A portable method of eliciting respect for social norms*

Erik O. Kimbrough† **Alexander Vostroknutov**‡

October 30, 2017

Abstract

Recent models of prosociality suggest that cooperation in laboratory games may be better understood as resulting from concern for social norms than from prosocial preferences over outcomes. Underlying this interpretation is the idea that people exhibit heterogeneous respect for shared norms. We introduce a new, abstract task to elicit a proxy for individual norm-following propensity by asking subjects to choose from two actions, where one is costly. We instruct subjects that "the rule is" to take the costly action. Their willingness to incur such a cost reveals respect for norms. We show that choices in this task are similar across five countries. Rule-following is correlated with norm-consistent behavior in dictator games, providing support for our interpretation.

JEL classifications: C91, D03
Keywords: experimental economics, norms, prosocial behavior, social preferences

*Thanks to Mahsa Akbari and Nidhi Dahiya for excellent research assistance; to the Social Sciences and Humanities Research Council of Canada Insight Grants (435-2015-0798) and Insight Development Grants (430-2012-0146) programs for providing research funds; and to Folco Panizza for collecting and analyzing Italian data. We received helpful comments from Cary Deck, Vernon Smith, and Bart Wilson. Thanks also to the Visiting Scholar's Program in the Walton College of Business at the University of Arkansas. The experiments were programmed using zTree (Fischbacher, 2007). All remaining errors are our own.

†Corresponding Author: Smith Institute for Political Economy and Philosophy, Chapman University, One University Drive, Orange, CA 92866, USA, e-mail: ekimbrou@chapman.edu

‡Center for Mind/Brain Sciences, University of Trento, Via delle Regole 101, 38123 Mattarello (TN), Italy. e-mail: a.vostroknutov@unitn.it

1 Introduction

Extensive experimental and empirical evidence suggests that many individuals readily incur costs in order to cooperate, donate, reciprocate and punish others, often in opposition to the axiom of self-interested payoff maximization. Earlier attempts to account for such observations relied on the idea that subjects had 'social preferences' defined over payoff distributions, rather than own payoffs (see e.g. Fehr and Schmidt, 1999), but evidence suggests that such accounts are incomplete, as prosocial and other-regarding behavior (e.g. giving in dictator games, rejection in ultimatum games) is known to be sensitive to payoff-irrelevant manipulations such as who is watching and how decision rights are allocated (Hoffman *et al.*, 1994, 1996; Cherry *et al.*, 2002). Moreover, recent evidence suggests that when individuals can conceal their selfishness, they regularly take advantage of this "moral wiggle room" and behave more selfishly than they otherwise would (Dana *et al.*, 2007; Andreoni and Bernheim, 2009; Lazear *et al.*, 2012).

This sensitivity of prosocial behavior to context cannot be rationalized via simple models in which utilities are defined only over payoffs, but one theoretical approach that better fits the available evidence suggests that prosocial behavior is driven by a desire to adhere to social norms (Cappelen *et al.*, 2007; López-Pérez, 2008; Kessler and Leider, 2012; Krupka and Weber, 2013; Kimbrough and Vostroknutov, 2016). In such models, individuals evaluate the appropriateness of own behavior by comparing it to exogenously defined and commonly known injunctive social norms, that is, to a set of shared beliefs about the appropriateness of possible actions available in the setting.[1] Each possible action has both direct payoff consequences and a normative valence, and the simplest such models assume these enter additively in the utility function. Social norms are assumed to vary with context, and when norms and self-interest conflict, individuals face tradeoffs between following the norm and maximizing their payoffs. Crucially, individuals may differ in the degree to which they suffer from norm violations (or gain from norm adherence), and this can be captured in a single parameter that assigns a weight to the normative component of utility. In such models, heterogeneous behavior within a given game can be explained via heterogeneity in these weights; while, heterogeneous behavior across games, or across contexts, can be explained by the fact that norms vary with the setting.

A key implication of these models is that observed behavior in social settings (i.e., where norms of conduct apply) cannot offer direct evidence on preferences since behavior deviating from payoff maximization is a joint product of both norms and the weight assigned to them. One major concern with the norm-based approach is that we currently lack a solid theory of how and why norms vary with context, which introduces a potentially dubious source of "researcher

[1] Other- and self-signalling models have also been proposed as a possible unifying account of prosocial behavior in games. In these models agents wish to maintain their reputations and behave prosocially in order to do so (e.g. Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009). As argued in Kimbrough and Vostroknutov (2016) such approaches are complementary to a norms-based approach since the ability to know which actions will help or harm a reputation is predicated on shared normative beliefs (i.e. injunctive norms).

degrees of freedom" into the framework. To alleviate this it is important to be able to pin down the norm in a given setting *ex ante*. To this end, Krupka and Weber (2013) introduced a method of directly eliciting shared beliefs about social norms using incentivized coordination games in which subjects are paid if they correctly identify others' normative beliefs about choices in a given game. Using this method, they showed that context-driven changes in norms can account for observed changes in behavior in a set of dictator game experiments reported in List (2007). Adapting their methods appropriately, it is possible to empirically identify social norms in any game (Kimbrough *et al.*, 2014), and thus while we may not be able to say something about the sources of norms in general, we can pin the norm down in any particular application and thereby learn something about the sources of social behavior.[2] Moreover, this norm elicitation method has been shown to be robust to asking subjects to identify the norm from the perspective of both second- and third-party roles in games as well to order-effects in within-subject designs (Erkut *et al.*, 2015; D'Adda *et al.*, 2016), suggesting that there is a single, consistent injunctive norm (on average), perceived by participants and third parties.

A second issue is that these models assume a single, portable weight placed on the normative component of utility, which is also observationally confounded with norms in social settings.[3] One solution is to attempt to measure (a proxy for) this parameter directly using an experimental task. Kimbrough and Vostroknutov (2016) introduced such an individual decision task, called the rule-following task, designed to measure how much weight individuals place on social norms. In their task norm- (or rule-) following is pitted directly against material interest in a setting in which social preference models and payoff-maximizing models make the same predictions, namely that all individuals will break the rule. Using this task and separately eliciting norms using variants of the task due to Krupka and Weber (2013), the authors found that rule-following behavior was correlated with norm-consistent behavior in public goods, trust, ultimatum and dictator games; in a followup paper, they also found that this task was correlated with norm consistent behavior in dynamic common pool resource games (Kimbrough and Vostroknutov, 2015).

