No	Yes

T

wo-stage least square regression estimates. Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Table B4 presents the analogous results for immoral consistency. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for the effect of moral consistency, the estimates on the coefficient for  $\widehat{A}_{i,1}$  are one-tailed.

The results in Table 4 suggest that nudging virtuous behaviour “now” may promote virtuous behaviour “later”. In other words, the nudge successfully crowds people into giving in Round 2, who would likely not have given in Round 2, by nudging them to give in Round 1.

### 3.2.1 Heterogenous Effects: Identity towards Altruism

In this section, we explore the heterogenous effects of identity on moral consistency. In Section Appendix B we augment our original model to include a term that represents past giving

behaviour,  $\Theta_A$ , which we use as a proxy for identity towards altruism. It is straightforward to show that  $\frac{\partial \gamma_A}{\partial \Theta_A} < 0$ , that is, moral consistency is *decreasing* in the strength of the individual's altruistic identity ( $\Theta_A$ ). Thus, the magnitude of the local average treatment effect,  $\beta^IV$ , will be greater for those with a weak identity than for those with a strong identity towards altruism.<sup>16</sup>

TABLE 5: MORAL CONSISTENCY & IDENTITY: ROUND 2 DONATION RATES & AMOUNTS

Propensity to Donate	Donation Amount			
	(1)	(2)	(3)	(4)
$\widehat{A}_{i,1} \times StrongIdentity$	-0.18 (0.39)	-0.14 (0.39)	0.04 (0.53)	0.1 (0.53)
$\widehat{A}_{i,1} \times WeakIdentity$	0.83*** (0.36)	0.85*** (0.35)	0.95*** (0.49)	0.95*** (0.47)
Strong Identity	0.36** (0.15)	0.36** (0.15)	0.37* (0.21)	0.37* (0.2)
Constant	0.14* (0.08)	0.17* (0.09)	0.15 (0.1)	0.21* (0.13)
Observations	415	415	415	415
$R^2$	0.03	0.04	0.12	0.14
Demographic Controls	No	Yes	No	Yes
$\chi^2$ test				
$\widehat{A}_{i,1} \times Strong = \widehat{A}_{i,1} \times Weak$	3.59	3.58	1.58	1.44

OLS regression estimates. Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Table B4 presents the analogous results for immoral consistency. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for the effect of identity on moral consistency, the estimates on the coefficient for  $\widehat{A}_{i,1} \times Weak$  and  $\widehat{A}_{i,1} \times Strong$  are one-tailed.

In columns (1) and (2) of Table 5 our dependent variable is the probability of donating in Round 2 and consistent with our model we find that subjects for whom altruism is a weakly held value are significantly more morally consistent than subjects for whom altruism is a strongly held value. However, in columns (3) and (4) our dependent variable is donation amount in Round 2 and we find that the difference between strong and weak altruists, while large in effect size, not significant.

One potential concern is that an individual who has only given to charity once or twice in the past year, but gave a large sum of money, would be classified as having a weak identity

<sup>16</sup>Benabou & Tirole's (2011) model, also drawing heavily from self-perception theory, predicts that when weakly-held values are encouraged, individuals respond in a confirmatory way (i.e., morally consistent). In other words, individuals for whom altruism is a weak facet of their identity will behave in a morally consistent way later, when nudged towards altruism now. Thus, the predictions from our model and the Bénabou and Tirole (2011) model about individuals with weak identities towards altruism are similar when  $\gamma_A > 0$ .

towards altruism under our definition. To address this potential problem, in Table B3, we re-classify a weak identity towards altruism as those who report giving 0 times in the past year. Again, we find that they are significantly more morally consistent than subjects who have given 4 or more times in the past year.

### 3.2.2 Moral Consistency with Delay

TABLE 6: RELEVANCE OF INSTRUMENT IN DELAY TREATMENTS: EFFECT OF DEFAULT TREATMENT ON ROUND 1 DONATION RATES

	(1)	(2)
Default Charity Delay	0.15*** (0.03)	0.15*** (0.04)
Constant	0.2*** (0.02)	-0.04 (0.09)
Observations	673	596
$R^2$	0.03	0.06
$F$ statistic	20.52	5.01
Demographic Controls	No	Yes

OLS regression estimates. Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for the effect of default charity nudge, the test of the hypothesis on the coefficient for Default Charity are one-tailed.

In Table 6, we present the first stage results for our Delay Treatment, which shows that the Default Charity treatment assignment is a relevant instrument for donation behaviour in Round 1.

**Result 4.** *We continue to find evidence of moral consistency, i.e.,  $\beta^{IV} > 0$ , implying  $\gamma_A > 0$ , with a one-week delay between Round 1 and Round 2 donation choices.*

In Table 7, we present our main result of the Delay Treatment. As in Table 4, we estimate the causal effect of the Round 1 choice on Round 2 choices by instrumenting for Round 1 behaviour with the Default Charity treatment assignment. One important difference in the Delay Treatment is that we need to take into account the potential for bias due to attrition of participants from Part 1 to Part 2 when measuring Round 2 donation behaviour. Specifically, 89% of our subjects who completed Part 1 returned a week later to complete Part 2. However, we found that participants assigned to the Default Charity Delay treatment were significantly less likely to return for Part 2 than subjects in the Default Cash Delay treatment (Default Charity Delay 86% returned versus Default Cash Delay 91% returned: t-statistic=1.78, p-value=.08). To account for this, we first estimate the probability of returning for Part 2

TABLE 7: MORAL CONSISTENCY WITH DELAY: ROUND 2 DONATION RATES & AMOUNTS

	Propensity to Donate		Donation Amount	
	(1)	(2)	(3)	(4)
$\widehat{A}_{i,1}$	0.33** (0.18)	0.22 (0.22)	0.65*** (0.22)	0.63*** (0.28)
Inverse Mills Ratio	-0.71** (0.29)	-0.17 (0.33)	-1.18*** (0.37)	-0.81* (0.42)
Constant	0.5*** (0.08)	0.25** (0.12)	0.55*** (0.1)	0.37*** (0.14)
Observations	673	596	673	596
$R^2$	0.19	0.2	0.33	0.38
Demographic Controls	No	Yes	No	Yes

Two-stage least square regression estimates that adjust for selection. Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for the effect of moral consistency, the estimates on the coefficient for  $\widehat{A}_{i,1}$  are one-tailed.

using the variables we collected in Part 1 (see Table A2), compute the inverse mills ratio and include the inverse mills ratio in our two-stage least squares instrumental variable regression (Wooldridge, 2015). We present these results in Table 7. Similar to our results in Table 4, we find that donating in Round 1 leads to a 25-35 percentage point increase in the probability of donating in Round 2 and \$0.64 (or a nearly 200%) increase in the amount donated in Round 2.

