KIEL INSTITUTE FOR THE WORLD ECONOMY

KIEL WORKING PAPER

Social Comparison Nudges – Guessing the Norm Increases Charitable Giving

No. 2058 November 2016

Simon Bartke, Andreas Friedl, Felix Gelhaar, and Laura Reh

Kiel Institute for the World Economy
ISSN 2195–7525

SOCIAL COMPARISON NUDGES — GUESSING THE NORM INCREASES CHARITABLE GIVING

Simon Bartke[a], Andreas Friedl,[a,b], Felix Gelhaar[a,b], and Laura Reh[c]

ABSTRACT

Social comparison nudges that employ descriptive norms were found to increase charitable giving. This paper finds that individuals who receive a descriptive norm donate significantly more when they have to guess the descriptive norm beforehand. We argue that guessing draws attention to the norm and therefore increases its effectiveness. Our results suggest that the effectiveness of nudges that use descriptive norms depends on how the a priori beliefs about the descriptive norm are updated.

Keywords: Social comparison nudge, attention, field experiment, charitable giving, social norms

JEL Classification codes: C93, D03, D64, H4

Corresponding author: Andreas Friedl, Kiel Institute for the World Economy, Kiellinie 66
24105 Kiel, Germany; Telephone: +46(0)431-8814285, E-mail: andreas.friedl@ifw-kiel.de

[a] Kiel Institute for the World Economy, Kiel, Germany
[b] Christian-Albrechts-University, Department of Economics, Kiel, Germany
[c] University of Cologne, Germany

The responsibility for the contents of this publication rests with the author, not the Institute. Since working papers are of a preliminary nature, it may be useful to contact the author of a working paper about results or caveats before referring to, or quoting, a paper. Any comments should be sent directly to the author.

1 Introduction

Modern societies are based to a great deal on implicit rules and norms. Information on the behavior of others, i.e. descriptive norms, can influence our own behavior (Festinger, 1954). Organizations that rely on donations increasingly take advantage of this by giving descriptive norm cues in their solicitation. They use so-called social comparison nudges (SCNs) that provide individuals with information about the behavior of relevant peers. Several studies find that descriptive norms positively impact donations (see e.g. Agerström et al., 2016; Alpízar and Martinsson, 2013; Frey and Meier, 2004; Martin and Randal, 2008; Shang and Croson, 2009).

The literature in social psychology finds that an increase in attention to a descriptive norm strengthens its influence on behavior (Chaiken and Eagly, 1989; Cialdini et al., 1990; Fazio, 1990; Harvey and Enzle, 1981; Melnyk et al., 2011; Petty and Wegener, 1999;). Krupka and Weber (2009) find that even without being provided a descriptive norm, giving a guess about peer behavior increases pro-sociality in a binary dictator game by increasing the attention to the norm. Following Krupka and Weber we argue that asking for a guess of the norm would also increase the attention to the provided descriptive norm.

The novelty of this paper is therefore to provide subjects with the descriptive norm directly after they guessed it. This way, subjects' beliefs about the norm, as expressed by their guess, are directly updated. This belief update is hypothesized to further increase the effectiveness of the provided descriptive norm. We test this hypothesis in a field experiment by comparing donation rates to a local charity. As descriptive norm we use the donation rate of the general public from a study of the German Ministry of Family Affairs. We find that asking for a guess of the descriptive norm before providing it significantly increases donations over merely providing it. We further find suggestive evidence that it matters for the donation decision whether subjects receive a belief update which lies below or above their previously stated guess about the descriptive norm.