While this approach provides evidence in favor of the norms-based account of prosociality, the rule-following task employed contextual cues (traffic lights) which may have introduced noise into the measurement of subjects' concern for social norms. This may limit the generality of the methods developed so far. Moreover, the norm-dependent utility specification relies on an additional untested assumption: independence of the weight that people put on following social norms and the beliefs they have about the shared norm.

To address the issue of contextual cues in the rule-following task, we present a new, abstract task which can be widely used to measure the propensity to adhere to social norms. We

[2] See the examples in Kimbrough and Vostroknutov (2015, 2016).

[3] With repeated observations of the same subject, making strong assumptions about the form of the utility function, it is possible to estimate this parameter from data. However, to our knowledge existing work has only attempted to estimate the population mean concern for norms.

have conducted a number of experiments that included this task in five countries (Canada, Italy, Netherlands, Turkey, USA), and for a subset of the rule-following data collected in three of those countries (Canada, Netherlands, USA), we also subsequently collected data in a variant of the standard Dictator game. The distribution of behavior in the rule-following task across the samples is quite similar, and the correlation between rule-following and norm-consistent behavior in the Dictator game is large in magnitude and highly significant overall, though we find some differences between countries. In addition, using separate groups of subjects in the USA and Canada, we have elicited norms in the Dictator game after also measuring rule following propensity via our new task. This allows us to test the independence assumption. While the overall normative beliefs of rule-followers and rule-breakers are quite similar - both types tend to believe that the equal split is the most normatively appropriate and that keeping the whole pie is the least normatively appropriate action - we do find some evidence that beliefs and rule-following behavior are not completely independent. More work needs to be done in order to understand this observation, though one possible explanation is a mild self-serving bias among rule-breakers. Overall, despite some complications, we conclude that the new task offers a simple, direct and portable method of eliciting individual-level respect for social norms.

2 The Task

2.1 The Rule Following Task

In the rule following task participants drag-and-drop balls one-by-one into two buckets: yellow or blue. For each ball in the yellow bucket they receive 10 cents and for each ball in the blue bucket they receive 5 cents (see Figure 1). The current earnings from the two buckets are shown above them. The total earnings in this task is the sum of earnings from the buckets. The position of the buckets on the screen is randomized across individuals.

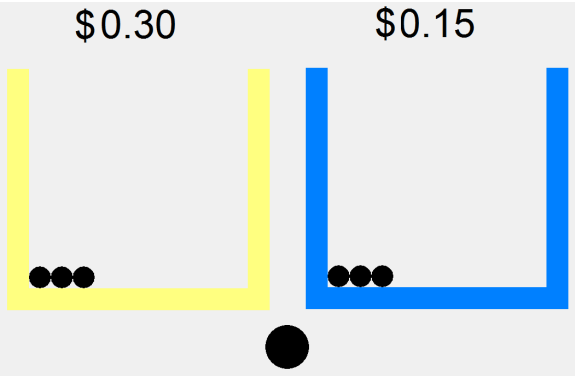


Figure 1: Presentation of the rule following task.

The instructions explicitly state that *"the rule is to put the balls into the blue bucket"* (see Appendix). Participants have a fixed number balls to allocate (say 100); thus, their earnings can

vary from \$5, if they follow the rule to the letter, to \$10, if they break the rule with each ball.[4] Across a variety of experiments we have elicited rule-following task behavior from 1030 subjects (44 NDL, 354 USA, 238 CAN, 336 TUR, 58 ITA).

2.2 The Dictator Game

In some experiments reported here, a Dictator game was played after the rule following task. A visually and structurally similar two buckets design was used (Figure 2). Subjects chose whether to put the balls in their "own" bucket or the bucket of "another player." The difference between the Dutch version and the US/Canadian version of the Dictator game was that in the Dutch version the buckets were originally empty, and the subjects had to make a choice to which bucket to put each of the 100 balls. In the US version, the buckets were filled with an equal number of balls and subjects had to decide *for each ball* whether to keep it in the bucket where it began or to move it to ("give to" or "take from") another bucket. All subjects made choices as if they were dictators, but subjects were randomly paired and half of the decisions were implemented. Instructions are available in the Appendix. In total we have 180 dictator observations (44 NDL, 70 USA, 66 CAN); this task was always performed immediately after the rule-following task, so that we can test the model.

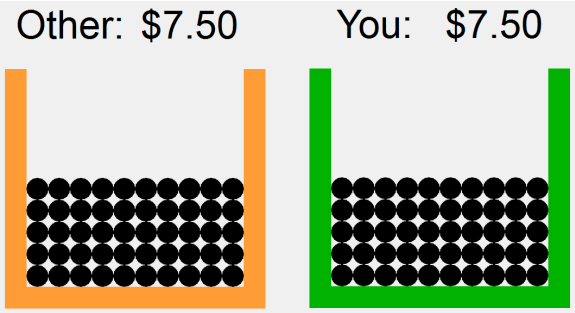


Figure 2: Presentation of the dictator game (USA and Canada).

3 How to Use the Rule Following Task

Following Kimbrough and Vostroknutov (2016), we use the cost incurred following the rule as a measure of an individual's propensity to follow norms in other environments. To give an intuition, consider a norm-dependent utility of subject i in the rule following task. For simplicity suppose that the blue bucket gives \$0 and the yellow bucket \$1 for each ball. Assume i gets linear consumption utility from money and incurs costs from not following the norm. The costs

[4] When subjects ask for clarification about the statement that *"the rule is..."*, we always answered that "this is the rule of the experiment," and when they ask whether anything will happen to them if they don't follow the rule, we told them that all information they need to make their decision is in the instructions.

are higher the less balls are in the blue bucket. The utility from having x balls in the yellow bucket can be written as follows:

$$U_i(x) = x + \phi_i g(x).$$

Here $\phi_i \geq 0$ is the propensity of i to follow norms and $g : \mathbb{R}_+ \to [-1, 1]$ is a function that assigns a normative *social appropriateness* (or inappropriateness) to each *action* x.[5] Function $g(x)$ is assumed to be the unique norm shared by all members of the society which, importantly, is assumed to be independent from individual parameter ϕ_i.

