Judgment and Decision Making (2024), **Vol. 19**:e9 1–21
doi:10.1017/jdm.2024.1

EMPIRICAL ARTICLE

Negative economic shocks and the compliance to social norms

Francesco Bogliacino[1], Rafael Charris[2], Camilo Gómez[3], and Felipe Montealegre[4]

[1]Dipartimento di Scienze Economiche, Università degli Studi di Bergamo, Bergamo, Italy; [2]Epidemiology Division, Department of Population Health, Grossman School of Medicine, New York University, New York, NY, USA; [3]CERGE-EI, a joint workplace of Charles University and the Economic Institute of the Czech Academy of Science, Prague, Czech Republic and [4]University of Bologna, Bologna, Italy

Corresponding author: Francesco Bogliacino; Email: francesco.bogliacino@unibg.it

Received: 10 August 2023; **Revised:** 23 October 2023; **Accepted:** 9 January 2024

Keywords: negative economic shocks; social norms; norm compliance; antisocial behavior; cooperation

Abstract

We study why suffering a negative economic shock, i.e., a significant loss, may trigger a change in other-regarding behavior. We conjecture that people trade off concern for money with a conditional preference to follow social norms and that suffering a shock makes extrinsic motivation more salient, leading to more norm violation. This hypothesis is grounded on the premise that preferences are norm-dependent. We study this question experimentally: after administering losses on the earnings from a real-effort task, we analyze choices in prosocial and antisocial settings. To derive our predictions, we elicit social norms for each context analyzed in the experiments. We find evidence that shock increases deviations from norms.

1 Introduction

A negative economic shock (an NES from here on) is a large financial loss on earnings or accumulated assets. NES can be the outcome of psychosocial stressors (divorces, job losses, and injuries) or traumatic events (violence and disasters). The literature has extensively studied NES to grasp the effect of poverty (Carvalho et al., 2016; Haushofer and Fehr, 2014; Mani et al., 2013). In fact, while comparing the rich and the poor is confounded by environmental and individual factors, under certain circumstances, NES may be a plausible source of income or wealth variation or can be manipulated in the lab. More recently, victims of NESs have tended to support extreme candidates at the elections, which has furthered the interest in how suffering an NES shapes behavior (Benczes and Szabó, 2023; Guiso et al., 2017; Inglehart and Norris, 2016; Rodrik, 2021).

When figuring out how NES affects social behavior, the literature mainly proposes variants of the collective threat argument. The latter goes as follows: facing the risk of aggregate shocks, societies develop tight cultural traits such as social sanctioning and adherence to social norms (Gelfand et al., 2017; Heinrich, 2020; Prediger et al., 2014; Roos et al., 2015; Szekely et al., 2021). For the aim of this article, a social norm is defined as a rule of behavior which is contingent and supported by expectations of other people's behavior and perception of appropriateness (Bicchieri, 2006). Similarly, exposure to warfare induces parochial prosociality as a form of collective insurance (Bauer et al., 2016). Variants of this argument are extended to NES experienced due to natural disasters (Botchway and Filippin,

2021; Cassar et al., 2017). This reasoning neglects a point, though. Suffering an NES independently affects whether subjects follow social norms because this experience alters the relative cost of norm compliance. Even conditional on the same threat, the simple fact that the impact is heterogeneous induces variations in norm compliance.

This article tries to address this issue. We conjecture that decision-makers (DMs) trade off money and compliance to social norms (Bicchieri, 2006; Kimbrough and Vostroknutov, 2018; Krupka and Weber, 2013; List et al., 2004). Following social norms is costly. Punishing transgressors, avoiding cheating, and abstaining from free-riding involve carrying out a cost. Suppose the DM trades off the concern for money and the conditional preference to follow the social norm. In that case, she will face an increasing marginal cost of norm compliance when experiencing an NES, leading to more norm violations.

We derive this prediction from a behavioral model. Assuming that norms enter the utility function and participants are heterogeneous in their psychological cost of compliance, we analyze optimal behavior in several binary decision problems ($d \in \{0, 1\}$) where a substantive norm applies, considering both antisocial and prosocial tasks. In all these settings, acting (i.e., $d = 1$) potentially harms the opponent. Sometimes the action is *prescribed* by the norm, as in settings where the norm recommends to (costly) punish. Sometimes this action is *proscribed* by the norm, as in settings where the norm recommends abstaining from cheating or stealing. The model predicts that we should observe more norm violations after experiencing an NES. We design 3 experiments to assess this prediction. The critical design choice is to manipulate NES by inducing significant losses (80%) on the earnings from a real-effort task (RET) (Bogliacino and Montealegre, 2020). Those participants randomly assigned to keep their money constitute the comparison group. After this initial stage, participants interact in one (or multiple) tasks, where we measure the change in norm compliance between the treatment and the control. The settings include stealing, cheating, joy of destruction (JoD), and cooperation.

Since the predictions are conditional on social norms, we elicit the normative expectations for each situation using Bicchieri and Xiao's (2009) methodology. Participants provide their personal normative beliefs (PNBs) over the action space of the DM. They are then asked to guess the modal response to the PNBs (under a simple incentive scheme). The latter responses measure the normative expectations and reveal whether a norm applies to the setting. Participants in this norm elicitation task did not make an actual choice to ensure that we elicit social norms separately from behavior (Krupka and Weber, 2013).

As predicted, subjects steal more and cheat more after suffering an NES. The increase in stealing is almost one-fourth of a standard deviation (sd) (calculated on the outcome variable in the control). In the die-under-the-cup task (Fischbacher and Föllmi-Heusi, 2013), where participants are paid according to the number that they *report* from rolling a die, they are 14% more likely to report 4 and 5, the number with the highest payoff. The effect is equivalent to around one-fourth of an sd. When we look at the JoD, the *decrease* in retaliation is as large as 50% of an sd, again supporting the prediction of the model. However, we cannot conclude that NES increases defection in the prisoner's dilemma (PD).

To document the robustness of our argument on shocks and norm compliance, we conducted a further experiment, where (a) we imposed a rule instead of relying on an elicited norm and (b) we separately controlled for the wealth effect. We found that the effect of an NES over norm compliance is robust and distinguishable from a pure wealth effect.

In Section 8 of this article, we will argue that our lab-based evidence is externally valid since it is consistent with a sizable amount of quasi-experimental evidence on NES and antisocial behavior (Bignon et al., 2017; Boonmanunt et al., 2020; Cortés et al., 2016; Dube and Vargas, 2013). More importantly, it outperforms alternative explanations for the same findings.

This article makes a contribution to the literature on the consequences of shocks. Shocks are multifaceted phenomena, where several factors concur to shape behavior. We studied individual-level shocks, whereas the bulk of the literature focuses on group shocks (Gelfand et al., 2017; Heinrich, 2020; Prediger et al., 2014; Roos et al., 2015; Szekely et al., 2021). Our contribution is complementary: experiencing the shock has a different effect from exposure to the threat. This differential behavioral effect of experience vs. information has been documented in works unrelated to social norms (D'Acunto

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

et al., 2021; Malmendier, 2021a, 2021b). By focusing on random shock, we also distinguish ourselves from the papers on framed shocks (e.g., natural disaster) or intentional (i.e., violent) shocks (Bauer et al., 2016; Bogliacino et al., 2023; Eckel et al., 2022). A further discussion on the different aspects of shock is provided in Section 8 below.