### 3.3 Exclusion Restriction & The Importance of Choice

Self-perception theory argues that an active choice is an important component linking behaviour and identity, noting that rejected alternatives reinforce the inferences an individual can make from their choice about their identity (Bem, 1972; Zanna, 1972). For example, an individual learns less about his altruistic identity if he is forced to donate rather than having had an active choice to keep the money for himself. One final assumption, the exclusion restriction assumption, is needed for valid identification of Hypothesis 3. The exclusion restriction states that the instrument ( $Z$ ) only affects outcome  $A_t$  through  $A_{t-1}$  (i.e., Round 1 donation behaviour ) and does not directly affect outcomes,  $A_t$ . In the language of and consistent with self-perception theory, this requires that it is the *choice* of donating in Round 1, which is influenced by the subject’s default position (i.e., the instrument,  $Z$ ), rather than the default position itself that affects Round 2 donation decisions (i.e., outcomes,  $A_t$ ). This

assumption is the hardest to justify without some evidence and is the motivation for our No Choice treatments that we will describe now.

### 3.3.1 No Choice Treatments

In the No Choice Treatments, subjects are assigned to either the Default Charity or Default Cash conditions, and follow the exact procedures as in the Choice Treatments, including the ownership questions and filler tasks, and are then asked to make the same Round 2 donation decision as the Choice treatments. The *only* difference between the No Choice treatments and the Choice treatments is that subjects in the No Choice treatments are not given the opportunity to switch their Round 1 default position. Thus, subjects in the Default Charity (No Choice) treatment are forced to make a donation in Round 1, while subjects in the Default Cash (No Choice) treatment are forced to keep the cash in Round 1.

For robustness and to isolate the importance of *choice* in driving spill-overs, we ran three variations of the No Choice treatments: (1) **Hypothetical Choice** in which they are asked to indicate what their Round 1 Choice would have been if given the opportunity; (2) **Hypothetical Scenario** in which they are asked to reflect upon, but not indicate, what their Round 1 choice would have been if given the opportunity; and (3) **No Information** in which they are not informed of any alternative to their assigned default position. The Hypothetical Choice treatment is identical to our Main treatment *except* for the choice is not executed and is thus not an active choice.

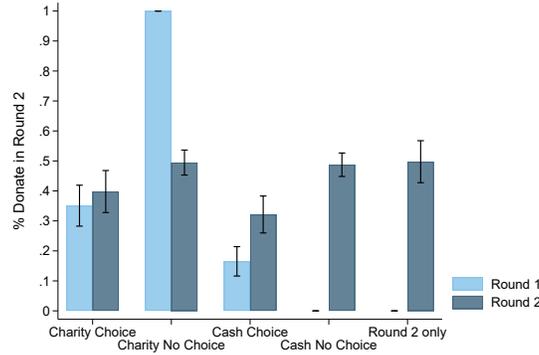
In a final control treatment (henceforth: Round 2 Only), subjects do not make a Round 1 decision and instead begin the experiment with the filler tasks and are then asked to make a Round 2 donation decision that is identical to the other treatments. The Round 2 Only treatment allows us to examine how subjects make Round 2 donation choices when they are only asked to give once.<sup>17</sup>

To demonstrate that *choice*, rather than treatment assignment, is the driving force behind Round 2 behaviour, we hypothesize and show that: (1)  $E[A_t | Z = 1, NoChoice] - E[A_t | Z = 0, NoChoice] = 0$  and (2)  $E[A_t | Z = 1] - E[A_t | Z = 0] > E[A_t | Z = 1, NoChoice] - E[A_t | Z = 0, NoChoice] = 0$ . We test and provide support for this hypothesis in this Section.

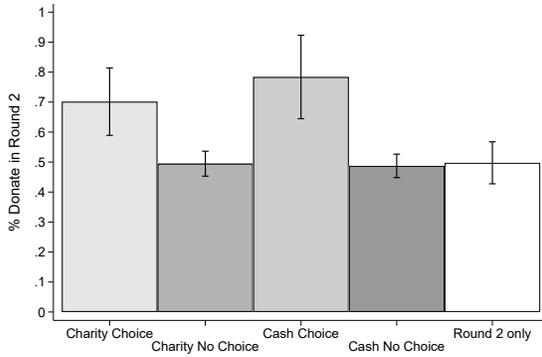
**Hypothesis 4. Exclusion Restriction Hypothesis:** *The default option treatment  $Z$  does not directly affect the decision to donate in Round 2. Instead, any effect of  $Z$  on Round 2 donation choices operates solely through the choice to donate in Round 1.*

<sup>17</sup>We replicate the summary statistics shown in Table 1 for the No Choice Treatments in Table C5.

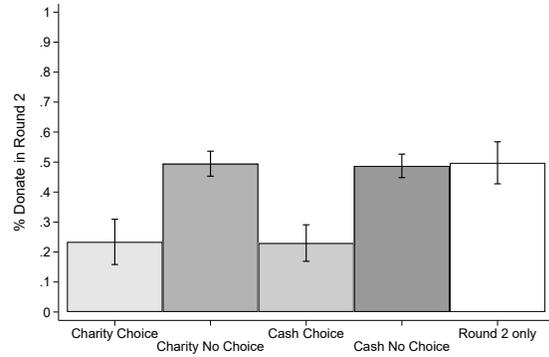
FIGURE 1: AVERAGE DONATION RATES



(A) Round 1 & Round 2



(B) Donated in R1



(C) Did Not Donate in R1

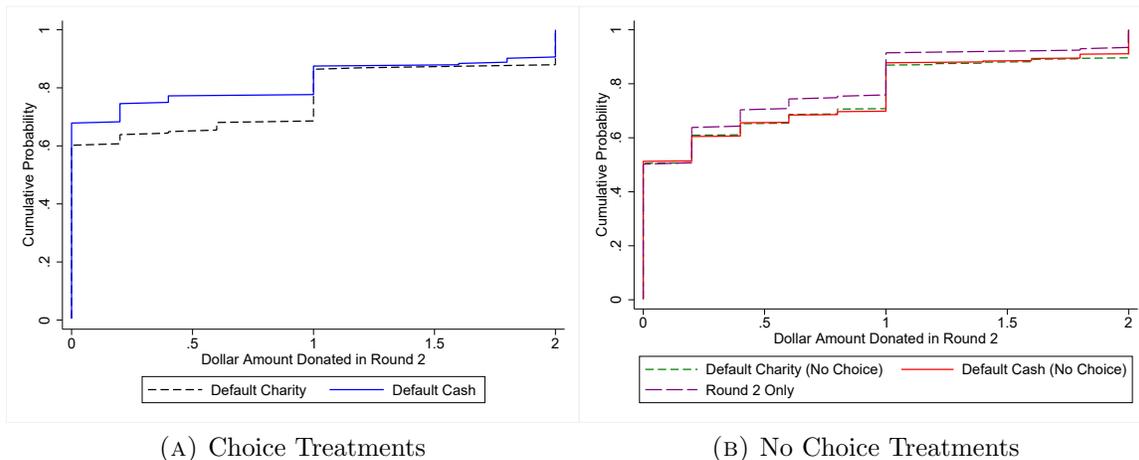
Figure 1a shows the average donation rates in Round 1 and Round 2 across the Choice and No Choice treatments. Figure 1b shows the Round 2 donation rates conditional on donating in Round 1. Figure 1c shows the mean Round 2 donation rates conditional on *not* donating in Round 1.