In case of the rule following task $g(x)$ is decreasing in x. Suppose that subject i maximizes her utility. Then, given appropriate assumptions on the shape of $g(x)$, we would have the following equation $\phi_i = -1/g'(x^*)$, where x^* is the optimal choice (observed by us in the experiment). Thus, we obtain a positive monotonic relationship between unobserved ϕ_i and observed x^*. In this sense we consider the observations in the rule following task as a proxy for ϕ_i.

Consider now a Dictator who keeps x dollars in the Dictator game and has the same norm-dependent utility $U_i(x)$. The only difference is that now $g(x)$ reflects the social appropriateness of action x in Dictator game. The social appropriateness of each action can be directly measured using the norm elicitation task introduced in Krupka and Weber (2013), which uses a coordination game to measure beliefs about norms. Figure 3 illustrates.

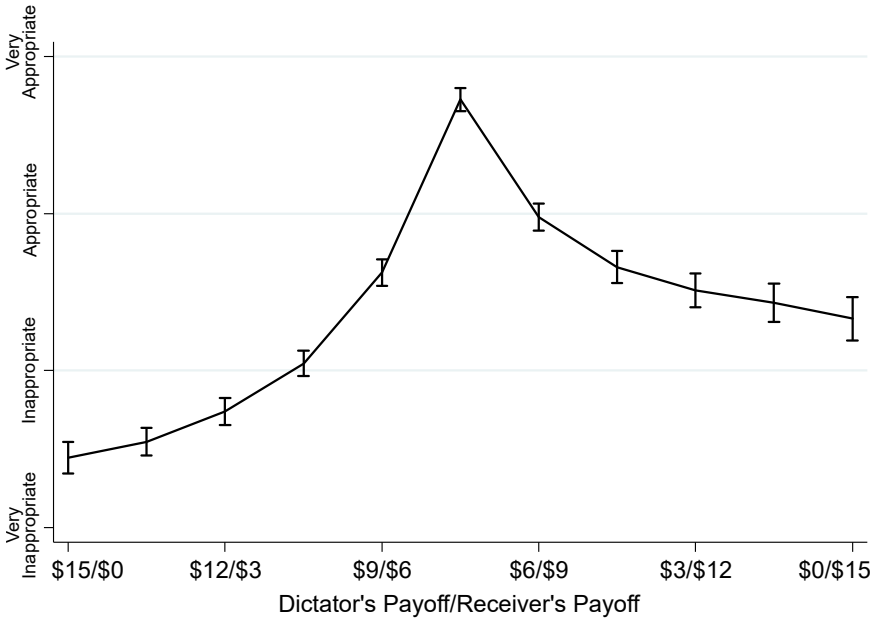


Figure 3: Elicited norms in the dictator game (USA data). Bars indicate +/- 1 SE.

The data was generated by asking separate set of subjects ($N = 88$, US/Canadian data), who did not actually play the game, to assign an appropriateness rating to a set of possible actions by dictators (here we report data from Kimbrough *et al.*, 2014). One of these actions is

[5] When defining norms for games it is crucial that $g(x)$ is a function defined on the action space and not on the outcome space, which coincide in case of the rule following task, but not in general.

chosen at random, and subjects are paid \$10 if their appropriateness rating matches the modal appropriateness rating and \$0 otherwise. This captures the notion of social norms as a set of shared beliefs about the appropriateness of actions in a given setting. Very socially inappropriate rating corresponds to function g equal to -1 and very socially appropriate rating corresponds to g equal 1. One can see that giving half in the Dictator game is considered highly appropriate and giving nothing is thought highly inappropriate. The parameter ϕ_i is measured by our rule following task. Thus, utility maximizers who have high parameter ϕ_i should tend to choose closer to the half/half distribution and utility maximizers with low ϕ_i should choose closer to keeping the entire pie.

4 Results

4.1 Portability of the Rule Following Task

Now we test whether the above utility specification has merit. Our first result concerns the distributions of rule following choices in five countries. Figure 4 shows the respective distributions in the Netherlands, USA, Canada, Turkey, and Italy. All statistical tests reported below are two-tailed.