Since norms are scripts that humans partially incorporate into their preferences (Gintis, 2007), it is not surprising that people manipulate or elude norms if allowed to do so (Andreoni and Bernheim, 2009; Bicchieri, 2010; Bicchieri et al., 2021; Chavez and Bicchieri, 2013). Dictator games are widely used in this literature to avoid confounds from strategic beliefs. Dana et al. (2007) introduce the concept of *moral wiggle room* to explain why when settings change, but the action space does not, subjects behave more egoistically. List (2007) documents a sizable behavioral change following minimal variation in the action space. Instead of relying on contextual changes, we provide evidence from indirect incentive effects.

The literature on shocks is now rapidly expanding. In experimental settings, the manipulation of losses or windfalls has been used to study poverty or scarcity, usually exploiting paycheck variation or natural experiments. This literature focused on the cognitive impact: Mani et al. (2013) found a negative effect in sugarcane farmers in India, Carvalho et al. (2016) found no effect, but paycheck variations are temporary, expected and expected to be temporary, while Bogliacino and Montealegre (2020) found a negative effect of NES on cognitive performance in the lab. Haushofer and Fehr (2014) claim that suffering NES (and in general, poverty) increases stress, which induces lower risk propensity (in the gain domain) and higher present bias, further worsening the cognitive performance in decision tasks. Several other papers analyze the impact on risk preferences, but fail to reach a consensus (Ayton et al., 2020; Bogliacino et al., 2021; Cassar et al., 2017). When it comes to social preferences, there exist several contributions on trust (Bejarano et al., 2018, 2021; Castillo and Hernandez, 2023; Friehe and Marcus, 2021). The impact on the compliance to social norms has been overlooked, though.

We now move to present the theoretical predictions and the experimental evidence. Formal proofs, the experimental protocols, and additional materials are available in the Supplementary Material (SM).

2 Shocks and norm compliance

2.1 Intuition and the main prediction

According to norm-dependent utility (Bicchieri, 2006; Kimbrough and Vostroknutov, 2018; Krupka and Weber, 2013; Levitt and List, 2007), individuals have preferences that encompass a psychological cost incurred by deviating from a norm. In this section, we apply this framework to a set of tasks that can be studied in a laboratory setting.

Consider a DM facing a binary choice involving a social norm. Taking the action is harmful, resulting in a loss for the counterpart, or preventing them from benefiting. The social norm is a behavioral rule that may recommend either abstaining from the action or carrying it out, depending on the context. The DM's preferences depend on their income, including assets and monetary payoff from their choice, but they also have a conditional preference to follow the social norm. When violating a norm, the subject incurs a psychological cost. DMs are heterogeneous in their sensitivity to norms.

What happens when a DM suffers an NES? An NES represents a reduction in the DM's asset position, increasing their concern for money. They become more focused on money due to either of the following reasons. (a) They may be risk-averse (Camerer, 1995; Starmer, 2000): if risk aversion holds, a reduction in asset position increases the amount that the DMs are willing to accept in exchange for income. (b) They may be loss-averse (Kahneman and Tversky, 1979): they evaluate their choice from a reference point, where losses loom larger than gains. Under loss aversion, DMs become more concerned about recovering from the experienced shocks and are more likely to violate the norm.

Now consider well-known experimental games: stealing, cheating, JoD, and PD. Cheating and stealing are highly similar. In the stealing game, participants decide whether to 'take' other people's money, without retribution. By definition, the payoff from taking is larger than from abstaining. In the

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

Table 1. *Theoretical predictions from the norm compliance model. $P(\cdot)$ is the belief over the action of the opponent.*

Setting	Social norm	Prediction
Stealing	Do not steal	$S(NES) > S(C)$
Cheating	Do not cheat	$C(NES) > C(C)$
Joy of destruction	Retaliation	$D(NES) < D(C) \vert P(D) = 1$
PD	Tit-for-tat	$C(NES) < C(C) \vert P(C) = 1$

cheating task, they may lie for personal gain: their payoff is proportional to the reported (private) draw of a die.

The JoD involves a simultaneous interaction in which participants decide between burning the counterpart's endowment (incurring a cost) or abstaining. In the standard calibration (Abbink and Herrmann, 2011), the initial endowment is 10, the cost of burning is 1, and the damage inflicted is 5.

The PD represents a classical social dilemma: subjects decide whether to cooperate or not. Cooperation holds collective benefits, yet it is in the subject's private interest to defect. Formally, the PD is a symmetric simultaneous game where the payoff from defection is higher than the payoff from both cooperating, and the payoff from non-cooperation by both parties is higher than the payoff of cooperating alone.

We assume the following norms for each setting: 'thou shalt not steal,' 'thou shalt not lie,' retaliate, and tit-for-tat (conditional cooperation). The first 2 are self-explanatory. In the JoD, the social norm of *retaliate* accounts for the evidence of costly money burning (Abbink and Herrmann, 2011). Retaliate implies burning when expecting the counterpart to do the same. The norm of conditional cooperation has been widely documented in social dilemmas (Gachter, 2007).

These norms are assumed to hold while deriving the predictions. In the following section, we discuss how we validated them through a norm elicitation experiment.

Norm compliance comes with a cost in all cases. In stealing and cheating, the norm recommends abstaining but complying prevents the subject from enjoying a larger income. In JoD, the norm recommends retaliating: if the subject expects to suffer a loss from the counterpart, they will bear the cost and burn the endowment. In PD, if the counterpart is expected to cooperate, the norm advises cooperation, but this incurs a cost, as defection is more profitable. In all 4 situations, NES is expected to increase norm violation, leading to specific predictions for each scenario, summarized in Table 1.

Please note that predictions for JoD and PD are contingent upon the belief that the counterpart will execute the action. The trade-off exists solely under this belief, not with the opposite one: if the DM anticipates the counterpart's abstention in JoD or defection in PD, norms and incentives align to recommend the same action. A formal mathematical treatment of the problem and the proofs of the results that support the predictions can be found in the SM Section A.

2.2 Testing the predictions: methods

The assessment of the predictions requires 2 steps. First, since the predictions rely on the assumptions that norms x and y apply to each setting, data on normative expectations should support these assumptions. These data should be gathered in a separate experiment.

Second, the assessment requires the controlled manipulation of an NES.

For the first step, we incentivized participants to truthfully report normative expectations, using a standard protocol (Bicchieri and Xiao, 2009). The elicitation is conducted on a separate sample of subjects who did not participate in the main experiments but were drawn from the same subject pool. In the instructions, we clarified to the participants that they belonged to the same subject pool. Since norms

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

are contingent (Bicchieri and Xiao, 2009), specifying the reference group is necessary for identification. Section 3 presents the design of this experiment and the results.

For the second step, we conducted several experiments with the same structure. In the first part, participants performed an RET for money. After receiving feedback on their performance, subjects were randomly assigned to either a treatment condition, where they experienced an NES, or a control condition, where they maintained what they had previously earned. Following this phase, participants performed one or several tasks, including stealing, cheating, JoD, and PD. Sections 4–6 present the design and results of this set of experiments.