### 3.3.2 The Role of Round 1 Choice on Round 2 Behaviour

In this section, we will show that there are no spill-overs when we compare the Round 2 behaviour of subjects in the Default Charity No Choice and the Default Cash No Choice. This finding is akin to showing that an exclusion restriction assumption holds; that is, the nudge in the Choice Treatments does not directly affect Round 2 behaviour, but instead the nudge operates through the Round 1 behaviour. In other words, the positive spill overs identified in the Choice treatments in Table 4 and Table 7 are not driven by the nudge but rather by the *active choice* made in Round 1.

Before proceeding to our analyses, we show the average donation rates across all treatments in Figure 1a. In Figures 1b and 1c, we show the Round 2 donation rates *conditional* on

FIGURE 2: DISTRIBUTION OF ROUND 2 DONATION AMOUNTS, BY TREATMENT ASSIGNMENT



Distribution of donation amounts in Round 2 by treatment assignment. Figure 2a shows the distribution of donation amounts for the Choice Treatments and Figure 2b shows the distributions for the No Choice treatments.

donating in Round 1 and not donating in Round 1, respectively. These figures immediately make clear a key feature of our two-period experimental design—what happens in Round 1 affects behaviour in Round 2 and there is a differential effect for the Default Charity and the Default Cash conditions. And just as we cannot compare Round 2 behaviour between the Default Cash and Default Charity treatments without controlling for Round 1 behaviour, we cannot compare Round 2 behaviour between the Choice and No Choice treatments without controlling for Round 1 behaviour. For example, if we want to know how removing choice in Round 1 affects donation rates in Round 2, we need to calculate the difference in *the change in donation rates* from Round 1 to Round 2 between the Default Charity No Choice and the Default Charity Choice (Angrist and Pischke, 2008). Using the averages reported in Figure 1a, we calculate that “no choice” in Round 1 results in a decrease in Round 2 donation rates of 55 percentage points in the Default Charity conditions.<sup>18</sup>

One final important observation about Figures 1a is that the average Round 2 behaviour of subjects in the No Choice treatments is not different from the average Round 2 behaviour in the Round 2 only treatment, where subjects *only* participate in Round 2. One potential interpretation, consistent with self-perception theory (Zanna, 1972), is that subjects in the No Choice treatments do not use their Round 1 behaviour to make inferences about the type of person they are when they are making their Round 2 choice.

<sup>18</sup>We show this in a regression framework in Table C6.

Next, we turn to our main question of the section—does the nudge directly affect Round 2 behaviour or does it operate *through* Round 1 choices. We begin our analysis by first showing the distributions of donation amounts in Round 2 by treatment assignment for the Choice and No Choice treatments in Figure 2. Consistent with our findings in Table 3, Figure 2a shows that the distribution of donations in Round 2 between the Choice Treatments—Default Cash and Default Charity—are significantly different (Kruksall-Wallis one-tailed test:  $\chi^2 = 3.01$ , p-value=0.04). On the other hand, Figure 2b shows that the distributions of donations for the No Choice treatments are strikingly similar and the Kruksall-Wallis test finds no statistical differences between the distributions of Round 2 donation amounts for the No Choice Treatments (One-tailed test:  $\chi^2 = 0.51$ , p-value=0.39). Further, the Round 2 donation behaviour of subjects in the Default Charity and Default Cash No Choice treatments is not statistically different from the Round 2 behaviour of subjects in the Round 2 Only treatment; that is, subjects in the No Choice conditions behave *as if* they never participated in Round 1.

**Result 5.** *We find that the treatment itself has no direct effect on Round 2 choices; that is, subjects in the Default Charity (No Choice) treatments do not behave significantly different in Round 2 than subjects in the Default Cash (No Choice) treatments.*

To bolster our findings from Figures 2, we conduct a series of quantile regressions at the 60th, 65th, 70th, 75th and 80th percentiles to further demonstrate the importance of choice in driving Round 2 behaviour. In Table 8, we pool together the Choice and No Choice treatments, using the Default Charity No Choice condition as the omitted category, and interacts Default Charity and Choice. The coefficient of interest in this table is the interaction term, Default Charity  $\times$  R1 Choice, which shows the *additional* Round 2 donation made by subjects in the Default Charity condition who also had the *choice* to donate in Round 1 (i.e., the Choice treatments). Consistent with Figures 2a and 2b, we find that the interaction is insignificant at the 60th percentile (and not shown, for all quantiles below the 60th percentile, indicating that in the Default Charity and Default Cash Choice treatments all subjects are donating \$0.00, while in the Default Charity and Default Cash No Choice treatments subjects are donating the same amounts—50 percent in both conditions donating \$0.00 and an additional 10% donating \$0.30. However, as we move up the distribution, a gap emerges in the amount given in the Choice conditions whereas no gap occurs in the No Choice conditions. For example, in the Choice conditions (see Figure 2a) 8 percent more subjects give at least \$0.20 in Default Charity (68%) than Default Cash (60%), and approximately 15 percent more subjects give at least \$0.80 in Default Charity (78%) than Default Cash (63%). And then as we move further up the distribution we see that in the

TABLE 8: THE IMPORTANCE OF CHOICE IN DRIVING POSITIVE SPILL OVERS: ROUND 2 DONATION AMOUNTS

	Donation Amount in Round 2					Tobit
	Percentiles					
	60th	65th	70th	75th	80th	
	(1)	(2)	(3)	(4)	(5)	(6)
Default Charity $\times$ Round 1 Choice	00 (0.13)	0.6*** (0.27)	1.00*** (0.26)	0.6*** (0.13)	00 (0.05)	0.29* (0.21)
Default Charity	00 (0.07)	00 (0.14)	-0.2 (0.13)	00 (0.07)	00 (0.02)	0.03 (0.1)
Round 1 Choice	-0.2** (0.09)	-0.4** (0.19)	-0.8*** (0.18)	-0.6*** (0.09)	00 (0.03)	-0.45*** (0.14)
Round 2 Only	00 (0.09)	00 (0.2)	-0.6*** (0.18)	-0.2** (0.09)	00 (0.03)	-0.09 (0.14)
Constant	0.2*** (0.05)	0.4*** (0.1)	1.00*** (0.09)	1.00*** (0.05)	1.00*** (0.02)	-0.1 (0.07)
Observations	1802	1802	1802	1802	1802	1802

Columns (1)-(5) show quantile regression estimates and column (6) shows tobit regression estimates censored below at 0 and above at \$2. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for effects of spill overs, the estimates on the coefficient for Default Charity  $\times$  R1 Choice are one-tailed, all other tests are two-tailed.