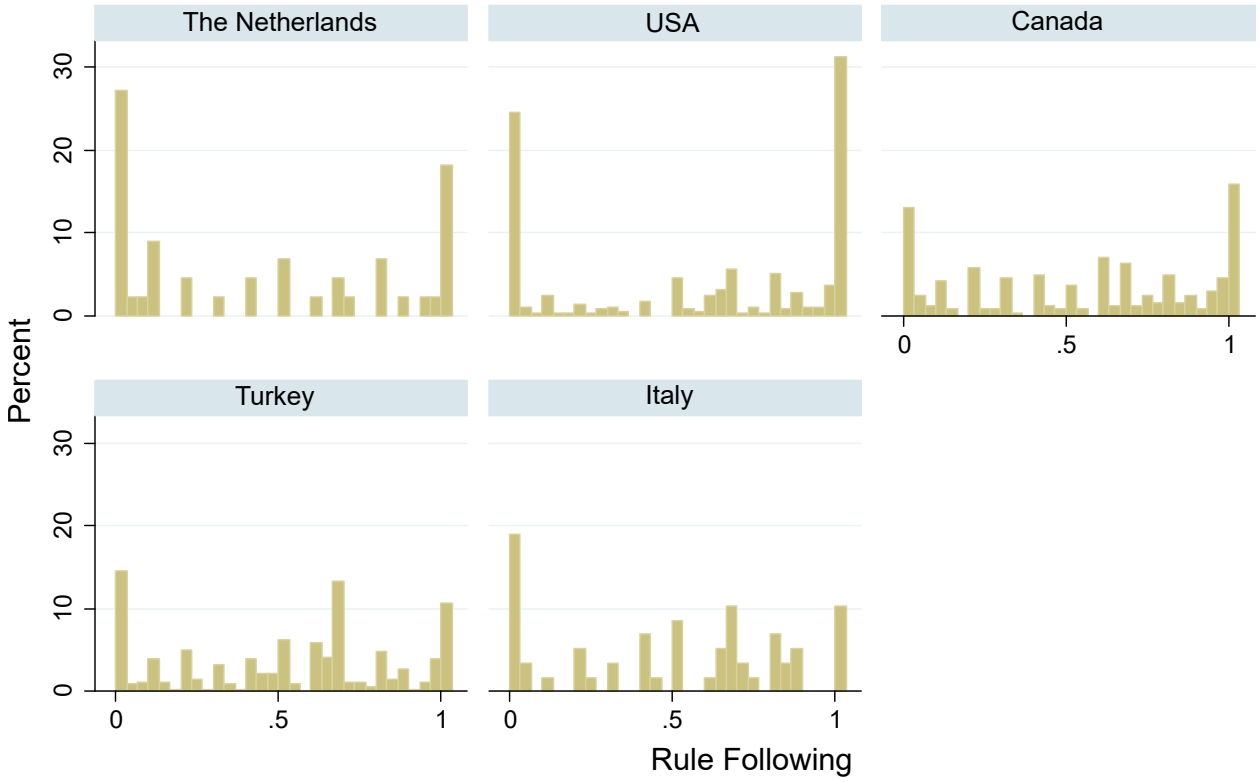


Figure 4: Rule following propensity (percentage of balls in blue bucket) in five countries. Number of observations: 44 in the Netherlands; 354 in USA; 238 in Canada; 336 in Turkey; 58 in Italy.

Almost all distributions show very similar patterns: there are peaks at putting almost all balls in the blue bucket (close to full rule following) and putting almost no balls in the blue bucket (close to full rule breaking). Turkey and Italy show deviation from this: there are much more subjects who choose intermediate number of balls. The distributions are not the same, Kruskal-Wallis test on five groups gives $p = 0.015$.

KS Tests	p-value
NDL-USA	0.079*
NDL-CAN	0.090*
NDL-TUR	0.061*
NDL-ITA	0.426
USA-CAN	0.002***
USA-TUR	0.000***
USA-ITA	0.001***
CAN-TUR	0.049**
CAN-ITA	0.259
TUR-ITA	0.609

Table 1: Kolmogorov-Smirnov tests of the equality of distributions of choices in the rule-following task.

Although, all distributions have point masses at full rule following and full rule breaking, the heights of these peaks differ. For example, in the US sample, the percentage of people who choose to put almost all balls into the blue bucket is almost double that of the Dutch and Canadian samples and almost triple that of Turkish and Italian samples. Table 1 shows the p-values of the Kolmogorov-Smirnov tests of the equality of distributions of choices in the rule-following task. USA sample is strongly significantly different from all countries except the Netherlands. This observation is also supported by a simple OLS regression analysis where (normalized) number of balls in a blue bucket is regressed on USA, Canada, Turkey, and Italy dummies (the Netherlands being a baseline). The coefficient on the USA dummy is 0.140 with $p = 0.029$, which signifies a 14% increase in rule following in the USA as compared to other countries. These findings can potentially be explained by the demographic differences in the samples (see Appendix A).

4.2 Rule Following and Norm Consistent Behavior in the Dictator Game

Next we report correlations between rule following task and Dictator game choices. In all sessions that also contained a Dictator game, the rule following task was presented first. Pooling data from all three countries, Spearman's rank correlation coefficient between the amount sent to the recipient in the dictator game and the share of balls placed into the rule-following bucket is highly significant and positive ($\rho = 0.31$, $p < 0.0001$, N=180). When we analyze the data within-country, we find a strong correlation between the behavior in the rule following task and the dic-

tator game in the Netherlands and USA, but not in Canada. Spearman's correlation coefficient in the Netherlands is $\rho = 0.42$, $(p = 0.006, N = 44)$; in the USA: $\rho = 0.51$, $(p < 0.0001, N = 70)$ and in Canada: $\rho = 0.06$, $(p = 0.62, N = 66)$. OLS regressions of the distance of the amount sent from the equal split and rule following also supports these findings (Table 2). Wald tests reveal that the coefficient on rule following is significant in the Netherlands ($\beta = -0.151, p = 0.035$) and the USA ($\beta = -0.171, p < 0.001$), but not in Canada ($\beta = -0.078, p = 0.251$), though the coefficient has the predicted sign. Moreover, in the USA the subjects give slightly less overall ($\beta = 0.087, p = 0.068$).

Distance of the amount sent from equal split norm		
RF	–0.138***	–0.151**
	(0.033)	(0.071)
USA	0.073**	0.087*
	(0.034)	(0.048)
USA × RF		–0.020
		(0.083)
CAN	0.018	–0.024
	(0.038)	(0.060)
CAN × RF		0.073
		(0.098)
constant	0.363***	0.368***
	(0.031)	(0.040)
N	180	180

Table 2: OLS regression of sharing in the dictator game. Errors are robust. * – $p < 0.1$; ** – $p < 0.05$; *** – $p < 0.01$.

Similar patterns are found in Turkish and Italian samples where different games were coupled with the rule following task. In Italy subjects chose in a series of mini-Dictator games which consisted of a choice between two allocations with constant sum of payoffs for all allocations. The average number of non-selfish choices (and more payoff to the other) correlates with the measure of rule following (Spearman's $\rho = 0.38$, $p = 0.0036, N = 58$). Norm-dependent utility can explain 84% of choices of subjects who chose non-selfish allocation at least once and 75% of choices of subjects who chose non-selfish action more than half of the time (Panizza *et al.*, 2017). In Turkey subjects were choosing in a repeated game and their rule following propensity was also significantly correlated with what could be seen as prosocial behavior in that environment (Gürdal *et al.*, 2017).