Finally, after assessing the predictions, we conducted a final test of the norm violation prediction in an experiment where the rule is imposed instead of being elicited. In this last experiment, we introduced additional treatments where we manipulated incentives ex ante (beforehand) instead of ex post (afterwards) with shocks, to identify the difference between wealth effects and shocks. Section 7 describes the procedures and results.

3 Eliciting social norms

The predictions are conditional on social norms. Following the definition by Bicchieri (2006), social norms should be supported by normative expectations. Normative expectations are second-order beliefs: what one expects others to think it ought to be done in a given contingency.

There are 2 main methods to elicit normative expectations: the coordination game by Krupka and Weber (2013) and the 2-step elicitation method by Bicchieri and Xiao (2009). The former asks participants to rate the actions available to the DM in terms of moral appropriateness but pays them to match the modal response. As in any coordination game, salience drives participants' strategic choices (Mehta et al., 1994). As a result, the incentives to coordinate should reveal the shared beliefs associated with norms.

The 2-step elicitation method by Bicchieri and Xiao (2009) recovers first-order and second-order beliefs. Subjects report their PNBs, as a singleton, over the action sets available to the DM. Then, they are paid to guess the response to the PNBs questions.

We send an online invitation to a sample drawn from the subject pool at the Unbiased Lab (Universidad Nacional) to fill in an online incentivized survey (it is available in SM Section C-IV). Data were gathered in February 2021.

Participants went through 2 parts. Part A elicited PNBs over the action space for the DM in each prosocial and antisocial task used in this article. The PNB is the 'personal opinion on what is the appropriate and morally correct action of Individual A, selecting one of the following options'. Each question included a description, the sample size, and the pool of participants. We described real experiments to make the description more accurate and to ensure that the pool of participants was salient. This was also instrumental in recovering the empirical expectations. We did not describe treatments but simply the decision problem. For cheating, stealing, and JoD, we used the experiments discussed in this manuscript. For PD, we did not have the data and to avoid deception we used the description of the same task from the experiment conducted by Bogliacino et al. (2023). The sequence of questions came in random order. In total, participants evaluated 6 decisions: 3 antisocial and 3 prosocial. The decisions include the trust and trustworthiness choices in a trust game, for which we elicited the normative expectations for another paper.

After stating their PNBs, in part B, participants were asked to predict the modal action among the respondents in the original experiment (empirical expectations) and the modal response to the PNBs questions among the respondents in the current experiment (normative expectations). Empirical expectations were collected for completeness. In each question, the order of the available options was randomized. Each participant made 12 predictions, and one was randomly selected for payment at the end. A correct guess was paid 25, 000 COP. The show-up fee was 10, 000 COP. The average time of completion was 35 minutes. We collected 109 observations. On average, participants earned 21, 000 COP (6 USD). Participants did not make decisions in the settings or participate in the experiments

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

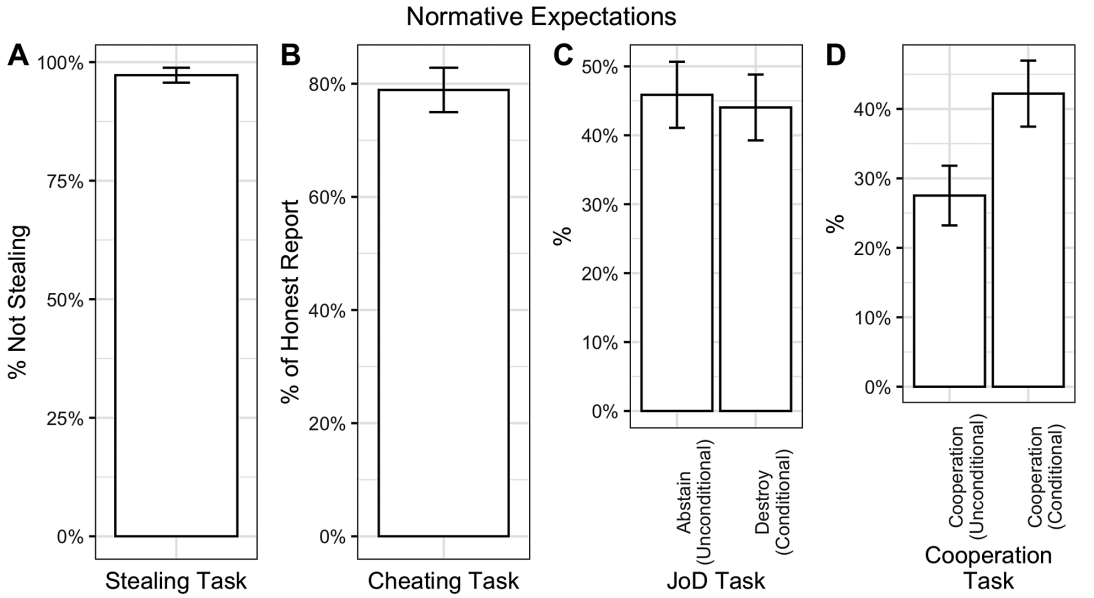

Figure 1. *The elicited normative expectations.*

described in the survey. This ensures the elicitation of normative expectations separately from behavior (Krupka and Weber, 2013).

These were the action sets of the DM in each setting described to the participants. For the stealing task, the decisions included stealing and not stealing. For the die-under-the-cup task, the action set included truthful reporting, reporting the first 3 numbers unconditionally, reporting 4 or 5 unconditionally, reporting 6 unconditionally, misreporting the drawn number plus or minus a maximum of 2 to own advantage, and misreporting the drawn number plus or minus a maximum of 2 to own disadvantage. For the JoD, the possible actions were burning unconditionally, abstaining unconditionally, choosing the same action as the counterpart (tit-for-tat), and choosing the opposite action of the counterpart. For the cooperation game, similarly to the JoD, the possible actions were cooperating unconditionally, defecting unconditionally, choosing the same action as the counterpart, and choosing the opposite action of the counterpart. We use the same framing used in the original experiment to avoid furthering experimenter demand.

We show the elicited normative expectations in Figure 1. For the stealing task (Panel A), Do Not Steal was the predicted PNB by 97.25% of the participants. In Panel B, truth-telling was predicted as the modal response to the PNB question for the cheating task by 78.90% of participants. In Panel C, for the case of JoD, the 2 modal normative expectations are non-burning unconditionally (45.87%) and tit-for-tat (44.04%). In Panel D, the modal normative expectation of cooperation is tit-for-tat (42.20%). To fit the page, Figure 1 plots the most frequent response to the normative expectations. In SM Section B, we report the tabulation of responses for each task. In Section 8, we discuss both normative and empirical expectations.

Once illustrated the social norms that apply to each setting, we move to the assessment of the predictions. We will now present 3 different experiments.

4 Experiment I

4.1 Experimental design and procedures

Experiment I is a standard between-subject design, with a treatment and a control condition. In the treatment condition, participants suffered an NES. The NES was an 80% loss on the accumulated

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

earning from an RET, experienced with a 50% probability. The probability was common information. The control includes those who did not experienced the shock. The RET was the Niederle and Vesterlund's (2007) task of summing sequences of 2-digit numbers and took place over 4 minutes. The assignment to the experimental conditions occurred at the individual level, within each session.