Choice Conditions the gap disappears at \$1.00 of giving, where in both conditions we see that about 83% both give at least \$1.00 and almost everyone else gives \$2.00. In contrast, inspection of Figure 2b shows that in the No Choice conditions there is never a gap in the distribution between Default Cash and Default Charity. The interaction estimates presented in Table 8 confirm that the gap between Default Charity and Default Cash in the Choice conditions is significantly larger than the (lack of any) gap in the No Choice conditions at the 65th, 70th and 75th percentiles. In column (5), we show the results from a Tobit regression model over the entire distribution of donations and we find qualitatively similar, but smaller and less precise results—the Default Charity is only effective at increasing donations in Round 2 relative to Default Cash when there was a *choice* in Round 1.

## 4 Conclusion

In this paper, we conducted an experiment to provide evidence that altruism begets altruism. We estimate a local average treatment effect, which is directly informed by self-perception

and our model of history dependence. We go beyond the existing literature and show that an increase in choosing to be altruistic now *causes* an increase in altruistic behaviour later, which we interpret as moral consistency.

Moral consistency helps to overcome decreases in giving that are typically associated with ask fatigue and multiple donation solicitations. On average, subjects in the Default Charity condition give \$0.53 in Round 1 and \$0.48 in Round 2 for a total average donation of \$1.01. Using the donation rates from the Round 2 Only condition, we know the average donation amount is \$.41 if an individual is only asked to give once. Thus, if there was no moral consistency and the Round 1 and Round 2 decisions were instead independent, then subjects in the Default Charity condition would donate 15% less. Of course, this difference is magnified more if we consider those subjects who have a weak identity towards altruism. When Round 1 and Round 2 decisions are linked through moral consistency, individuals in the Default Charity condition with a weak identity towards altruism donate a total of \$0.98 to charity. However, if Round 1 and Round 2 decisions were independent, then these same subjects would donate 27% less.

We also believe the findings in this paper generate interesting questions for future research. For example, one interesting question for future research may study whether different types of nudges or a longer length of time between asks result in similar patterns of moral consistency. We obtain exogenous variation in our Round 1 giving by using a default option nudge, but studying whether reminding individuals about social norms aRound giving, also a popular nudge, also generates moral consistency would be of great academic and practical interest.

## References

- Adena, Maja and Steffen Huck. 2019. "Giving once, giving twice: A two-period field experiment on intertemporal crowding in charitable giving." *Journal of Public Economics* 172:127–134.
- Akerlof, George A and Rachel E Kranton. 2005. "Identity and the Economics of Organizations." *Journal of Economic perspectives* 19 (1):9–32.
- Andreoni, James and Marta Serra-Garcia. 2021. "Time inconsistent charitable giving." *Journal of Public Economics* 198:104391.
- Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association* 91 (434):444–455.
- Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics*. Princeton university press.
- Anik, Lalin, Lara B Akinin, Michael I Norton, and Elizabeth W Dunn. 2009. "Feeling good about giving: The benefits (and costs) of self-interested charitable behavior." *Harvard Business School Marketing Unit Working Paper* (10-012).

- Bem, Daryl J. 1972. "Self-perception theory." In *Advances in experimental social psychology*, vol. 6. Elsevier, 1–62.
- Bénabou, Roland and Jean Tirole. 2011. "Identity, morals, and taboos: Beliefs as assets." *The Quarterly Journal of Economics* 126 (2):805–855.
- Benartzi, Shlomo and Richard H Thaler. 2007. "Heuristics and biases in retirement savings behavior." *The journal of economic perspectives* 21 (3):81–104.
- Benjamin, Daniel J, James J Choi, and Geoffrey W Fisher. 2010. "Religious identity and economic behavior." Tech. rep., National Bureau of Economic Research.
- Blanken, Irene, Niels van de Ven, and Marcel Zeelenberg. 2015. "A meta-analytic review of moral licensing." *Personality and Social Psychology Bulletin* 41 (4):540–558.
- Cairns, Jason and Robert Slonim. 2011. "Substitution effects across charitable donations." *Economics Letters* 111 (2):173–175.
- Cappelen, Alexander W, Gary Charness, Mathias Ekström, Uri Gneezy, and Bertil Tungodden. 2017. "Exercise improves academic performance." .
- Cassar, Alessandra, Giovanna d’Adda, and Pauline Grosjean. 2014. "Institutional quality, culture, and norms of cooperation: Evidence from behavioral field experiments." *The Journal of Law and Economics* 57 (3):821–863.
- Castillo, Marco, Ragan Petrie, and Anya Samek. 2017. "Time to Give: A Field Experiment on Intertemporal Charitable Giving." .
- Choi, James J, David Laibson, Brigitte C Madrian, and Andrew Metrick. 2003. "Optimal defaults." *The American Economic Review* 93 (2):180–185.
- Clotfelter, Charles T. 1980. "Tax incentives and charitable giving: Evidence from a panel of taxpayers." *Journal of Public Economics* 13 (3):319–340.
- Cronqvist, Henrik and Richard H Thaler. 2004. "Design choices in privatized social-security systems: Learning from the Swedish experience." *The American Economic Review* 94 (2):424–428.
- Damgaard, Mette Trier and Christina Gravert. 2018. "The hidden costs of nudging: Experimental evidence from reminders in fundraising." *Journal of Public Economics* 157:15–26.
- Eckel, Catherine C, David H Herberich, and Jonathan Meer. 2018. "It’s not the thought that counts: A field experiment on gift exchange and giving at a public university." *The Economics of Philanthropy* .
- Engl, Florian, Arno Riedl, and Roberto A Weber. 2018. "Spillover effects of institutions on cooperative behavior, preferences, and beliefs." .
- Exley, Christine L. 2015a. "Excusing selfishness in charitable giving: The role of risk." *The Review of Economic Studies* 83 (2):587–628.
- . 2015b. "Using charity performance metrics as an excuse not to give." Tech. rep., Working paper.
- Frey, Bruno S and Stephan Meier. 2004. "Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment." *American Economic Review* 94 (5):1717–1722.
- Gee, Laura and Jonathan Meer. 2019. "The Altruism Budget: Measuring and Encouraging Charitable Giving." *NBER Working Paper* (w25938).