4.3 Testing the Independence of RF Behavior and Normative Beliefs

In this final section we test the assumption that subjects' rule following propensities and elicited normative beliefs are independent, which underlies our simple version of the norm-dependent

utility. The subjects whose norm elicitation data were presented in Figure 3 above also completed the rule following task. This allows us to test if there is a systematic difference in rule followers' and rule breakers' beliefs about the prevailing norm. Figure 5 shows elicited norms separately for rule followers and rule breakers (split by the median rule following). Note that both groups rate the purely selfish action as the least appropriate of all actions and the equal split as the most appropriate of all actions, suggesting basic agreement about the relative appropriateness of the actions. Nevertheless, rank-sum tests show significant differences in the normative evaluations of the first five possible dictator actions from giving nothing to giving almost half ($p = 0.048$, $p = 0.001$, $p < 0.001$, $p < 0.001$, $p = 0.016$ from left to right), with rule breakers systematically rating these actions as slightly less inappropriate (or slightly more appropriate) than rule followers. This suggests that rule followers and rule breakers do not fully agree on the *strength* of the norm. One plausible interpretation is a mild self-serving bias for rule breakers who deem less generous actions more socially appropriate than rule followers in order to rationalize their selfishness (Di Tella *et al.*, 2015).

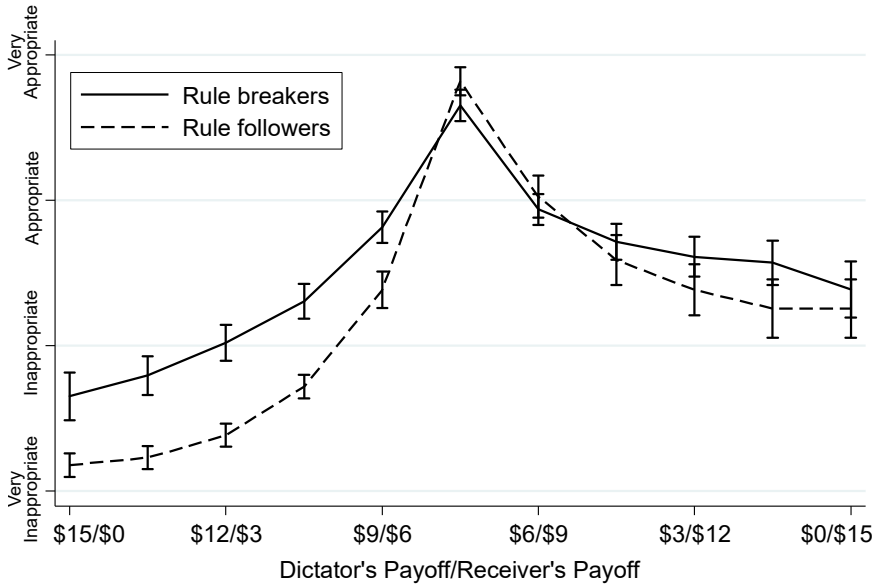


Figure 5: Elicited norms in DG for rule followers and rule breakers. Bars indicate +/- 1 SE.

This evidence points towards the possibility that the assumption in the norm-dependent utility that ϕ_i and g are independent might not hold. One way around this problem is norm elicitation *within* the same subjects who actually play the game and not in a separate group drawn from the same subject pool, as was customary since Krupka and Weber (2013). Indeed, D'Adda *et al.* (2016) and Thomsson and Vostroknutov (2017) have successfully used within subject norm elicitation. The difference from the standard analysis, in which the average norm is used for all subjects, is that now the elicited norm becomes subject dependent and reflects different "opinions" of subjects about the social norm in the environment. When the norm, the rule following parameter, and the choice in the DG are available for each subject, more sophisticated analysis can be performed than when the average norm from a separate experiment is used.

5 Discussion

In this paper we introduce a new method of eliciting individual concern for social norms (in a non-strategic setting), and we provide evidence that prosocial behavior in laboratory experiments is consistent with norm-dependent preferences: a set of preferences under which utility depends on a combination of own payoff and (heterogeneously, according to the parameter ϕ_i) how closely own action adheres to a normative standard. Intuitively, the role of social norms here is quite close to Adam Smith's notion of the 'impartial spectator', in whose eyes an agent wishes his actions to appear praiseworthy (or at least not blame-worthy); see Smith (1759).[6] As with standard models of preferences, agents exhibit non-satiation in own reward, but when social norms suggest non-payoff maximizing behavior (and when agents care about social norms), they will deviate from own-reward maximizing in the direction of the normatively appropriate action. Thus, such a model can be analyzed using standard tools, with choices in individual decisions and in games resulting from optimization by agents heterogeneously weighing the dictates of self-interest against the normative restraints imposed by sociality.

The rule following task we describe is not perfect. In particular, we do not find the correlation between the norm following and prosociality in one out of five countries where experiments were run (Canada). In addition, we see that elicited norms and rule following parameter seem to be correlated in, at least, the Dictator game that we used. In our opinion, this does not indicate that the method is flawed, but rather that there are things we still do not understand about how adherence to social norms explains social behavior. The inconsistencies just mentioned suggest new paths for research and that more work remains be done.

References

ANDREONI, J. and BERNHEIM, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, **77** (5), 1607–1636.

BÉNABOU, R. and TIROLE, J. (2006). Incentives and prosocial behavior. *The American Economic Review*, **96** (5), pp. 1652–1678.

CAPPELEN, A. W., HOLE, A. D., SØRENSEN, E. Ø. and TUNGODDEN, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, **97** (3), 818–827.

CHERRY, T. L., FRYKBLOM, P. and SHOGREN, J. F. (2002). Hardnose the dictator. *American Economic Review*, **92** (4), 1218–1221.

[6] "[A person] desires, not only praise, but praise-worthiness; or to be that thing which, though it should be praised by nobody, is, however, the natural and proper object of praise. He dreads, not only blame, but blame-worthiness; or to be that thing which, though it should be blamed by nobody, is, however, the natural and proper object of blame." (Smith, 1759, §. 3.1)

D'ADDA, G., DROUVELIS, M. and NOSENZO, D. (2016). Norm elicitation in within-subject designs: Testing for order effects. *Journal of Behavioral and Experimental Economics*, **62**, 1–7, cEDEX Working Paper.

DANA, J., WEBER, R. A. and KUANG, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, **33** (1), 67–80.

DI TELLA, R., PEREZ-TRUGLIA, R., BABINO, A. and SIGMAN, M. (2015). Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review*, **105** (11), 3416–3442.