2 Methods

The field experiment was conducted at the main train station in the city of Kiel, Germany in September 2015. Our sample consists of 263 observations (131 females) in three treatments. The participants were recruited from the general public waiting for public transport, mainly on their way to or from work. We conducted our experiment on work days, both in the morning between 7 and 9 a.m. and in the evening between 4 and 6 p.m. One solicitor that was unaware of the study's hypotheses carried out all treatments of the experiment. Following a written protocol, the solicitor approached subjects by asking whether they would be willing to participate in a two minute survey. Only individuals that made the impression of commuting alone were approached and the solicitor continued with individuals out of earshot in order to minimize potential social image effects. Participants were reimbursed with a scratch lottery ticket. The tickets are well known in Germany, cost €1, and we explicitly mentioned the jackpot prize of €60,000. This facilitated the approach of subjects and is likely to have increased participation. It further provides an easy, non-strategic decision environment with a binary choice to either donate the (unscratched) ticket or not.

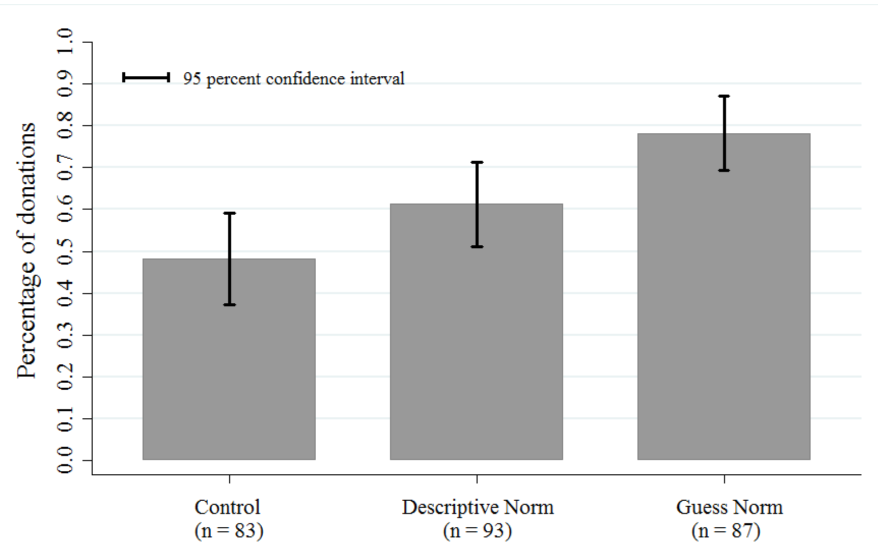
Participants faced one of three treatments; *Control*, *Descriptive Norm* or *Guess Norm*. In the *Control* treatment participants were asked whether they wanted to donate their ticket to a local child relief organization. In the *Descriptive Norm* treatment participants were told the following prior to the donation decision: "According to a survey of the Ministry of Family Affairs about 2/3 of the population in Germany make charitable donations per year." In the *Guess Norm* treatment participants first had to give a guess on the question: "What percentage of the population in Germany do you think makes charitable donations per year?" Only after they stated their guess the solicitor provided the same statement as in *Descriptive Norm*. All donated tickets were put into a non-transparent donation box to avoid social cues. After the donation decision a short follow-up questionnaire was conducted (translated versions of the experimental protocol and questionnaire can be obtained from the authors upon request).

3 Results

We are interested in how the decision to either donate or keep the scratch ticket is affected by the different treatments. We use Fisher's exact test to test for the difference in proportions of donations between treatments. We report p-values from two sided testing throughout, except mentioned otherwise. Our main results are summarized in Figure 1.

In line with previous results in the literature investigating SCNs, we find that subjects who are merely informed about the donation behavior of others (*Descriptive Norm*: mean = 61% sd = 5%) donate the scratch ticket more often than subjects in the control group (*Control*: mean = 48%, sd = 5%) who received no descriptive norm. This difference in donation behavior of 13% is marginally statistically significant (p = 0.096).

