

**Krupka, E. & Weber, R. (2013): Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?**

*Journal of the European Economic Association*

**Introduction**

- Why are social norms relatively absent in economics?
  - They are difficult to measure and quantify and therefore, it is hard to obtain precise predictions on their influence on behavior.
  - They were used in previous research as a tool for explaining anomalous behavior (involuntary unemployment, tipping, costly punishment, etc.).
  - Usually studied indirectly and used as a post hoc explanation.
- Recent experiments have shown that minor contextual features lead to variance in choices and outcomes in dictator games, consistent with variations in and preference for compliance with social norms.
- The paper introduces a novel incentivized method for identifying social norms separately from behavior, especially the behavior directed to others in both variants of the dictator game.
- Assumption 1: Individuals care about behaving in compliance with social norms.
- Assumption 2: Decision-maker's utility is based on the money they receive and the extent to which their action complies with social norms (taking socially appropriate actions).

**Defining and Identifying Social Norms**

- **Two important features:**
  - Prescribe or proscribe behaviors and actions, rather than outcomes
  - Jointly recognized or collectively perceived
- The article focuses on **Injunctive norms** (“ought” to do rules), rather than Descriptive norms (customs or actions regularly taken by members of a group). The authors claim that these kinds of norms are sufficient for explaining a considerable amount of variation in other-regarding behavior.
- **Injunctive norms:** collective perceptions, among members of a population, regarding the appropriateness of different behaviors; members are expected to be aware of them.
  - The power of social norms comes from:
    - “... [the] willingness of people within the population to punish (or reward) others’ deviation from (or adherence to)”
    - “...the experience of positive or negative emotions produced by one’s own adherence or deviation from a social norm.”

- **Formalized definition (Choice model):**

- Set of  $K$  actions available to a decision-maker:  $A = \{a_1, \dots, a_K\}$
- Social norm:  $N(ak) \in [-1,1]$  defines the degree to which an action is collectively perceived as appropriate or inappropriate.
- Assume that an action is collectively recognized as:
  - Appropriate, then:  $N(ak) > 0$
  - Inappropriate, then:  $N(ak) < 0$
- $u(ak) = V(\pi(ak)) + \gamma N(ak)$ 
  - $\pi(ak)$  is the payoff from taking a certain action.
  - $\gamma$  is the degree to which an individual cares about adhering to social norms.
  - $\gamma = 0$  then the individual does not care about social norms and will select a payoff-maximizing action and vice versa.
- Example:
  - *Two choice environments:  $A = \{a_1, a_2\}$  and  $A' = \{a_1', a_2'\}$ , such that  $\pi(a_1) = \pi(a_1') > \pi(a_2) = \pi(a_2')$*
  - *If there are no social norms in either environment ( $N(ak)=N(ak')=0$ , for  $k=1,2$ ) the decision maker will choose  $a_1$  in the first environment and  $a_1'$  in the second.*
  - *If social norms differ between the two choice environments, the individual may select actions corresponding to different payoffs in the two environments. For example, if  $N(a_1) = N(a_1') < N(a_2) < N(a_2')$ , then for some values of  $\gamma$  a decision maker will select  $a_1$  in the first environment and  $a_2'$  in the second environment.*
- Actions vary in the degree they are perceived as socially appropriate, rather than having a binary classification.
  - *“For example, while there may be social agreement that it is always appropriate to arrive on time in many Western cultures, there may be some instances in which arriving five minutes late is less socially inappropriate (meeting friends at a bar) than others (arriving at a funeral).”*

### **Incentivized elicitation method (Coordination game)**

- Present responders with a description of a choice environment (including all possible actions).
- Ask respondents to make a judgment on the social appropriateness of each action (four-point scale: “very socially inappropriate”, to “very socially appropriate”).
- Respondents play a *matching coordination game*, with the goal of anticipating the social appropriateness rating of others.
- The elicitation method will yield a representation of a social norm if:
  - (a) there is a general social agreement that some actions are more or less socially appropriate.
  - (b) respondents attempting to tacitly match others’ responses relying on such shared perceptions to help them do so.