ERKUT, H., NOSENZO, D. and SEFTON, M. (2015). Identifying social norms using coordination games: Spectators vs. stakeholders. *Economics Letters*, **130**, 28–31.

FEHR, E. and SCHMIDT, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, **114** (3), 817–868.

FISCHBACHER, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, **10** (2), 171–178.

GÜRDAL, M., TORUL, O. and VOSTROKNUTOV, A. (2017). The survival of redistributive institutions, in preparation, Boğaziçi University and University of Trento.

HOFFMAN, E., MCCABE, K., SHACHAT, K. and SMITH, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, **7** (3), 346–380.

—, — and SMITH, V. L. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review*, **86** (3), 653–60.

KESSLER, J. B. and LEIDER, S. (2012). Norms and contracting. *Management Science*, **58** (1), 62–77.

KIMBROUGH, E. O., MILLER, J. B. and VOSTROKNUTOV, A. (2014). Norms, frames and prosocial behavior in games, mimeo.

— and VOSTROKNUTOV, A. (2015). The social and ecological determinants of common pool resource sustainability. *Journal of Environmental Economics and Management*, **72**, 38–53.

— and — (2016). Norms make preferences social. *Journal of the European Economic Association*, **14** (3).

KRUPKA, E. L. and WEBER, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, **11** (3), 495–524.

LAZEAR, E. P., MALMENDIER, U. and WEBER, R. A. (2012). Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*, **4** (1), 136–163.

LIST, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, **115** (3), 482–493.

LÓPEZ-PÉREZ, R. (2008). Aversion to norm-breaking: A model. *Games and Economic Behavior*, **64** (1), 237–267.

PANIZZA, F., VOSTROKNUTOV, A. and CORICELLI, G. (2017). Transmission of social norms, in preparation, University of Trento.

SMITH, A. (1759). *The Theory of Moral Sentiments*. Liberty Fund: Indianapolis (1982).

THOMSSON, K. and VOSTROKNUTOV, A. (2017). Small-world conservatives and rigid liberals: Attitudes towards sharing in self-proclaimed left and right. *Journal of Economic Behavior and Organization*, **135**, 181–192.

Appendix (for online publication)

A Demographics

In this appendix we explore the connection between our behavioral measures and various demographic characteristics. Table A1 shows summary statistics of the demographic data.[7] The samples are relatively similar in age, percent male and years of study, with the Dutch students having slightly fewer years at University on average. Also, though Economics and related fields are the most common in all five countries, the US sample has fewer other Social Science students and somewhat more in the Arts and Humanities. The Dutch sample has the lowest share of students studying Natural Sciences. Self-reported ethnic composition varies widely across the samples, with the Dutch sample being mostly composed of Europeans, the American sample mostly composed of whites and "Americans", and the Canadian sample divided roughly 1/3 to 2/3 between "Canadians" and Asians. The Italian sample is composed almost exclusively of Italians, and the Turkish sample is composed entirely of Turks.

	Full RF Sample					DG Sample		
	NDL	USA	CAN	TUR	ITA	NDL	USA	CAN
Num. Subjects	44	354	238	336	58	44	70	66
Rule Following	44.5%	58.5%	54.8%	52.0%	49.6%	44.5%	60.3%	61.2%
Dictator Giving						19.9%	18.8%	22.7%
Demographics								
Males	52.3%	57.6%	55.0%	56.6%	51.7%	52.3%	58.6%	57.6%
Years of Study	2.11	2.76	2.80		2.66	2.11	2.63	2.70
Age	21.75	21.77	21.84	21.35	22.0	21.75	21.44	22.08
Field of Study								
Economics/Business	63.6%	58.2%	38.7%	38.4%	63.8%	63.6%	58.6%	39.4%
Social Sciences	15.9%	5.9%	19.3%		29.3%	15.9%	10.0%	4.5%
Arts & Humanities	9.1%	13.0%	8.4%			9.1%	10.0%	9.1%
Natural Sciences	11.4%	22.6%	32.8%		5.2%	11.4%	21.4%	45.5%
Other				61.6%	1.7%	11.4%	21.4%	45.5%
Ethnicity								
"American"		48.6%	1.7%				44.3%	
"Canadian"		0.3%	30.3%					22.3%
White/European	88.6%	24.6%	5.5%		100%	88.6%	27.1%	4.5%
Hispanic		6.0%	1.7%				8.6%	3.0%
Black		7.3%	1.3%				7.1%	
Asian	11.4%	13.0%	58.8%	100%		11.4%	12.9%	70.0%

Table A1: Summary statistics of behavioral and demographic data.

Table A2 shows an OLS regression of the (normalized) number of balls in the blue bucket (i.e. the extent of rule following) regressed on economics/business field, ethnicity, gender, and age. The baseline of the regression is female, undeclared ethnicity with field of study being not economics or related fields (see Table A3 for details). Wald tests indicate that subjects in the US are more rule-following than both Italian and Dutch subjects. All other country comparisons are insignificant. And, consistent with our observations in Kimbrough and Vostroknutov (2016), males are more rule breaking.

[7] See Table A3 for the classification of fields of study and nationalities.

Rule Following	β	Std.Err.
Economics/Business	–0.014	(0.024)
"American"	0.145	(0.258)
"Canadian"	0.151	(0.254)
White	0.167	(0.258)
Hispanic	0.252	(0.265)
Black	0.103	(0.264)
Asian	0.138	(0.252)
Male	–0.083***	(0.024)
Age	0.001	(0.003)
USA	0.156**	(0.069)
Canada	0.122	(0.078)
Turkey	0.102	(0.077)
Italy	0.047	(0.076)
Constant	0.305	(0.271)
N	1030	

Table A2: OLS regression of the rule following propensity. Errors are robust. * – $p < 0.1$; ** – $p < 0.05$; *** – $p < 0.01$.