Figure 1: Percentage of lottery ticket donations for three treatments



The most pronounced effect on the share of donations is observed in *Guess Norm*. We find clear evidence that making subjects state a guess about the descriptive norm before receiving it leads to a higher share of donations compared to both *Control* and *Descriptive Norm*. In particular, while *Descriptive Norm* increases donations only marginally compared to providing no descriptive norm, subjects under *Guess Norm* donate in 78% of the cases. This share is significantly higher (*Guess Norm*: mean = 78%, sd = 4%) than under *Descriptive Norm* ($p = 0.016$). The difference in donations between *Guess Norm* and *Control* therefore is 30% (vs. 13% between *Descriptive Norm* and *Control*) and statistical significant ($p < 0.001$).

Table 1 depicts the marginal effects at the mean that incremental changes in certain independent variables have on the dependent variable "donation" as estimated by logistic regression with robust standard errors. Model (1) examines the effect of the *Descriptive Norm* treatment relative to the *Control* treatment, leaving out the *Guess Norm* observations. This model controls for time of day (morning or evening), gender, and age. These covariates were selected in order to control for the time slots when we collected our observations, the body of evidence which suggests gender differences in charitable giving (Andreoni and Versterlund, 2001; Mesch et al., 2011) and the fact that age is the only predictor apart from our treatment variables that has a stable significant influence on donation decisions, respectively. Likewise, Model (2) estimates the impact of the *Guess Norm* treatment relative to the *Control* treatment in the presence of these covariates, but leaving out the *Descriptive Norm* observations. The estimated marginal effects of the treatments are robust to other model specifications including different regressors (weekday, weather condition, type of occupation, and solicitor experience, i.e. number of session) as can be seen in the appendix. We see that the marginal effects of the *Descriptive Norm* and *Guess Norm* variables in these two models are consistent with the treatment effects described above.

Previous work that investigated how descriptive norms affect behavior suggests that heterogeneous treatment effects can occur when social norm information is used to influence behavior. It was found that subjects who learned that their behavior deviates from the behavior of the peer group are likely to adjust their behavior towards the norm. When this adjustment occurs in the direction that is not intended by the nudge, this is coined "boomerang effect" (Clee and Wicklund, 1980). This effect was found in different applications in the field (Schultz et al., 2007; Costa and Kahn, 2011). From this, we hypothesize that subjects with a belief update from guessing below the descriptive norm are more likely to donate than subjects with a belief update from guessing above.

Table 1: Marginal effects

VARIABLES	(1)	(2)
Guess Norm	-	0.316***
	(-)	(0.0714)
Descriptive Norm	0.139*	-
	(0.0753)	(-)
Evening	0.0289	-0.0590
	(0.0765)	(0.0776)
Female	-0.0585	0.0550
	(0.0765)	(0.0778)
Age	-0.0042	-0.0057*
	(0.0029)	(0.0030)
Observations	176	170

Robust standard errors in parentheses (OIM)
*** p<0.01, ** p<0.05, * p<0.1
NOTE: Treatments, evening, female are categorical variables; age is at its mean value

When we look at the stated guesses, we find that 86% of them are below our descriptive norm. Only 12 participants (14%) in *Guess Norm* guessed higher than the descriptive norm (mean = 74.25, sd = 5.48), while 75 guessed lower (mean = 29.33, sd = 15.09). When we compare donation decisions of participants that guessed below the norm (mean = 81%, sd = 4%) with donation decisions of participants that guessed above the norm (mean = 58%, sd = 14%) we find participants that guessed above are marginal significantly less likely to donate (p = 0.084, one-sided). Our finding suggests that descriptive norms not only interact with previous behavior in shaping decision making, but also with previous beliefs about the norm.

4 Conclusion

This paper reports the results of a field experiment testing the influence of a descriptive norm on donation behavior. We contribute to the literature by studying how increasing attention to the descriptive norm affects donations. We find that asking for a guess of the descriptive norm before providing it significantly increases donations over merely providing it.

We additionally find tentative evidence that subjects with guesses above the norm donate less than those with guesses below the norm. Explanations for this include the boomerang effect and moral licensing (see Merrit et al., 2010 in general and Tiefenbeck et al., 2013 for moral licensing in SCNs). Our results extend both explanations by suggesting that the effectiveness of SCNs depends on how the descriptive norm updates the a priori beliefs that individuals hold about the descriptive norm.