- Gneezy, Ayelet, Alex Imas, Amber Brown, Leif D Nelson, and Michael I Norton. 2012. “Paying to be nice: Consistency and costly prosocial behavior.” *Management Science* 58 (1):179–187.
- Gneezy, Uri, Elizabeth A Keenan, and Ayelet Gneezy. 2014. “Avoiding overhead aversion in charity.” *Science* 346 (6209):632–635.
- Gupta, Neeraja, Luca Rigott, and Alistair Wilson. 2021. “The Experimenters’ Dilemma: Inferential Preferences over Populations.” *arXiv preprint arXiv:2107.05064* .
- Huck, Steffen, Imran Rasul, and Andrew Shephard. 2015. “Comparing charitable fundraising schemes: Evidence from a natural field experiment and a structural model.” *American Economic Journal: Economic Policy* 7 (2):326–69.
- Imbens, Guido W and Joshua D Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62 (2):467–475.
- Karlan, Dean and John A List. 2007. “Does price matter in charitable giving? Evidence from a large-scale natural field experiment.” *American Economic Review* 97 (5):1774–1793.
- Kessler, Judd B and Katherine L Milkman. 2016. “Identity in charitable giving.” *Management Science* 64 (2):845–859.
- Kessler, Judd B, Katherine L Milkman, and C Yiwei Zhang. 2019. “Getting the rich and powerful to give.” *Management Science* 65 (9):4049–4062.
- Kessler, Judd B and Alvin E Roth. 2012. “Organ allocation policy and the decision to donate.” *The American Economic Review* 102 (5):2018–2047.
- . 2014. “Getting more organs for transplantation.” *The American Economic Review* 104 (5):425–430.
- Khan, Uzma and Ravi Dhar. 2006. “Licensing effect in consumer choice.” *Journal of Marketing Research* 43 (2):259–266.
- Landry, Craig E, Andreas Lange, John A List, Michael K Price, and Nicholas G Rupp. 2010. “Is a donor in hand better than two in the bush? Evidence from a natural field experiment.” *American Economic Review* 100 (3):958–83.
- List, John A and David Lucking-Reiley. 2002. “The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign.” *Journal of political Economy* 110 (1):215–233.
- Madrian, Brigitte C and Dennis F Shea. 2001. “THE POWER OF SUGGESTION: INERTIA IN 401 (k) PARTICIPATION AND SAVINGS BEHAVIOR.” *Quarterly Journal of Economics* 116 (4).
- Meer, Jonathan. 2013. “The habit of giving.” *Economic Inquiry* 51 (4):2002–2017.
- Meer, Jonathan and Benjamin A Priday. 2019. “Tax Prices and Charitable Giving: Projected Changes in Donations Under the 2017 TCJA.” Tech. rep., National Bureau of Economic Research.
- Meier, Stephan. 2007. “Do subsidies increase charitable giving in the long run? Matching donations in a field experiment.” *Journal of the European Economic Association* 5 (6):1203–1222.
- Monin, Benoit and Dale T Miller. 2001. “Moral credentials and the expression of prejudice.” *Journal of personality and social psychology* 81 (1):33.
- Mullen, Elizabeth and Benoît Monin. 2016. “Consistency versus licensing effects of past moral behavior.” *Annual review of psychology* 67.

- Nisan, Mordecai and Gaby Horenczyk. 1990. "Moral balance: The effect of prior behaviour on decision in moral conflict." *British journal of social psychology* 29 (1):29–42.
- Nisan, Mordechai. 1985. "Limited morality: A concept and its educational implications." *Moral education: Theory and application* :403–420.
- Peysakhovich, Alexander and David G Rand. 2016. "Habits of virtue: Creating norms of cooperation and defection in the laboratory." *Management Science* 62 (3):631–647.
- Ploner, Matteo and Tobias Regner. 2013. "Self-image and moral balancing: An experimental analysis." *Journal of Economic Behavior & Organization* 93:374–383.
- Pollak, Robert A. 1970. "Habit formation and dynamic demand functions." *Journal of political Economy* 78 (4, Part 1):745–763.
- Sachdeva, Sonya, Rumen Iliev, and Douglas L Medin. 2009. "Sinning saints and saintly sinners the paradox of moral self-regulation." *Psychological science* 20 (4):523–528.
- Shang, Jen and Rachel Croson. 2009. "A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods." *The Economic Journal* 119 (540):1422–1439.
- Shu, Suzanne B and Joann Peck. 2011. "Psychological ownership and affective reaction: Emotional attachment process variables and the endowment effect." *Journal of Consumer Psychology* 21 (4):439–452.
- Snowberg, Erik and Leeat Yariv. 2021. "Testing the waters: Behavior across participant pools." *American Economic Review* 111 (2):687–719.
- Staiger, Douglas and James H Stock. 1997. "Instrument variables regression with weak instruments." *Econometrica* 65 (3):557–586.
- Strahilevitz, Michal A and George Loewenstein. 1998. "The effect of ownership history on the valuation of objects." *Journal of consumer research* 25 (3):276–289.
- Sunstein, Cass and Richard Thaler. 2008. "Nudge." *The politics of libertarian paternalism. New Haven* .
- Thaler, Richard H and Cass R Sunstein. 2003. "Libertarian paternalism." *The American Economic Review* 93 (2):175–179.
- Wooldridge, Jeffrey M. 2015. *Introductory econometrics: A modern approach*. Cengage learning.
- Yörük, Barış K. 2014. "Does giving to charity lead to better health? Evidence from tax subsidies for charitable giving." *Journal of Economic Psychology* 45:71–83.
- Zanna, Mark P. 1972. "Inference of belief from rejection of an alternative action." *Representative Research in Social Psychology* .

# Appendix A Appendix A: Tables and Figures

This section is meant for online publication only.

FIGURE A1: DONATION EXPERIMENT SCREENSHOTS

Thank you again for participating.



In addition to the participation payment of \$1 that you will receive, you have also earned another \$1.

The extra \$1 will be added to your final payment, so you will receive a total of \$2 which includes \$1 for participating in this study plus this extra \$1 bonus.