Field of Study	**Definition**
Economics (baseline)	Economics, finance, business, international business, business economics, accounting, marketing, econometrics (and operations research), human decision sciences, supply chain management, banking, actuarial, economics and psychology, business and psychology, human resource management, public administration
Social Sciences	Psychology, sociology, cognitive science, hospitality management, anthropology, law, political science
Arts & Humanities	Arts, journalism, University College Maastricht, music, languages, european studies, architecture, sports studies, higher education, international relations, apparel studies, history, geography
Natural Sciences	Computer science, informational systems, biology, animal science, environmental sciences, engineering of all types, (bio)chemistry, physics, mathematics, biomedical research, health sciences, communication disorders, kinesiology, dietetics, nutrition studies, nursing
Ethnicity	**Definition**
N/A (baseline)	Not specified
"American"	Ethnicity specified as "American"
"Canadian"	Ethnicity specified as "Canadian"
White	All White Europeans including Spanish and Portuguese, Caucasian or white Americans and Canadians
Hispanic	All Spanish (Portuguese) speaking non-Europeans (Latin and Southern America)
Black	African Americans, African Canadians and Africans
Asian	All nations from Middle East and Asia and Native Americans

Table A3: Classification of fields of study and ethnicities.

B Instructions

B.1 RF Task Instructions

General information You are now participating in a decision making experiment. If you follow the instructions carefully, you can earn a considerable amount of money depending on your decisions and the decisions of the other participants. Your earnings will be paid to you in CASH at the end of the experiment

This set of instructions is for your private use only. During the experiment you are not allowed to communicate with anybody. In case of questions, please raise your hand. Then we will come to you and answer your questions privately. Any violation of this rule excludes you immediately from the experiment.

Part I In Part I of this experiment, you will decide how to allocate 100 [50] balls between two buckets.

Your task is to put each of the balls, one-by-one, into one of the two buckets: the blue bucket or the yellow bucket. The balls will appear in the center of your screen, and you can allocate each ball by clicking and dragging it to the bucket of your choice. For each ball you put in the blue bucket, you will receive 5 cents, and for each ball you put in the yellow bucket, you will receive 10 cents.

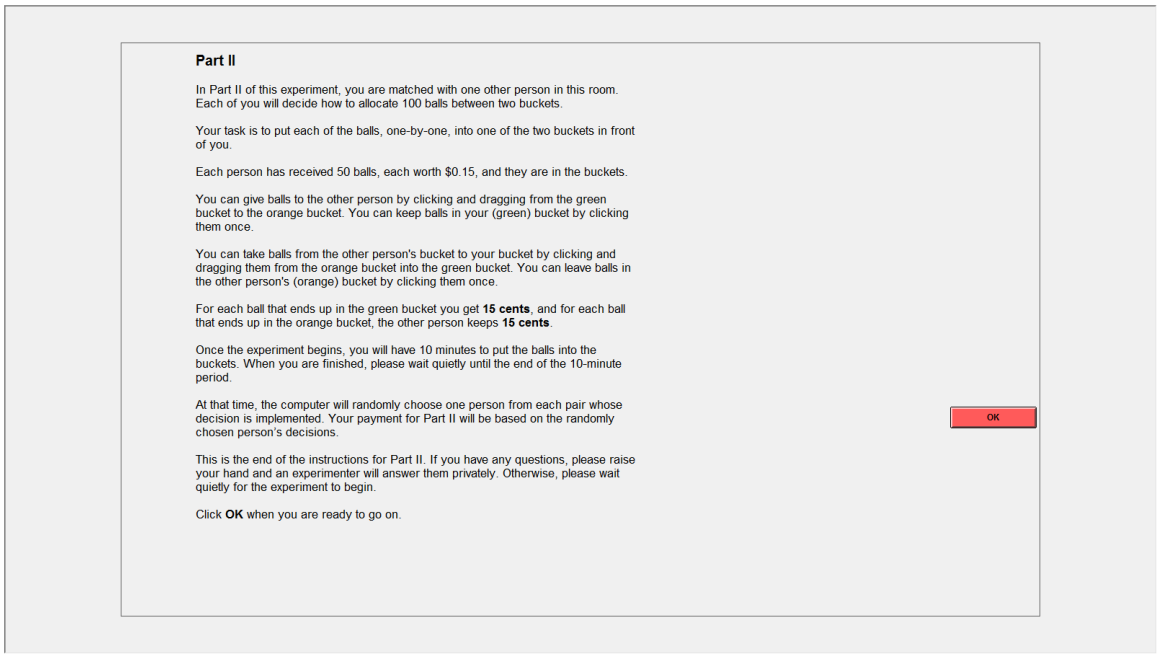
The rule is to put the balls in the blue bucket.

Once the experiment begins, you will have 10 minutes to put the balls into the buckets. When you are finished, please wait quietly until the end of the 10-minute period.

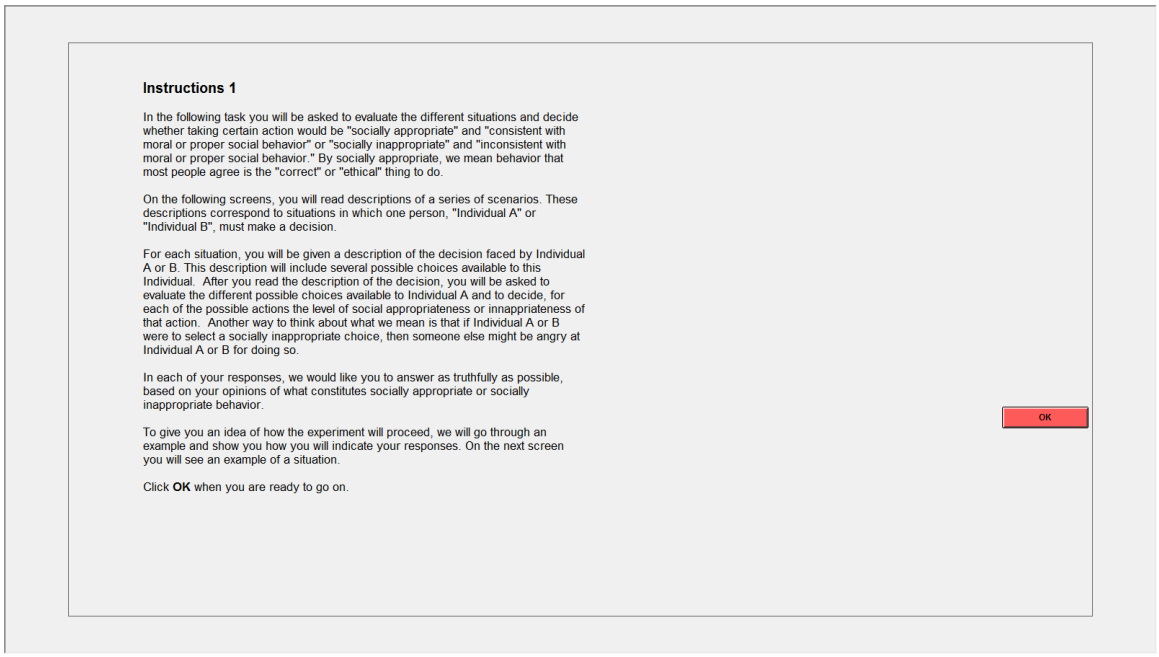
Your payment from Part I will be based on your decisions: it is the sum of payments from the blue and yellow buckets.