To ask for a guess is an easy and low-cost practice that is applicable not only to charitable giving but other policy areas in which SCNs have been applied. These include promoting energy (Allcott, 2011; Nolan, et al., 2008) and water conservation (Ferraro and Price, 2013), retirement savings (Beshears et al., 2015), and increasing voting participation (Green and Gerber, 2008; Margetts et al., 2011). For a better understanding on why and how norms shape behavior more research is needed on the interaction between beliefs about and the degree of attention to norms.

References

Agerström, J., Carlsson, R., Nicklasson, L. & Guntell, L., 2016. Using Descriptive Social Norms to Increase Charitable Giving: The Power of Local Norms. *Journal of Economic Psychology*, 52, pp.147–53. doi:10.1016/j.joep.2015.12.007.

Andreoni, J. & Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, 116, pp.293-312.

Allcott, H., 2011. Social Norms and Energy Conservation. *Journal of Public Economics*, 95, pp.1082–1095.

Alpízar, F. & Martinsson, P., 2013. Does It Matter if You Are Observed by Others? Evidence from Donations in the Field. *Scandinavian Journal of Economics*, 115, pp.74–83.

Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C. & Milkman, K.L., 2015. The effect of providing peer information on retirement savings decisions. *The Journal of finance*, 70, pp.1161-1201. doi:10.1111/jofi.12258

Chaiken, S. & Eagly, A.H., 1989. Heuristic and systematic information processing within and beyond the Persuasion Context., in: Uleman, J.S. & Bargh, J.A. (Eds.), *Unintended thought*, New York: Guilford, pp.212–252.

Cialdini, R.B., Reno, R.R. & Kallgren, C.A., 1990. A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. *Journal of personality and social psychology*, 58, pp.1015–1026.

Fazio, R.H., 1990. Multiple process by which attitudes guide behavior: The MODE mofel as an integrative framework. *Advances in Experimental Social Psychology*, 23, pp.75–109. http://dx.doi.org/10.1016/S0065-2601(08)60318-4.

Ferraro, P.J. & Price, M.K., 2013. Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *Review of Economics and Statistics*, 95, pp.64–73. http://dx.doi.org/10.1162/REST_a_00344\nhttp://www.mitpressjournals.org/doi/abs/10.1162/REST_a_00344.

Festinger, L., 1954. A theory of social comparison processes. *Human Relations*, 7, pp.117–140.

Frey, B.S. & Meier, S., 2004. Social Comparisons and Pro-social Behavior: Testing in a Field Experiment “Conditional Cooperation”. *The American Economic Review*, 94, pp.1717–1722.

papers3://publication/uuid/59EA7CEC-2288-44F3-BF35-21032C468844.

Green, D. P., & Gerber, A. S., 2008. *Get out the vote: How to increase voter turnout*, Washington DC: Brookings Institution Press.

Harvey, M.D. & Enzle, M.E., 1981. A cognitive model of social norms for understanding the transgression-helping effect. *Journal of Personality and Social Psychology*, 41 p.866-75.

Krupka, E. & Weber, R.A., 2009. The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology*, 30, pp.307–320. http://dx.doi.org/10.1016/j.joep.2008.11.005.

Margetts, H., John, P., Escher, T., & Reissfelder, S. (2011). Social information and political participation on the internet: an experiment. *European Political Science Review*, 3, pp.321-344.

Martin, R. & Randal, J., 2008. How is donation behaviour affected by the donations of others? *Journal of Economic Behavior and Organization*, 67, pp.228–238.

Melnyk, B.M. & Fineout-Overholt, E., 2011. *Evidence-based practice in nursing & healthcare: A guide to best practice*, Philadelphia: Lippincott Williams & Wilkins.