Your Extra Bonus Earnings: \$1

(A) ROUND 1, CASH ENDOWMENT

Thank you again for participating.



In addition to the participation payment of \$1 that you will receive, we will also make a donation on your behalf to a charity called CARE.

Your Donation Amount: \$1.50

This \$1.50 will be added to all of the donations of every participant to make a single payment to CARE. If you wish to receive confirmation of this donation, you will have the opportunity to indicate this in the survey at the end of this study.

(B) ROUND 1, CHARITY ENDOWMENT



You now have the option to give away your extra \$1 bonus. If you decide to give away your bonus, we will instead make a \$1.50 donation to CARE on your behalf. In that case, you will have a \$1 bonus from the participation fee and a \$1.50 donation to CARE. We will add your \$1.50 donation to all of the donations of other participants to make a single payment to CARE. If you wish to receive confirmation of this donation, you will have the opportunity to indicate this at the end of this study.

- I do not want you to take \$1 away from my bonus.
- I want you to take away my entire \$1 bonus in order to donate \$1.50 to CARE.

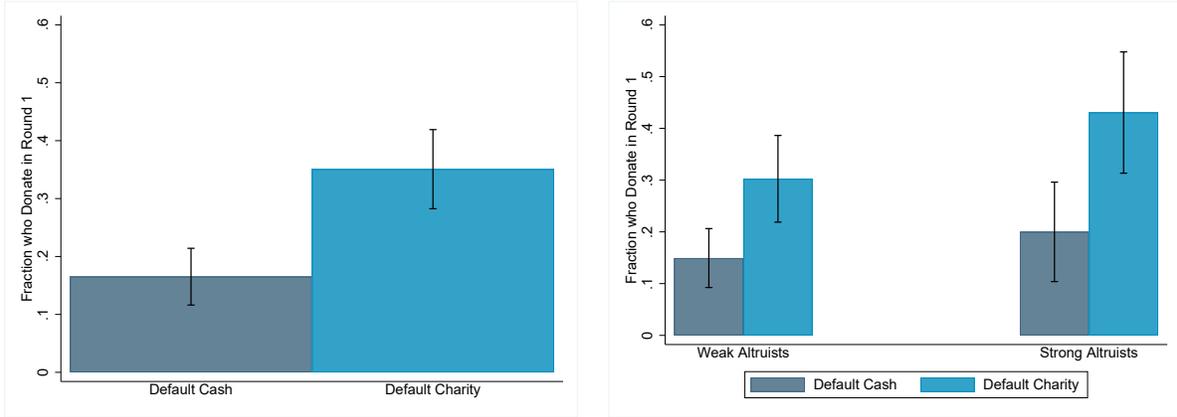
(C) SWAP CASH FOR DONATION

You now have the option to take away your \$1.50 donation to CARE. If you decide you do not want us to make a \$1.50 donation to CARE on your behalf, we will instead add \$1 to your bonus that you will get at the end of the survey. In that case, you will have a total of \$2 for your bonus consisting of \$1 of bonus for participation and \$1 for this decision, and we will not make any donation to CARE on your behalf.

- I do not want you to take away my \$1.50 donation to CARE.
- I want you to take away my entire \$1.50 donation to CARE in order to add \$1 to my bonus payment.

(D) SWAP DONATION FOR CASH

FIGURE A2: FIRST-STAGE: ROUND 1 DONATION RATES



(A) ROUND 1 RATES

(B) ROUND 1 RATES BY CONVICTIONS

Average donation rates in Round 1 by treatment assignment with 95% confidence intervals. Figure A2a shows that subjects assigned to the Default Charity treatment are 19 percentage points more likely (more than a 100 percent increase) to donate to charity in Round 1 than subjects assigned to the Default Cash treatment (t-test:  $p\text{-value} < .0001$ ). In Figure A2b we look at the effect of the default option nudge by subjects' number of charitable donations in the past 12 months. In general, we find that the Default Charity condition significantly increases donation rates regardless of the subjects' past donation history. For example, subjects with strongly-held (weakly-held) views towards altruism are 23 (15) percentage points more likely to give under the Default Charity than Default Cash condition (t-test:  $p\text{-value} = .003$  and  $p\text{-value} = .002$ , respectively).

TABLE A1: ROUND 2: MULTIPLE PRICE LIST FOR DONATION EXPERIMENT

Option 1:	Add \$1.00 to your bonus and Donate \$0 to Save the Children.
Option 2:	Add \$.90 to your bonus and Donate \$.20 to Save the Children.
Option 3:	Add \$.80 to your bonus and Donate \$.40 to Save the Children.
Option 4:	Add \$.70 to your bonus and Donate \$.60 to Save the Children.
Option 5:	Add \$.60 to your bonus and Donate \$.80 to Save the Children.
Option 6:	Add \$.50 to your bonus and Donate \$1.00 to Save the Children.
Option 7:	Add \$.40 to your bonus and Donate \$1.20 to Save the Children.
Option 8:	Add \$.30 to your bonus and Donate \$1.40 to Save the Children.
Option 9:	Add \$.20 to your bonus and Donate \$1.60 to Save the Children.
Option 10:	Add \$.10 to your bonus and Donate \$1.80 to Save the Children.
Option 11:	Add \$0 to your bonus and Donate \$2.00 to Save the Children.

TABLE A2: SUMMARY STATISTICS, DELAY TREATMENT

	All Delay Treatments	Default Charity Delay	Default Cash Delay
VARIABLES COLLECTED IN PART 1			
Never Heard of CARE	.76 (.43)	.77 (.42)	.75 (.44)
Risk Taking	1.18e-09 (1.00)	.07 (1.00)	-.07 (1.00)
Duration of Part 1 (sec)	614.44 (567.27)	621.08 (399.97)	608.36 (686.34)
Observations	673	322	351
VARIABLES COLLECTED IN PART 2			
Female	.49 (.50)	.48 (.50)	.51 (.50)
Age	42.17 (12.80)	42.50 (13.15)	41.87 (12.50)
Unemployed	.04 (.20)	.04 (.20)	.04 (.20)
Employed full-time	.70 (.46)	.70 (.46)	.70 (.46)
Employed part-time	.14 (.34)	.14 (.34)	.14 (.35)
Retired	.05 (.22)	.05 (.22)	.05 (.21)
Income < \$10,000	.04 (.20)	.04 (.19)	.05 (.21)
Income > \$150,000	.07 (.25)	.08 (.26)	.06 (.24)
Observations	596	278	318

Means reported with standard deviations in parentheses. The variables collected in Part 1 are used to correct for attrition between Part 1 and Part 2.