## Experiment 1 (Identifying Social Norms in Payoff-Equivalent Environments):

- “Standard” versus “bully” dictator game: both types offer the same choices to the decision-maker but vary the actions (“*giving* versus *taking*”).
  - **Standard:** The dictator receives \$10 and is given the choice to transfer (i.e., give) part of the money to a recipient (e.g, the dictator transfers \$1 to the recipient and keeps \$9: outcome \$9, \$1).
  - **Bully:** Both the dictator and recipient receive \$5. The dictator is given the choice to transfer (i.e., take) part of the money from the recipient to him/herself (e.g, the dictator transfers \$4 from the recipient to him/herself: outcome \$9, \$1).
- The authors argue that it is possible that social norms governing the two sets of actions are contextually different, despite the fact that outcomes are the same.
- 199 subjects at Carnegie Mellon University, the University of Pittsburgh, and the University of Michigan
- **Description of the Experiment**
  - Participants were told that one of the situations for which they were to provide appropriateness ratings would be selected at random at the end of the session and that one of the possible action choices in this situation would also be randomly selected.
  - If for this action choice, the participant’s appropriateness rating was the same as the modal response in the session, then that participant would receive an additional payment (\$5 in Pittsburgh, \$10 in Michigan) at the conclusion of the session (i.e. incentivized elicitation method).
  - They rated actions in either the bully or standard dictator game variants.
- **Results**
  - There is a substantial social agreement that the action that produces equal payoffs (\$5, \$5) is very socially appropriate in either environment.
  - Maximizing A’s own payoff and leaving the other person with nothing (\$10, \$0) is the most socially inappropriate action in either variant.
  - Actions involving “taking” are generally less appropriate than those involving “giving”. Subjects are able to anticipate the rating of others in modal responses (when the ratings differ in two variants).
  - For every outcome from (\$9, \$1) to (\$6, \$4), the mean rating for the corresponding action is higher in the standard (giving) context than in the bully (taking) context, and these differences are all highly statistically significant.
  - For example, for the wealth allocation (\$8, \$2), the modal response in the standard context for giving \$2 to the other person is “somewhat inappropriate”. But in the bully context, where the same outcome involves taking \$3 from the other person, there is a social agreement that the action is “very inappropriate”.

## Experiment 2 - Evaluating predictions with behavioral choice data

- A different set of subjects in one of the two choice environments (i.e., standard or bully DG)
- **Standard dictator game:**
  - Each dictator received a yellow envelope labeled “money for you” that contained ten \$1 bills. The other group (recipients) received empty white envelopes labeled “money for other person”.
  - Instructions were read aloud describing the choice, in which dictators would make a (double-blind) anonymous decision of how much of the \$10 to put in their envelope to share with the paired recipient.
  - The dictator privately allocated money between the two envelopes placed the white envelope labeled “money for other person” inside the box and left with whatever remained in the yellow envelope (This procedure allowed individual decisions to be anonymous).
- **Bully variant of the dictator game:**
  - Procedures were identical except that the two envelopes handed out at the beginning of the experiment each contained five \$1 bills.
  - The instructions informed dictators that they would be able to give up to \$5 to or take up to \$5 from the other person.
- **Behavioral Predictions**
  - Prediction 1: *More agents will select the action producing the equal-split (\$5, \$5) allocation in the bully game than in the standard game.*
    - In the bully game, every combination with higher payoffs than (5,5) for the DM will be less socially appropriate than in the standard game.  
→ So choosing an action that produces equal-split will be more attractive.
    - This depends on: (i) utility loss from foregoing the higher monetary payoff,  $V(\$5) - V(\pi(ak)) < 0$ , on (ii) the degree to which the two actions differ in social appropriateness,  $N(a(\$5,\$5)) - N(ak) > 0$ , and on (iii) concern for norm compliance  $\gamma$ .
    - The difference in norm ratings is always higher in the bully context than in the standard, according to the results in Table 1.
  - Prediction 2: *Conditional on not selecting the action producing the equal-split (\$5, \$5) allocation, more agents will select the action producing the payoff-maximizing (\$10, \$0) allocation in the bully game than in the standard game.*
    - The difference between  $N(a(\$10,\$0))$  and  $N(ak)$  is smaller in the bully context than in the standard dictator game, for all  $ak \in \{a(\$6,\$4), \dots, a(\$9,\$1)\}$ .
    - If individuals select something other than an equal split, then by the prediction, they will take all the wealth.