After the treatment, the participants played the stealing task and the JoD (Abbink and Herrmann, 2011) in random order. In the JoD, participants can burn half of the endowment of the counterpart at their own cost. The decision is simultaneous. The initial endowment is 10 ECUs and the cost of burning is 1 ECU. In the Stealing task, participants can appropriate 80% of the earnings from the RET, from another participant.

Three details of the experimental design are worth mentioning. First, to avoid a positive endowment shock after the NES, the assignment of the 10 ECU of the JoD preceded the RET. Second, we elicit beliefs on whether the counterpart was affected by shock and whether the counterpart was going to burn. To reduce the likelihood of hedging (Blanco et al., 2010), incentives for beliefs were smaller (1 ECU if correct).

Third, the victim in the Stealing task was a participant to another experiment occurring simultaneously. We could not allow stealing within the session, as this was instrumental to manipulate intentional shocks in another experiment, and as this would have induced possible retaliatory behavior changing the nature of the task. Additionally, 2 antisocial tasks with opponents within the same session could generate compensatory behavior.

This is how incentives were determined. Participants received the show-up fee and the gain from the RET immediately after the session's end. The RET should always be paid to maintain salience of the treatment. The money from the other tasks (and the beliefs) was paid 1 week later. This is the rationale for this methodological choice. Stealing affected a participant in another session, so logistically we had to call them back. To avoid asymmetry in the JoD and the Stealing incentives, and to make sure that we do not get attrition, we decided to pay both tasks with a delay.

The timeline was the following: all subjects played the RET, after which they were randomly assigned to control or treatment, and then they played the JoD and Stealing in random order. After reading the general instructions aloud, we asked participants to follow the specific instructions on the computer screen for each task. Subjects could raise their hand at any time if they had any questions. A final questionnaire was handed out to the participants.

In total, we recruited 184 undergraduate students from the Unbiased subject pool. Invitations were randomized. Sessions took place in the lab, in presence, around October 2019.

Out of the 184 participants, 92 were in the NES condition and 92 in the control. The average session had 20 participants and there were 9 sessions in total. The exchange rates were 1,000 COP per ECU. On average, each participant earned $17,000$ ($\pm 5,200$) COP (approximately USD 5 at the time). The experiment is programmed in oTree (D. L. Chen et al., 2016) and the English version of the protocol is available in SM Section C-I.

4.2 Results

Participants, on average, solved 5.21 ($\pm$2.44) problems, and the performance is not different across experimental conditions ($\chi^2 = 9.18, p = 0.759$).

In Figure 2 (left panel), we report the average stealing rate by condition, with a 95% confidence interval. On average, stealing increases from 62% to 73.9% in presence of an NES (Fisher's exact = 0.114 1-sided $p = 0.057$). We report a 1-sided test, since we are postulating a direction for the alternative hypothesis

To have an additional test of the prediction that stealing increases with NES, we run an OLS regression, adjusting for order. Since the experiment includes 2 antisocial tasks, there may be significant order effect, although the order is randomized and should not interfere with the treatment. Column 1 of Table 2 reports the results. The increase in stealing is both economically relevant, around 25% of an

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

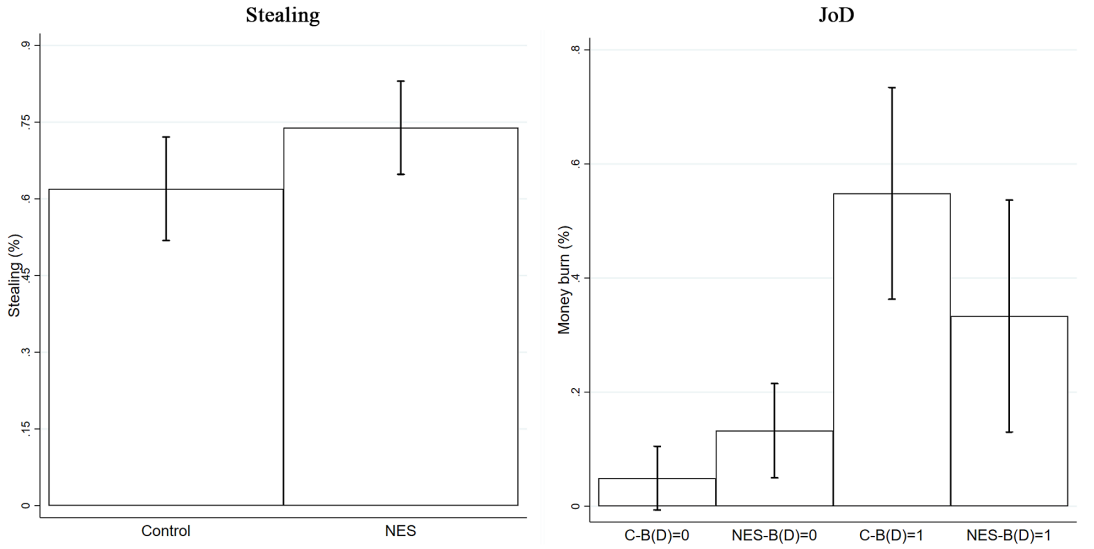

Figure 2. *The impact of NES in the stealing and JoD tasks.*

Table 2. *OLS estimates of effect of NES on stealing and JoD.*

	(1) Stealing	**(2) JoD**
NES	0.120	
	(0.068)	
Order	0.126	0.073
	(0.070)	(0.052)
NES-B(D)=0		0.080
		(0.049)
C-B(D)=1		0.486***
		(0.096)
NES-B(D)=1		0.277**
		(0.102)
Constant	0.545***	0.011
	(0.067)	(0.037)
N	184	184
R^2	0.03	0.21
F-test	3.00	8.95***

Note: Robust standard errors shown in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

sd of the outcome in the control condition, and statistically significant ($F(1, 181) = 3.07$, $p = 0.040$ 1-sided).

The behavior in the JoD is shown in Figure 2 (right panel). Participants in the control burn 4.91% of the time when they believe that the counterpart will not burn. This rate is not statistically different from 0. When they expect the other to burn, the rate increases to 54.38%. This experimental fact documents the social norm of retaliation plotted in Figure 1 and replicate the standard results in JoD experiments (Abbink and Herrmann, 2011). However, under the same belief, the shock lowers down the likelihood to retaliate to only 33.33%, i.e., there is a 21% reduction, which is both economically relevant

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

(52% of an sd computed for the control condition) and almost statistically significant ($F(1, 179) = 2.44$, $p = 0.059$ 1-sided).

The supporting regression is reported in column 2 of Table 2. We run an OLS regressions with 3 dummies: an indicator variable $NES - B(D) = 0$, an indicator variable $C - B(D) = 1$, and an indicator variable $NES - B(D) = 1$. The label $B(D)$ represents whether the belief on the counterpart was 0 or 1. The labels *NES* and *C* represent treatment and control. The omitted category is $C - B(D) = 0$. Again, distribution tests provide the same results: the rate of burning changes significantly across treatment and beliefs ($\chi^2(3) = 36.64, p < 0.001$) and the likelihood to burn, conditional on the belief $B(D) = 1$ is lower for NES than control (Fisher's exact = 0.172, 1-sided $p = 0.094$).[1]

To summarize, as predicted, we detected a positive effect of NES on stealing and a negative effect of NES on retaliation.