Merritt, A.C., Effron, D.A. & Monin, B., 2010. Moral self-licensing: When being good frees us to be bad. *Social and personality psychology compass*, 4, pp.344–357.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V., 2008. Normative social influence is underdetected. *Personality & Social Psychology Bulletin*, 34, pp.913–923. http://doi.org/10.1177/0146167208316691

Petty, R.E. & Wegener, D.T., 1999. The elaboration likelihood model: Current status and controversies., in: Chaiken, S. & Trope, Y. (Eds), *Dual-process theories in social psychology*, New York: Guilford

Shang, J. & Croson, R., 2009. A Field Eexperiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. *The Economic Journal*, 119, pp.1422–1439. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.2009.02267.x/full.

Socio-Economic Panel (SOEP), 2015. SOEP 2014: Documentation of Person-related Status and Generated Variables in PGEN for SOEP v31. *SOEP Survey Papers*, 292, Series D. Berlin: DIW Berlin.

Tiefenbeck, V., Staake, T., Roth, K., & Sachs, O., 2013. For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, 57, pp.160-171.

Appendix

The following analysis presents several robustness checks to the regression analysis in the main text. We examine the treatment effects of the *Descriptive Norm* and *Guess Norm* treatments by means of logistic regression models that investigate the predictive power on donation decisions of the two treatment dummies and several other potentially relevant covariates. More precisely, in the regression models in this appendix we additionally control for the type of occupation, the weather condition, weekday effects and solicitor experience (number of session).

The categories of the variable "type of occupation" were obtained by asking subjects about their occupational status. Based on this information we categorized each subject into an occupational position by using the "Current Employment Status" classification system in its hierarchically summarized version on page 11 from the Socio-Economic Panel (SOEP) (2015). This led to the following categories (with respective subject counts in parentheses): pensioner (6), currently in education (22), apprentice (10), self-employed (6), manual laborer (19), employee (in services, 148) and civil service (49). Three subjects were unwilling to report their occupation. Moreover, we cannot report any observations in our sample for the SOEP occupational position categories of "unemployed" and "military / social service". The weather condition was also categorized into one of four categories (as categorized by our solicitor for every session that she conducted): sunny (79 observations, partly cloudy – sunny (62 observations), cloudy (100 observations) and rainy (22 observations). The "rainy" category serves as the baseline category in the regression analysis below. In order to control for weekday effects, we categorized weekdays into three categories of roughly equal size: The "beginning of the week" category contains the days Monday and Tuesday and has 92 observations, the "middle of the week" category contains the days Wednesday and Thursday and has 99 observations and the "end of week" category contains Fridays and has 72 observations. These three categories do not differ in size significantly by means of a Chi-squared test $\chi^2\,(2) = 3.86, p = 0.145$. The "end of week dummy" serves as the baseline category in the models below. Solicitor experience could bias our results if increased solicitation experience over time leads a change in solicitation success. To control for time trends in our data, we add the session number as a regressor to the models. In total, 22 sessions were conducted in chronological order with session number 1 being the first session that was conducted.

Model (A1) describes a logistic regression analysis with robust standard errors clustered at the session level of the two treatment variables *Descriptive Norm* and *Guess Norm* and examines the effect of the regressors gender, age and time of day when the donation decision was made on the dependent variable "donation of scratch ticket to charity" which are shown in Models (1) and (2) presented in the main text. Moreover, Model (A1) contains the additional regressors time of week, weather condition, type of occupation and number of session that were described above. Model (A2) in Table A1 depicts the marginal effects logistic regression with robust standard errors (OIM) with the identical list of regressors as in Model (A1). This way, Models (A1) and (A2) demonstrate the (marginal) treatment effects of the *Guess Norm* and *Descriptive Norm* treatments as well as the effects of the described control variables on donation decisions for the full sample which has not been analyzed in the main text. Similarly, Models (A3) and (A4) contain the same full list of regressors, but leave out *Guess Norm* and *Descriptive Norm* observations, respectively. With this model structure, Models (A3) and (A4) present the logistic regressions with additional control regressors that serve as a robustness check to the marginal effect Models (1) and (2) shown in the main text with standard errors clustered at the session level. By means of a specification link test we find no evidence that adding complexity to Models (1) and (2) in the main text as well as Models (A1) – (A4) in the appendix through squaring the prediction leads to a significant increase in predictive power. Furthermore, we find no evidence for critical multicollinearity. Testing our regressors for correlation with one another, no regressor shows a tolerance between 0.2 or a variance inflation factor (VIF) above 5. The mean VIF over all regressors of Models (A1) – (A4) is 2.42.