## Appendix B Additional Moral Consistency Results

### Appendix B.1 Identity & Moral Consistency

We augment the model presented in Section 2.4 to include a term that captures past private consumption and charitable giving such that

$$U(c, A) = u(c_t, c_{t-1}(\Theta_c)) + \alpha v(A_t, A_{t-1}(\Theta_A)) \quad (7)$$

where  $\Theta_c$  and  $\Theta_A$  represent a composite of private consumption and charitable giving up to and including time  $t - 2$ , respectively. Thus, today's utility depends on the choices the individual makes today as well as all past choices. The parameter  $\alpha \in [0, 1]$  governs the intensity of the individual's preference for altruism and warm glow. The functions  $u(\cdot)$  and  $v(\cdot)$  are concave in consumption and donations to charity, respectively. At time  $t$ , a subject chooses  $(c_t, A_t)$  taking their previous choices,  $(c_{t-1}, A_{t-1})$ , as given.

$$\max_{c_t, A_t} U(c_t, A_t \mid \bar{c}, \bar{A}) = \max_{c_t, A_t} u(c_t - \gamma_c c_{t-1}(\Theta_c)) + \alpha v(A_t - \gamma_A A_{t-1}(\Theta_A)) \text{ subject to } I = c_t + p \times A_t \quad (8)$$

From the first order conditions we find that

$$A_t(Z = 0) - \gamma_A A_{t-1}(\Theta_A, Z = 0) = A_t(Z = 1) - \gamma_A A_{t-1}(\Theta_A, Z = 1) \quad (9)$$

We assume that  $\frac{\partial A_{t-1}}{\partial \Theta_A} \Big|_{Z=1} \geq \frac{\partial A_{t-1}}{\partial \Theta_A} \Big|_{Z=0}$ . Rearranging and taking expectations of equation 9, we obtain

$$\frac{E[A_t \mid Z = 1] - E[A_t \mid Z = 0]}{E[A_{t-1} \mid \Theta_A, Z = 1] - E[A_{t-1} \mid \Theta_A, Z = 0]} = \gamma_A \quad (10)$$

To test these hypotheses about identity from section 3.2.1, we will use the same specification as equation 6, but with an interaction between endogenous regressor ( $A_{t-1}$ ) and the strength of conviction towards altruism, either weak ( $\Theta_A = 0$ ) or strong ( $\Theta_A = 1$ ), and instrument for Round 1 donation behaviour using the assignment to the Default Charity treatment interacted with the strength of the conviction. Our specification for this hypothesis is therefore given by

$$A_{i,2} = \delta_0 + \delta_1^{IV} \widehat{A}_{i,1} \times \mathbf{1}[\Theta_A = 0] + \delta_2^{IV} \widehat{A}_{i,1} \times \mathbf{1}[\Theta_A = 1] + \delta_3 \mathbf{X} + \varepsilon_i, \quad (11)$$

**Hypothesis 5. *Altruism as a Weak Facet of Identity I:***

$\delta_1^{IV} > 0$ : *individuals who hold altruism as a weak facet of their identity will behave morally consistently.*

**Hypothesis 6. *Altruism as a Weak Facet of Identity II:*** *if  $\gamma_A > 0$ , then individuals who hold altruism as a weak facet of their identity will behave more morally consistent than individuals who hold altruism as a strong facet of their identity; that is,  $\delta_1^{IV} > \delta_2^{IV}$ .*

Bénabou and Tirole (2011) also predict that when strongly-held convictions are challenged, individuals will be more likely to respond in a contradictory way to the challenge to restore their self-image. This means that for those individuals who have a strong identity towards altruism but are nudged towards selfishness (i.e., the Default Cash condition), Bénabou and Tirole (2011) predicts that individuals will respond by being more altruistic in the future. We formally state this in Hypothesis 7.

**Hypothesis 7. *Altruism as a Strong Facet of Identity:***

(i)  $\lambda_2^{IV} < 0$ : individuals who hold altruism as a strongly-held facet of their identity will respond in a contradictory way to a nudge towards selfishness.

TABLE B3: LOCAL AVERAGE TREATMENT EFFECTS: ROUND 2 DONATION RATES & AMOUNTS

<i>Panel A: Moral Consistency</i>				
	Propensity to Donate		Donation Amount	
	(1)	(2)	(3)	(4)
$\widehat{A}_{i,1} \times StrongValue$	-0.18 (0.39)	-0.13 (0.39)	0.04 (0.53)	0.12 (0.53)
$\widehat{A}_{i,1} \times WeakValue$	1.17** (0.63)	1.26** (0.67)	1.50* (0.96)	1.58* (0.99)
Strong Value	0.42*** (0.15)	0.42*** (0.15)	0.43* (0.22)	0.43** (0.22)
Constant	0.07 (0.08)	0.09 (0.1)	0.1 (0.12)	0.08 (0.16)
Observations	245	245	245	245
$R^2$	.	.	0.002	0.03
$\chi^2$ test				
$\widehat{A}_{i,1} \times Strong = \widehat{A}_{i,1} \times Weak$	3.30*	3.21*	1.75	1.69
<i>Panel B: Immoral Consistency</i>				
	Propensity to Keep		Keep Amount	
	(1)	(2)	(3)	(4)
$\widehat{c}_{i,1} \times StrongValue$	0.17 (0.24)	0.19 (0.25)	0.02 (0.27)	0.06 (0.26)
$\widehat{c}_{i,1} \times WeakValue$	0.46 (0.46)	0.46 (0.45)	0.75 (0.48)	0.79 (0.49)
Strong Value	0.23 (0.45)	0.2 (0.44)	0.52 (0.47)	0.52 (0.48)
Constant	0.51 (0.42)	0.52 (0.41)	0.2 (0.43)	0.17 (0.44)
Observations	245	245	245	245
$R^2$	0.03	0.05	0.002	0.03
$\chi^2$ test				
$\widehat{c}_{i,1} \times Strong = \widehat{c}_{i,1} \times Weak$	.32	.27	1.75	1.69

This table replicates Table 5 and Table B4 but we have redefined a Weak Identity towards altruism as those subjects who report that they gave 0 donations in the past year. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Given that we have a directional hypothesis for the effect of identity on moral consistency, the estimates on the coefficient for  $\widehat{A}_{i,1} \times Weak$  and  $\widehat{A}_{i,1} \times Strong$  are one-tailed.