5 Experiment II

5.1 Experimental design and procedures

Experiment II is a between-subject design, with 2 conditions. The assignment to treatment is administered at the individual level. In the treatment, the participants face a loss of 80% of the accumulated earnings from an RET. As in Experiment I, we use the non-shocked as the control group.

This experiment had been conducted online, because of the COVID-19 pandemic. To ensure control, the RET was changed with respect to Experiment I. Logging in from home, some participants may access a calculator, voiding of meaning the RET. Instead, we adopted a 4-minute transcription task. In the transcription task, participants are required to copy fragments of 35 characters in dedicated boxes. The software did not allow copying and pasting. The text was written in Tagalog (the text was the Theory of Moral Sentiments; Smith, 1759), to ensure that performance depended on effort and not on accumulated knowledge.

After the RET and the assignment to the experimental conditions, the main task was a 'cheating game' based on Fischbacher and Föllmi-Heusi's (2013) die-under-the-cup. In this task, participants rolled a die privately and reported their results. In the instructions, an online die was made available through a link, highlighting that the link pointed outside the domain of the experiment and the observability from the experimenters, but the same instructions stressed that participants could use any available dice. The payoff was calculated as 2, 000 COP times the reported number (from 1 to 5) and 0 for a reported six. After the second task, participants had to answer some demographic questions.

This experiment was conducted online. We sent random invitations to a sample from the Unbiased subject pool, excluding those that took part in previous experiments with NES. We sent out a link for participation, with included instructions. Experiment II was conducted following the same timeline as Experiment I: RET, assignment to experimental condition, and main decision.

Participants received a show-up fee, the earnings from the RET (to make the treatment salient) and the die-under-the-cup. In total, we recruited 158 participants. Data collection occurred in June 2020; on average, each participant earned around 25, 000 COP (around 7 USD).

The experiments is programmed with oTree (D. L. Chen et al., 2016). The experimental protocol can be found in SM Section C-II.

5.2 Results

On average, participants tried 11.91 ± 3.78 transcriptions, completing successfully 9.17 ± 4.15 of them. There is no difference between treatment and control ($\chi^2 = 26.20$, $p = 0.199$ and $\chi^2 = 23.13$, $p = 0.231$, respectively).

[1]Since we pre-registered an OLS regression, we can confirm that the results do not change ($F(1, 295) = 1.32$, $p = 0.126$ 1-sided).

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

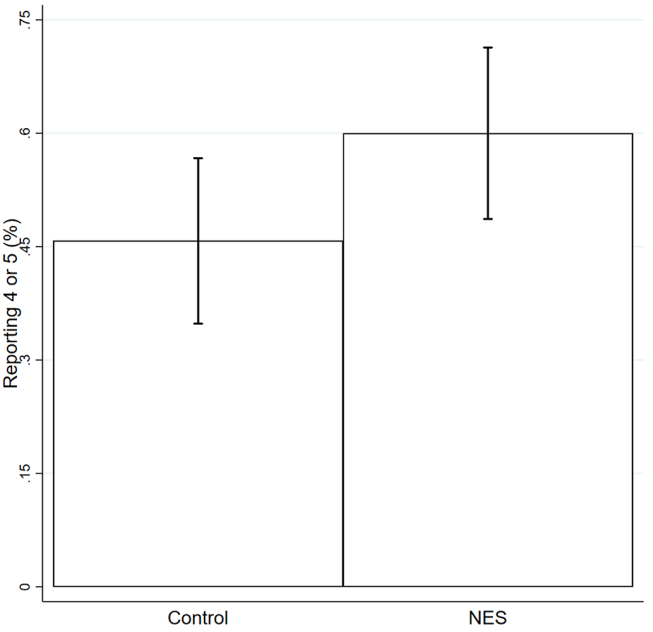

Figure 3. *The impact of NES in the die-under-the-cup task.*

There was a considerable amount of cheating: we neatly reject the null hypothesis that the observed data come from a fair dice ($\chi^2 = 37.11$, $p < 0.001$).

Since we do not observe the original draws, we cannot test for cheating directly. A conservative test assesses whether the payoff of the reported draw is larger for NES than control. Notice that the outcome is the payoff and not the draw because reporting larger numbers is cheating only for numbers lower than 5. A Mann–Whitney–Wilcoxon (MWW) test is almost significant ($z = -1.701$, $p = 0.089$).

For a more precise test, we can measure how the likelihood of reporting the numbers with the highest payoff (4 or 5) differs between treatment and control, as in Bogliacino and Montealegre (2020).

Figure 3 shows the mean outcome, broken down by experimental condition. In the control, the likelihood of reporting 4 or 5 is 45.78%. It increases to 60% in the presence of an NES. The difference is as large as 28% of an sd of the outcome in the control condition and is statistically significant ($t = -1.79$, $p = 0.037$ 1-sided, controlling for unequal variance).

6 Experiment III

6.1 Experimental design and procedures

In Experiment III, we test the effect of NES on cooperation. This is an online experiment with a 2×1 between-subject design. In the first part, participants performed a transcription task as in Experiment II. After knowing their performance, they either kept all the money (control) or suffered an 80% loss (NES). To avoid deception but allow for a surprise effect, in explaining the incentives for the RET, we warned that in the second part of the task, the total payment could change. The surprise effect is the first design change with respect to the previous 2 experiments.

Table 3 shows the normal form of the PD. In the instructions, strategies C and NC were labeled green and blue. The second design innovation with respect to Experiments I and II consisted in the introduction of this procedure: (a) participants were asked to declare their strategy based on the belief that the counterpart was playing either C or NC and (b) subsequently, they were required to indicate

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

Table 3. *Experiment III: the stage game.*

1/2	C	NC
C	8, 8	0, 10
NC	10, 0	4, 4

which scenario they considered more likely: the counterpart playing C, NC, or choosing randomly. Participants were aware that their answers in (a) would determine their behavior. This approach allowed us to elicit belief-contingent decisions, which represents the most rigorous test of our predictions and effectively addresses the endogeneity of belief.

Since we asked 4 comprehension questions with feedback, we excluded those that made more than one mistake. The total number of observations is 297, of which 146 are in the control and 151 in the NES condition. Data collection took place in March–April 2022 (PreAnalysis Plan recorded as AsPredicted #89448 reported in SM Section D-I).

We recruited participants via the REBEL Lab at Universidad del Rosario. Participants are invited using ORSEE (Greiner, 2015). The protocol is programmed in oTree (D. L. Chen et al., 2016) and can be found in SM Section C-III.

6.2 Results

On average, participants successfully completed 8.88 ± 4.02 transcriptions. There is no difference between treatment and control ($\chi^2 = 25.18$, $p = 0.120$).

These are the main results (Figure 4). Under the belief that the counterpart would not cooperate, participants chose cooperation 9.58% of the time. The shock marginally decreased cooperation to 9.27%. The difference was not statistically significant (Fisher's exact= 1.000, $p = 0.541$ 1-sided). Coming to the main outcome of interest, under the belief that the counterpart would cooperate, participants chose cooperation 59.58% of the time. The shock increased defection and lowered cooperation to 52.98%. The difference is not statistically significant (Fisher's exact= 0.293, $p = 0.151$ 1-sided) and represents slightly less than 15% of an sd in the control.