Table A1: Models testing for robustness

VARIABLES	Model (A1) Full model	Model (A2) Full model (marginal effects)	Model (A3) Full model (*Descriptive Norm* only)	Model (A4) Full model (*Guess Norm* only)
Guess Norm	1.377***	0.315**	-	0.988**
	(0.410)	(0.125)	(-)	(0.493)
Descriptive Norm	0.624*	0.143*	0.659*	-
	(0.350)	(0.0734)	(0.370)	(-)
Evening	0.0884	0.0202	0.324***	-0.383
	(0.223)	(0.0667)	(0.113)	(0.339)
Female	0.0696	0.0159	-0.0765	0.307
	(0.215)	(0.0647)	(0.263)	(0.309)
Age	-0.0326***	-0.00745***	-0.0341***	-0.0334***
	(0.00816)	(0.00281)	(0.0105)	(0.0114)
Beginning of week	0.609**	0.139	1.199***	-0.143
	(0.310)	(0.108)	(0.245)	(0.432)
Middle of week	0.0822	0.0188	0.163	-0.754**
	(0.188)	(0.0865)	(0.167)	(0.335)
Sunny	0.862**	0.197	1.455***	0.522
	(0.337)	(0.135)	(0.477)	(0.520)
Partly cloudy	0.736	0.168	1.990***	-0.250
	(0.512)	(0.157)	(0.430)	(0.626)
Cloudy	0.776	0.177	1.811***	0.616
	(0.479)	(0.128)	(0.522)	(0.590)
Pensioner	2.037*	0.466	1.312	2.102
	(1.165)	(0.295)	(1.250)	(1.313)
In education	0.105	0.0241	-1.092	0.839
	(0.997)	(0.167)	(1.088)	(1.167)
Apprentice	-0.282	-0.0645	-1.029	-0.0285
	(1.139)	(0.201)	(1.238)	(1.295)
Self-employed	0.602	0.138	14.60***	0.488
	(0.994)	(0.235)	(0.801)	(0.942)
Employee (services)	0.169	0.0386	-0.496	0.600
	(0.789)	(0.116)	(0.759)	(0.915)
Public service job	0.265	0.0606	-0.357	0.773
	(0.815)	(0.131)	(0.842)	(0.962)
Session	0.0290	0.00663	0.00719	0.0595
	(0.0231)	(0.00816)	(0.0187)	(0.0390)
Constant	-0.361	-	-0.693	0.128
	(0.985)	(-)	(0.991)	(1.205)
Observations	263	263	176	170

Robust standard errors in parentheses (Models A1, A3, A4: clustered at session level; Model A2: OIM)
*** p<0.01, ** p<0.05, * p<0.1