This is the end of the instructions for Part I. If you have any questions, please raise your hand and an experimenter will answer them privately. Otherwise, please wait quietly for the experiment to begin.

B.2 DG Instructions (USA)



B.3 Norm Elicitation Instructions (USA)



Instructions 2

In what follows, you will read a description of each scenario. Then, you will indicate your appropriateness rating by placing a check mark in the corresponding box.

At the end of the experiment today, the computer will randomly select one of the situations. For this situation, the computer will also randomly select one of the possible choices that the Individual could make. Thus, we will select both a situation and one possible choice at random. Each choice is equally likely to be selected.

For the choice selected, we will determine which response was selected by the most people here today.

If you give the same response as that most frequently given by other people, then you will receive an additional $10. This amount will be paid to you, in cash, at the conclusion of the experiment.

For instance, if we were to select the example situation from the last screen and the possible choice "Leave the wallet where it is," and if your response had been "somewhat socially inappropriate," then you would receive $10, in addition to the $7 participation fee, if this was the response selected by most other people in today's session. Otherwise you would receive only the $7 participation fee.

Your payment depends only on whether your response corresponds to the most common response to the chosen and DOES NOT depend on the decisions made by individual A or B.

OK

Please click **OK** when you are ready to go on.

If you have any questions, please raise your hand and wait for the experimenter.

Example Situation

Individual A is at a café. While there, Individual A notices that someone has left a wallet at one of the tables. Individual A must decide what to do. Individual A has four possible choices: take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the bartender.

Individual A can choose only one of these four options. The table on the right presents a list of the possible actions available to Individual A. For each of the actions, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please click on the corresponding cell.

Please make sure you make an assessment for each possible choice in each row of the table.

In what follows, you will be asked to assess the appropriateness of the actions in two other scenarios. For each action please indicate the extent to which you believe taking that action would be "socially appropriate" and "consistent with moral or proper social behavior" or "socially inappropriate" and "inconsistent with moral or proper social behavior." By socially appropriate we mean behavior that ***most people*** agree is the "correct" or "ethical" thing to do.

Individual A's Action	Very Socially Inappropriate	Somewhat Socially Inappropriate	Somewhat Socially Appropriate	Very Socially Appropriate
Take the wallet				√
Leave the wallet where it is			√	
Give the wallet to the bartender		√		
Ask others nearby if the wallet belongs to them	√			

OK

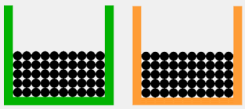
Scenario 1

Individual A has been invited to an experiment and paired with another anonymous individual B so that neither individual will ever know the identity of the other individual with whom he/she is paired.

In the experiment, Individual A will make a choice, the experimenter will record this choice, and then both individuals will be informed of the choice and paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither individual will receive any other money for participating in the experiment.

Individual A observes two buckets on the computer screen. One of the buckets is green and labeled "You" and one of the buckets is orange and labeled "Other". Each bucket contains 50 balls.

You: $7.50 Other: $7.50



Individual A reads the following instructions: *"Each person has received 50 balls, each worth $0.15, and they are in the buckets. You can give balls to the other person by clicking and dragging from the green bucket to the orange bucket. You can keep balls in your (green) bucket by clicking them once. You can take balls from the other person's bucket to your bucket by clicking and dragging them from the orange bucket into the green bucket. You can leave balls in the other person's (orange) bucket by clicking them once. For each ball that ends up in the green bucket you get* ***15 cents****, and for each ball that ends up in the orange bucket, the other person keeps* ***15 cents****."*

Now, look at the table on the right side of the screen and consider eleven possible actions that Individual A could take.

For each of the actions, please indicate whether you believe choosing that action is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate.

Remember: when we select a scenario and an action for payment, you will only receive the additional $10 if your response is the same as the most frequent response made by other people.

Individual A's Action	Very Socially Inappropriate	Somewhat Socially Inappropriate	Somewhat Socially Appropriate	Very Socially Appropriate
Take 50 balls to the green bucket, leave 0 balls in the orange bucket: A gets $15, B gets $0			√	
Take 40 balls to the green bucket, leave 10 balls in the orange bucket: A gets $13.5, B gets $1.5				√
Take 30 balls to the green bucket, leave 20 balls in the orange bucket: A gets $12, B gets $3			√	
Take 20 balls to the green bucket, leave 30 balls in the orange bucket: A gets $10.5, B gets $4.5		√		
Take 10 balls to the green bucket, leave 40 balls in the orange bucket: A gets $9, B gets $6	√			
Keep 50 balls in the green bucket, leave 50 balls in the orange bucket: A gets $7.5, gets B $7.5		√		
Keep 40 balls in the green bucket, Give 10 balls to the orange bucket: A gets $6, B gets $9			√	
Keep 30 balls in the green bucket, Give 20 balls to the orange bucket: A gets $4.5, B gets $10.5				√
Keep 20 balls in the green bucket, Give 30 balls to the orange bucket: A gets $3, B gets $12			√	
Keep 10 balls in the green bucket, Give 40 balls to the orange bucket: A gets $1.5, B gets $13.5		√		
Keep 0 balls in the green bucket, Give 50 balls to the orange bucket: A gets $0, B gets $15	√			

OK