To wrap up, there is a marginal reduction in cooperation due to NES, but the effect is not statistically significant at the conventional level.

7 Experiment IV: evidence from rule following

Across 3 different experiments, NES induce more norm violations. The evidence is robust for antisocial tasks and less so for PD (Figure 4). To provide a robustness check, we designed an additional experiment. This experiment improves on the previous designs in 3 ways: (a) it separately controls for a wealth effect, (b) it imposes the rule to follow instead of relying on elicited social norm, and (c) it allows for an intensive margin to increase statistical power.

Participants faced an RET under 4 possible conditions: (1) each correct transcription is paid 1 point, the potential exposure to 80% loss is announced but not administered (Control); (2) each correct transcription is paid 1 point, the potential shock is announced and administered (Shock treatment); (3) each correct transcription is paid 1 point (High treatment) and there is no exposure to shock; (4) each correct transcription is paid 0.2 points (Low treatment) and there is no exposure to shock. Treatments (1) and (2) followed the procedures in Experiment III: the first part was described as having 2 phases, where the payment could change between phases 1 and 2. Treatments (3) and (4) have no second phase.

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

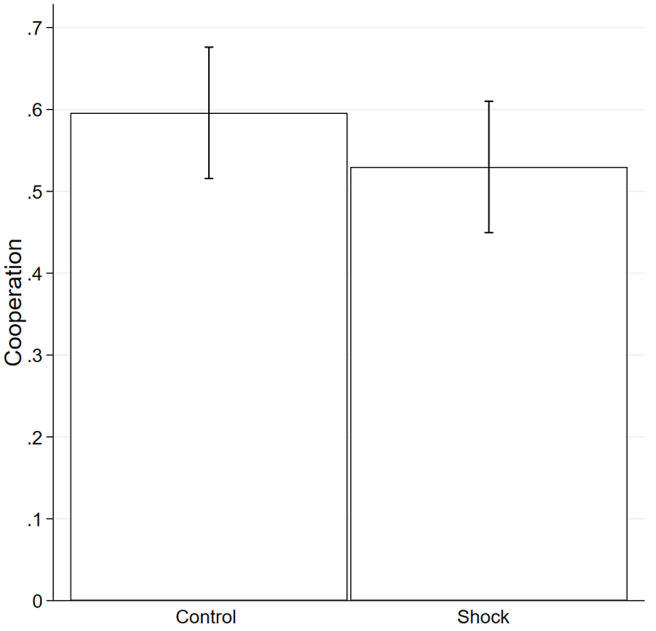

Figure 4. *The impact of NES in the prisoners' dilemma.*

In the second part, the main task is the Rule Following Task (Kimbrough and Vostroknutov, 2018). Participants should guide a stick man across a path with 5 traffic lights. Endowed with 30 seconds, each second worth 1 point, they were paid for the seconds left when they reached the end of the path. The rule announced to participants but not enforced was to wait for the green at each traffic light. It took 5 seconds for a traffic light to change from red to green. The protocol was programmed in oTree (D. L. Chen et al., 2016) and can be found in SM Section C-V.

This is a between-subject online experiment. Random invitations were sent to those participants in our database that did not participate in any other experiment with shocks. The analysis was pre-registered (Aspredicted #80601, available at SM Section D-II). We paid both tasks and a show-up fee. Notice that we expected to reach 300 participants, but due to the end of the term and the fact that we could not use participants from other experiments, we had to close the data collection at 210 participants. However, the variability was lower than expected.

These are the results. The performance in the RET was marginally different across treatments ($\chi^2 = 74.81$, $p = 0.094$). The main outcome variable was the number of seconds spent crossing the path. On average, subjects spent 5.92 seconds (4.80 sd). In the control, the outcome was 7.51 (±6.80). In treatment (2) (the NES condition in Experiment I-III), the outcome was 4.95 (±2.73). The difference is almost statistically significant in an MWW test ($z = -1.929$, $p = 0.053$), but the test is 2-sided. The effect is around 38% of the sd in the control.

In treatment (3), the outcome was 5.89 (±4.34), whereas in treatment (4), it was 5.43 (±4.28). An MWW test does not reject the null hypothesis ($z = -0.809$, $p = 0.418$). In other words, the wealth effect did not generate the same result as the shock. Results are plotted in Figure 5.

Since we pre-registered also OLS regression using the seconds as the outcome variable, with robust standard errors, we report them in SM Section B. A (1-sided) *t*-test of the effect of NES, controlling for unequal variance, returns $t = 2.46$ (1-sided $p = 0.008$). The difference between the shock effect (dummy for treatment 2) and the wealth effect (the difference in performance between treatments (3) and (4)) is 2.94 seconds and is statistically significant ($t = -2.27$, $p = 0.024$).

We can finally summarize all our experimental evidence in Table 4.

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

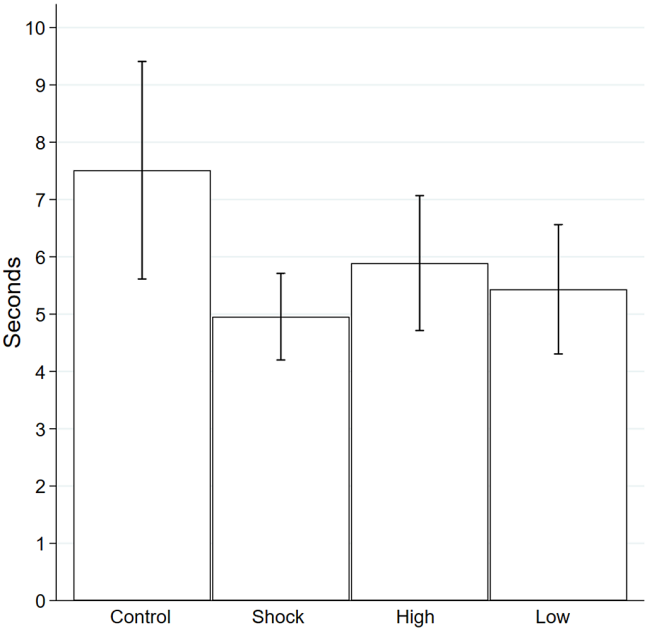

Figure 5. *Results from the rule following task.*

Table 4. *Summary of predictions and results. $P(\cdot)$ is the belief over the action of the opponent.*

Setting	**Prediction**	**Effect (% sd in C)**
Stealing	$S(NES) > S(C)$	24.48
Cheating	$C(NES) > C(C)$	28.37
Joy of destruction	$D(NES) < D(C) \vert P(D) = 1$	51.86
PD	$C(NES) < C(C) \vert P(C) = 1$	13.42
Rule following	$T(NES) < T(C)$	37.53

8 Discussion

In this section, we discuss norm dependent utility theory in relation to our experimental results and evidence from previous studies, both experimental and quasi-experimental. We then provide a critical discussion of some of the assumptions. Finally, we then contrast this explanation with alternative theories, particularly with regard to endogenous self-image theory.