NOTE: In Model (A2) continuous variables are evaluated at their mean values

The results from Model (A1) demonstrate that the treatment effects of both treatment dummies described in the main text are robust to the inclusion of several control variables and in the full sample. In this model, we continue to find that the *Guess Norm* treatment remains a strongly significant positive influence on donation decisions even in the presence of additional control variables. The coefficient estimate for *Guess Norm* is significant at the p = 0.001 level, while the coefficient estimate for the *Descriptive Norm* treatment is still significant, yet only at the marginal level at the p = 0.075 level, thus confirming the results presented in the main text. Model (A2) presents the marginal effects to the Model (A1). The estimated marginal effects associated with the *Guess Norm* and *Descriptive Norm* treatments mirror almost perfectly those estimated in the shorter marginal effects models (1) and (2) from the main text despite the fact that Model (A2) contains a longer list of regressors and observations over all three treatment conditions (*Guess Norm, Descriptive Norm, Control).* Whereas Model (A2) estimates a marginal effect of the *Guess Norm* (*Descriptive Norm)* treatment on increasing the likelihood of donation of 31.5% (14.3%), Model (1) ((2)) estimates this effect to be 31.6% (13.9%). Finally, Models (A3) and (A4) show that the treatment effects are also robust to the full list of covariates in a model specification that holds out the *Guess Norm* and *Descriptive Norm* observations respectively, which is identical to the specification presented in the main text. We take this as evidence that especially *Guess Norm* increases the likelihood of a donation significantly and across several model specifications. Likewise, the treatment effect of *Descriptive Norm* is robust across different model specifications, yet considerably smaller than the effect of *Guess Norm.*

In terms of significant influences of covariates other than the two treatment dummies on the donation decision, Models (A1) – (A4) lead to the following insights: Throughout all model specifications we continue to find a stable and significant negative effect of age on the decision to donate.[1] However, as can be seen by the arguments in the previous paragraph, this effect of age affects the size and significance of our treatment effects only marginally. Model (A3) indicates that *Descriptive Norm* works particularly well in increasing the likelihood of donation in the evening. This finding is not supported in the *Guess Norm* treatment, however.

[1] *Incorporating non-linear age regressors does not lead to an improved model fit.*

We speculate that this difference could be affected by the relationship between changes in attention over the time of day and how different ways of providing a descriptive norm interact with this fluctuations in attention. This speculation is driven by the insight presented in the main text, that *Guess Norm* has a stronger influence on focusing attention to the descriptive norm than *Descriptive Norm*. We leave an investigation of this relationship to further research. We likewise find heterogeneous effects for time of the week in our data between the two treatments. Whereas there seems to exist a positive effect of "beginning of the week" on donation decisions under *Descriptive Norm*, this effect is negative under *Guess Norm*. To the contrary, the effect of Friday's on donations is positive under *Guess Norm*, whereas it is negative under *Descriptive Norm.* Almost throughout all models (except for partly cloudy under *Guess Norm*) there is a positive influence of every non-rainy weather condition on donation behavior. In Model (A1), sunny weather has a significant positive influence at the $p = 0.011$ level on donation likelihood. In Model (A3), all non-rainy weather conditions have a significantly positive influence on donation decisions at the below 1% significance level. We can only speculate that a positive psychological concept like good mood is induced through better weather conditions which affect donations positively. However, this potential weather effect is not present in the *Guess Norm* treatment which could suggest that this treatment focuses attention on the descriptive norm in such a way that the weather condition surrounding the decision situation exerts only insignificant influence. Further research could try to confirm this insight and to identify the causal mechanism behind it if it persists. In terms of how affiliations with different types of occupation affect the likelihood to donate we find that over all models in the appendix, identifying as a pensioner and being self-employed unanimously increases the likelihood to donate, whereas the occupational position as an apprentice decreases donation likelihood. While the directions of the effects for these occupations are identical across models, none of them are statistically significant, except for the self-employed category in Model (A3), which is however only populated with one observation in this model and therefore almost not insightful. Importantly, we want to point out that across all four models in the appendix, the variable "session" does not show a significant influence on donation decision. Given that our solicitor was blind to the study's hypotheses, we therefore conclude that (a) our solicitor has not learned a systematic behavior over time which influenced subjects' decisions and (b) finding (a) is also robust at the within-treatment level which means that she also did not learn a systematic behavior influencing donation decisions over the sessions of one particular treatment. To conclude, Models (A1) – (A4) demonstrate that the marginal effects we present in the main text are (a) robust to a longer model (Models (A3) and (A4)) and robust to a longer model that also contains both treatments (Models (A1) and (A2)).