## Appendix B.2 Immoral Consistency

To estimate the causal effect of keeping the money in Round 1 ( $c_{i,1}$ ) on the likelihood of also keeping the money in Round 2 ( $c_{i,2}$ ) we instrument for keeping the money in Round 1 using

the assignment to the Default Cash condition. We use a similar interaction as in equation 11.

$$c_{i,2} = \lambda_0 + \lambda_1^{IV} \widehat{c}_{i,1} \times \mathbf{1}[\Theta_A = 0] + \lambda_2^{IV} \widehat{c}_{i,1} \times \mathbf{1}[\Theta_A = 1] + \lambda_3 \mathbf{X} + \varepsilon_i, \quad (12)$$

In Table ??, we estimate equation 12 to examine whether there is evidence of immoral consistency; that is, does keeping the cash in Round 1 cause an increase in keeping the cash in Round 2. We do not find evidence consistent with immoral consistency on the extensive margin (columns (1) & (2)), but columns (3)& (4) show that keeping the cash in Round 1 causes subjects to keep more cash in Round 2.

**Result 6.** *We find no evidence that subjects who have a strong identity towards altruism behave in a morally balanced manner when nudged towards selfishness.*

Columns (2) & (4) test for the second part of the Bénabou and Tirole (2011) hypothesis, which states that subjects who are nudged away from a strongly-held value will respond in a contradictory manner. Thus, we hypothesized that subjects for whom altruism is a strongly-held value, but are nudged towards selfishness, would less selfish (or more altruistic) in Round 2. However, we do not find support for this hypothesis.

TABLE B4: LOCAL AVERAGE TREATMENT EFFECTS FOR IMMORAL CONSISTENCY: ROUND 2 DONATION RATES & AMOUNTS

	Propensity to Donate				Donation Amount			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{c}_{i,1}$	0.15 (0.16)	0.15 (0.16)	.	.	0.3* (0.16)	0.31* (0.16)	.	.
$\widehat{c}_{i,1} \times StrongIdentity$	.	.	0.17 (0.24)	0.2 (0.25)	.	.	0.02 (0.27)	0.05 (0.26)
$\widehat{c}_{i,1} \times WeakIdentity$	.	.	0.11 (0.23)	0.09 (0.22)	.	.	0.48** (0.24)	0.47** (0.24)
Strong Identity	.	.	-0.09 (0.25)	-0.13 (0.24)	.	.	0.27 (0.27)	0.24 (0.26)
Constant	0.78*** (0.12)	0.72*** (0.14)	0.82*** (0.18)	0.78*** (0.19)	0.57*** (0.12)	0.53*** (0.14)	0.45** (0.19)	0.42** (0.2)
Observations	415	415	415	415	415	415	415	415
$R^2$	0.08	0.1	0.09	0.1	0.21	0.22	0.12	0.14
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
$\chi^2$ test								
$\widehat{c}_{i,1} \times Strong = \widehat{c}_{i,1} \times Weak$ (p-value)			.03 (.86)	.11 (.74)			1.58 (.21)	1.44 (.23)

Demographic controls include gender, employment status (Full-time, part-time, retired or unemployed), and income. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively.

## Appendix C Additional No Choice Treatment Results

TABLE C5: SUMMARY STATISTICS, NO CHOICE TREATMENTS

	Treatment Conditions		
	Default Charity	Default Cash	Round 2
Altruism Strongly Held Value	.26	.34	.38
Past donations $\geq 4$	(.44)	(.47)	(.49)
Female	.51	.52	.54
	(.50)	(.50)	(.50)
Age	37.46	37.22	38.36
	(14.31)	(11.63)	(11.24)
Unemployed	.09	.06	.08
	(.28)	(.24)	(.26)
Employed full-time	.64	.64	.59
	(.48)	(.48)	(.49)
Employed part-time	.14	.17	.17
	(.35)	(.35)	(.37)
Retired	.03	.03	.02
	(.18)	(.16)	(.14)
Income < \$10,000	.05	.04	.05
	(.21)	(.19)	(.22)
Income > \$150,000	.04	.04	.05
	(.19)	(.19)	(.21)
Observations	558	630	199

Means reported with standard deviations in parentheses.

TABLE C6: DIFFERENCES-IN-DIFFERENCES: THE IMPORTANCE OF CHOICE IN DRIVING POSITIVE SPILL OVERS

	Probability to Donate			
	Default Charity Only		Default Cash Only	
	(1)	(2)	(3)	(4)
Round 2	0.05***	0.05***	0.16***	0.16***
	(1.45e-15)	(1.41e-15)	(8.50e-16)	(5.00e-16)
No Choice	0.65***	0.65***	-0.17***	-0.16***
	(1.69e-15)	(0.0008)	(6.98e-16)	(0.002)
Round 2 $\times$ No Choice	-0.55***	-0.55***	0.33***	0.33***
	(1.98e-15)	(1.88e-15)	(9.45e-16)	(7.23e-16)
Constant	0.35***	0.31***	0.17***	0.14***
	(1.30e-15)	(0.03)	(5.80e-16)	(0.01)
Observations	1498	1498	1708	1708
$R^2$	.	.	.	.
Demographic Controls	No	Yes	No	Yes

Columns (1)-(4) show differences-in-differences regression estimates. Columns (1) and (3) estimate a random effects estimator, while columns (2) and (4) include individual specific controls. Robust standard errors in parentheses and \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively.

We first consider the role of choice in driving positive inter-temporal effects on donation

behaviour. To do this, we consider the difference-in-difference that we referenced in Section 3.3.2 that compares the change in donation rates between Round 1 and Round 2 for the Choice and No Choice treatments. Consistent with the “back of the envelope” calculation performed in Section 2.3, we find that *restricting* choice in Round 1 in the Default Charity condition results in a 55 percentage point reduction in Round 2 donation rates. We show this by constructing a panel of giving, where each subject appears twice in the data set—Round 1 and Round 2. Using a random effects estimator, we separately analyze how the interaction of time and choice affects donation rates for the Default Charity conditions and then also the Default Cash conditions.<sup>19</sup> Table C6 presents our results.

In columns (1) and (2), we estimate the change in donation rates between Round 1 and Round 2 for the Default Charity conditions. We find that, on average, subjects are more likely to give in Round 2 and that subjects give more in the No Choice treatment (i.e., because they have No Choice). However, the significant negative coefficient on the interaction term in columns (1) and (2) suggest that subjects in the Default Charity No Choice treatments reduce their donation rates significantly more in Round 2 more than the Default Charity Choice treatment. Columns (3) and (4) display the same regression models for the Default Cash treatments and find that subjects in the Default Cash No Choice treatment *increase* their donation rates significantly more than subjects in the Default Cash Choice treatment.

---

<sup>19</sup>We separately estimate the effect of experimentally *removing* choice for the Default Charity and the Default Cash treatments to avoid a three-way interaction and simplify the interpretation.