8.1 Other evidence from quasi-experimental studies

Table 4 summarizes the main results. Our evidence is consistent with several stylized facts. The observation that NES may generate antisocial behavior, particularly crimes against property, has not gone unnoticed. Compelling quasi-experimental evidence documents a positive relationship between NESs and antisocial behavior. For example, Dube and Vargas (2013) used the change in coffee prices to study variations in crime in communities highly dependent on the harvest for their income. Cortés et al. (2016) used the collapse of the Ponzi scheme in Colombia to detect variations in a portfolio of criminal activities. Moreover, Bignon et al. (2017) exploited regional variation in the exposure to

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

phylloxera in wine-producing regions in France to identify an increase in property crime. Additionally, Dix-Carneiro et al. (2018) used the trade liberalization shock in Brazil to estimate the causal impact of the shock on criminal activity. Even weather shocks have been linked to an increase in property crimes (Mehlum et al., 2006).

For these stylized facts, norm-dependent utility outperformed other theories. In the JoD, the norm of retaliation generates a trade-off between compliance and income (money burning is costly). Although grounded in the same reasoning as stealing and cheating, the model predicts *less* antisocial activity and is consistent with our controlled evidence. Contrarily, theories of crime like strain theory (Merton, 1938) cannot predict this finding. Strain theory states that the frustration caused by NES should increase all antisocial behaviors (stealing *and* money burning). Additionally, psychological theories of NES cannot fully account for the experimental evidence: NES induce cognitive load (Bogliacino and Montealegre, 2020; Mani et al., 2013), but there is no consensus on the relationship between cognitive load and social preferences (Alós-Ferrer and Garagnani, 2020; Kvarven et al., 2020; Rand, 2016). Haushofer and Fehr (2014) claim that people experiencing shocks may face increasing risk aversion due to stress. Increasing risk aversion induces more norm violation, ceteris paribus. The evidence on the relationship between shocks and risk aversion is inconclusive, though (Bogliacino et al., 2021; Chuang and Schechter, 2015).

8.2 Other evidence from experimental studies

Our paper makes a contribution to the set of studies that measure behavior using induced value in the lab or in lab-in-the-field and manipulate shocks or use plausible exogenous variations.

In our opinion, the strongest piece of evidence comes from Boonmanunt et al. (2020), which combines pre- and post-harvest variation with a social norm manipulation. In line with our predictions, a social-norm-reminder intervention reduced cheating after harvest, but not before. This result points to the fact that people who are more likely to think in terms of trade-offs because of shocks (de Bruijn and Antonides, 2022) are less responsive to a social norm intervention, which supports our main argument.

A couple of previous studies were inconclusive: Aksoy and Palma (2019) measured cheating around paycheck variation, but could not detect any significant variation. Bogliacino and Montealegre (2020) exogenously manipulated NES and measure cheating with die-under-the-cup task, finding no effect, but the presence of 4 tasks may have diluted the incentives.

As we discussed in the introduction, a well-established literature gathered consistent evidence that exposure to collective threats such as natural disasters, violence, or prolonged scarcity makes prosocial behavior more likely to occur (Bauer et al., 2016; Bogliacino et al., 2023; Cassar et al., 2017) or money burning (Prediger et al., 2014), but this literature differs in that these shocks are framed and involve variation of other elements, such as social identities.

We delve into discussing 2 experimental studies in more depth. Bejarano et al. (2021) manipulate NES on the endowment in a trust game in a couple of papers: in one, they induce a loss on the first mover, and in the second, on the second mover. Their main methodological choice is to induce shock on unearned endowment (differently from us), to make it common knowledge, and to separately manipulate wealth (as in our Experiment discussed in Section 7). Part of the difference may reflect alternative design choices (in particular, the earned vs. unearned endowment), but the choice of the task also plays a role.

The trust decision is fundamentally different from the one at the core of our model: since the norm for the trust game is likely to be tit-for-tat, norms and incentives recommend the same action, and thus the shock should produce no effect. Things are different for trustworthiness, where the most probable norm is *to share*, which generates a clear trade-off between incentives and norms. As a result, a way to test our theory is to analyze the trustworthiness decisions when negative shocks are administered on the second mover (Bejarano et al., 2018). A comparison between their Shock-Equal (equivalent to our Control condition) and Shock-Unequal (equivalent to our NES condition) treatments shows that the amount returned goes down by 17 pp (from 0.39 to 0.19) and the share who returned nothing (which would be the strictest form of non-compliance) goes up by 19 pp (from 0.10 to 0.29).

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

Finally, Pickup et al. (2020) look at the consequences of taking norm-consistent decisions when it is revealed ex post that this involved a cost, instead of a benefit: their design remunerates norm-inconsistent choices with a fixed benefit and norm-consistent choices with a mixed lottery. Those participants who faced the loss instead of the gain (in the lottery) adjusted their beliefs and their punishment/reward behavior to strengthen the importance of the norm. The paper is not meant to test our main hypothesis but rather the existence of cognitive dissonance when people face regrets for adhering to established norms. With this in mind, it still manipulates shocks and elicits behavior in settings with substantive norms.

We agree that this result appears contradictory with respect to our prediction. In fact, the authors report a random shock leading to a subsequent increase in reward and punishment associated with the norm. There are 2 confounding factors that should be considered when reading the results from the point of view of our theoretical framework. First, the loss is manipulated via the lottery but is determined by the initial decision to comply. This methodological choice (correct from the point of view of the authors' research question) leverages regret in a way that is very different from our settings. Second, we suspect that framing the loss as norm-related should make people more concerned with norms, which generates exactly the opposite effect of losses in our settings. Of course, one could not capture this effect within our model (as θ is a parameter that does not vary within subject), but we can easily think of other intervening behavioral factors (emotions, beliefs updating, etc.).

8.3 Plausibility of the assumptions

In order to investigate whether the norm violation hypothesis looks plausible, over the next paragraphs, we discuss some of the assumptions.

8.3.1 Conditional preference to conform

Our model is predicated upon the assumption that people are motivated by norms, besides self-interest. We model the conditional preference to conform with the assumption that the cost of violating a norm enters the utility function. Contrary to this assumption, humans may feel also positive emotions or self-esteem from *following* norms. In other words, we are equating the avoidance of a cost to receiving a benefit. Behaviorally, this may sound like a restrictive assumption, but it has no bearing on our results.

In fact, one could think at Krupka and Weber's (2013) norm-dependent utility where subjects are motivated by material payoffs (where either concavity or loss aversion hold) and by the perceived appropriateness of the options available in the action set. This latter term is $\theta N(a)$ where θ is the sensibility to the norm and $N(\cdot) \in [-1, +1]$ are the normative expectations, elicited with the coordination method. The main prediction holds (as formally shown in SM Section A.4). The intuition is simply that the choice still involves a comparison between material interests and norm compliance concerns.

8.3.2 Nature of the reference point

As discussed in Section 2, the main prediction rests on risk aversion or loss aversion, 2 behavioral traits that are common in the average subject, even in studies with representative samples (Chapman et al., 2023).

Based on the results of experiment IV (described in Section 7), we believe that between the 2, the most plausible assumption is the presence of reference dependence and loss aversion. The definition of NES that we believe applies to this theory is a loss that is significant and salient. Our experimental approach leverages the sense of entitlement generated by real effort, which makes the endowment a natural candidate for a reference point.