- **Results**
- Prediction 1:
  - Standard condition: 8 out of 48 participants (17%) gave \$5 to the recipient.
  - Bully treatment: 18 out of 49 participants (37%) neither took from nor gave money to the recipient.
  - \*  $p < .001$
- Prediction 2:
  - In the standard game, 40 participants gave less than \$5 to the recipient, and of these 16 (40%) gave \$0.
  - In the bully variant, 31 participants took money from the recipient, and of these 16 (52%) left the recipient with \$0.
  - \*  $p = .03$
- Shared an amount from \$1 to \$4:
  - Standard game: 24 of 52 subjects (46%).
  - Bully variant: 15 of 54 subjects (28%).
  - \*  $p = 0.05$
- Mean amount allocated to the recipient:
  - Standard Game = \$2.46
  - Bully variant = \$3.11
- The ratio,  $2\gamma/\beta$ , identifies how much money an individual is willing to sacrifice to take an action that is very socially appropriate ( $N(ak) = 1$ ) rather than one that is very socially inappropriate ( $N(ak) = -1$ ). This ratio indicates that subjects are willing to pay \$5.66 to comply with social norms.

### Reanalyzing Previously Collected Dictator Game Data

- The authors show that elicited ratings are consistent with the behavior in the experiments and that the results produced by specific variants of the dictator game can be accounted for by changes in the social appropriateness of actions.
- **Dictator Game With a Sorting Option**
  - Lazear, Malmendier, and Weber (2012) explored a variant of the dictator game in which subjects could opt to not play the game (by “passing”), in which case the dictator received \$10 and the other participant received \$0 without learning that a dictator game could have been played.
  - **Findings:** Subjects appear to care the same about social norms in the dictator game based on whether subjects were required to play the game (standard) or had the option of not playing the game but then choosing to play (Sorting). Opting out is considered

far less socially inappropriate. Subjects are willing to pay large and very similar amounts to comply with social norms (\$5.68).

- **Dictator Game with Additional Taking Options**

- List (2007): In a standard variant, dictators divided \$5 between themselves and another participant, in \$0.50 increments. In a “Take \$1” variant, dictators could alternatively take \$1 from the recipient, an option selected by many participants.
- **Findings:** Giving small amounts to the recipient is more socially appropriate if the DM could have taken money instead (Take \$1). Introducing the option of Taking \$1, makes “keeping all” less socially inappropriate and it may be attractive to DMs because they can get a higher payoff than in the standard game. Again, norm compliance is fairly constant across the experimental conditions, but subjects are willing to pay approximately \$2.67 to take an action that is very socially inappropriate rather than one that is very socially appropriate.

- **The Stability of Preferences Across Experiments**

- The authors claim that subjects are generally willing to pay around \$5 to take actions that are socially appropriate.
- When including binary dictator game controls, where dictators were unaware of the payoff consequences, but could get this information easily without any cost, the authors found that many dictators did not choose to acquire information and increased selfish behavior. The analysis predicts that the proportion of such fair behavior declines substantially in the hidden information treatment and that many subjects will choose to remain willfully ignorant.

## **Conclusion**

- Novel incentivized method for identifying social norms using coordination games
- Elicited social appropriateness ratings can predict behavioral changes across variants of the dictator games when DM cares about norm compliance and payoff.
- **Limitations**
  - Subjects could coordinate in ways that have nothing to do with norms
  - Under significantly higher stakes, predictions are very likely to require different sets of weights on money and social appropriateness.
  - Behavior may be more complex in other games (public goods games and trust games, with reciprocity and uncertainty).