Our theory is predicated on a rather traditional distinction between material interest and other 'motivations' (typically belief dependent) that enter as arguments in the utility functions (Battigalli and Dufwenberg, 2022). As a result, the reference point is described as a vector of goods or a monetary

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

payoff. However, a more complex representation is arguably possible. The status quo for a person losing a job is a case in point: it can be construed as a monetary loss in earnings, but also include the loss of the sense of stability and achievements or the loss of status. The theory is grounded on the existence of a trade-off, and there are multiple ways to conceptualize it.

8.3.3 Nature of the shock

In this manuscript, NESs are manipulated at the individual level and do not involve framing. This assumption is crucial to make sure that we manipulate concern for money keeping everything else equal. The literature on disasters, violence, and tight and loose cultures tends to see shocks as group-level threats. This is common to field studies to lab studies that manipulate shocks (Szekely et al., 2021).

In principle, the distinction between group-level shock and individual-level shock is likely to matter, because the 2 leverage social identities to a different degree. Collective threats deepen the in-group dimension. This may lead to differential behavior toward group members in terms of trust, cooperation, and punishment (Chen and Li, 2009).

Any additional framing of the shock may potentially lead to further behavioral consequences. Natural disasters may evoke memories of other people stepping up to help; violence may generate a sense of injustice or a violation of the self. Understanding these elements of a taxonomy of shocks requires further theoretical and empirical analysis.

8.3.4 Conflicting norms

In the JoD game, subjects are split and half of them believe that unconditionally abstaining from the destruction of payoffs is the norm, whereas in the PD game, about 30% of the subjects consider unconditional cooperation to be the social norm. Although it is true that our model does not include heterogeneity in terms of the norm, this is irrelevant to our results.

Notice that in both cases—unconditional abstention and unconditional cooperation—the strategy recommended by norms and incentives remains the same regardless of the shock. By random assignment, those subjects that have these normative expectations will be equally split between treatment and control, and cannot confound our inference.

8.3.5 Normative and empirical expectations

A related question is whether normative and empirical expectations are aligned. Given that norms are dependent on both beliefs (Bicchieri, 2006), one could gain additional insights from exploring those data. Unfortunately, since the elicitation of the empirical expectations was not the primary focus, we gathered a rather coarse measure, asking participants to predict the modal response (and never conditioning on beliefs). As a result, we can only speculate whether normative and empirical beliefs are consistently supporting the same norm. We plotted the empirical expectations in Figure 2 of SM Section B.

In the Stealing task, 30.28% of subjects expect stealing to be the modal decision, this should support the social norm of abstaining. In the JoD, 77.98% of participants predict the mode to be 'Destroy'. This is consistent with the norm of retaliation, in the sense that behavior and empirical expectations are self-sustaining. In the PD, the beliefs over the modal action are split, 46.79% of participants predict cooperation. We cannot state whether this supports the claim that conditional cooperation is the norm, but of course it is perfectly consistent with it (especially considering that the interaction is one shot). A review of existing studies on social dilemmas shows that conditional cooperators usually constitute from half to 60% of subjects (Thöni and Volk, 2018), but the review is based on the dynamic Voluntary Contribution Game, whereas here we are considering a dichotomous and one-shot PD.

The least informative case is the cheating task. In the latter, subjects expect the modal decision to be reporting 4 (32.11%) or 5 (44.04%)—obviously, we could not elicit beliefs of cheating because we did not observe the outcome of the dice. This implies that subjects expect cheating to occur, but we cannot distinguish from these data whether they expect a moderate or a large amount of cheating (both are consistent with subjects predicting a mode of 4 or 5).

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

8.4 Alternative interpretations

We showed that our model of norm compliance explains a number of results in both our experiments and previous studies. A natural question is whether the results can be explained by other theories.

In particular, a major critique concerns the possibility that instead of defining ex ante the social norm that applies to the setting and study compliance as a result of the shock, we model social expectations as endogenous variables that are updated with the shock.

On the one hand, we do defend the plausibility of our approach. In settings in which social norms are meaningful explanatory variables, they have to show some persistence. Bicchieri (2006) uses the metaphor of the script, to grasp the role of norms as a guidance and a blueprint for action. This makes sense only to the extent that there are shared beliefs supported by enough confidence. Of course, we are not denying that unrest and revolutions may drastically reshape norms, but this does not apply to individual-level shocks that did not target groups or threaten social identities. Additionally, on a more practical level, we are studying very specific contingencies, for which norms are well rooted: in 2 of them (cheating and stealing), social norms are reinforced by moral norms (Sugden, 2018), whereas for the other 2 (retaliation and cooperation), norms may be supported by evolutionary mechanisms such as emotions that curb self-interest or enforce commitment (Darwin, 1871; Frank, 1988; Smith, 1759).

What we do not deny, is the presence of different psychological justifications, that a DM may use to convince the internal observer or the members of the reference groups, trying to conceal the violation of the norm. This leads to the second part of our argument.

Instead of social norms, we may think of preferences defined over material consequences and the social image (Andreoni and Bernheim, 2009; Benabou and Tirole, 2006; Grossman, 2015; Grossman and van der Weele, 2017). The social image is a second-order belief over the (unobservable) behavioral traits owned by the DM, conditional on the outcome observed. In equilibrium, these beliefs are updated consistently with the action taken. In the SM, we present a simple model where preferences have these characteristics and shocks are not fully observable (which is what happens in our experiments). Based on this analysis, we claim that, in the population, there will be 3 classes of people (we simplify the narrative here but formal results are in SM Section A.6): those that deviate from the norms regardless of the shock (the bad), those who comply because they did not receive the shock (the good), and those that would have complied had the shock spare them (the unlucky).

In this context, beliefs are fully endogenous and the social motivation of subjects is not additively separable from extrinsic motivation (Bowles and Polania-Reyes, 2012), but the mechanism and the interpretation are essentially the same: a subset of DMs trade-off social image and concern for material payoff. In equilibrium, these subjects have the psychological justification that, although they pool with those who do not care about the norm, there is enough chance to be seen as an unlucky victim instead. When they are spared by the loss, the concern for social image is still too large and the DM chooses to comply, but when loss aversion kicks in, it leads to norm violation.

Finally, it is easy to suggest that other explanations may account for some of the results. For example, a trade-off between loss aversion and anger may explain the effect of NES on retaliation. Relying on these partial motivations would typically explain some but not all of the results, as we suggested in Section 8.1 about strain theory. We believe that there is some added value in providing a simple explanation to an array of stylized facts.

9 Concluding remarks

What happens to norm compliance when subjects face an NES? We explore this question experimentally. According to norm-dependent utility theory, subjects are expected to deviate more from social norms. To assess this prediction, we randomly administer shocks on the earnings from an RET and elicit the behavior of participants in several tasks where a rule of behavior is grounded on normative expectations. The experiments demonstrate how experiencing major losses can be a source of norm violation.

https://doi.org/10.1017/jdm.2024.1 Published online by Cambridge University Press

This article bears implications for the ongoing discussion on global challenges such as pandemics, global warming, and war threats. Theories of cultural evolution suggest that the Western culture evolved through a peculiar WEIRD—an acronym for Western, Educated and from an Industrialized, Rich, and Developed country—psychology and set of social norms, forged in response to the Marriage and Family Program of the Church (Heinrich, 2020; Heinrich et al., 2010). If vulnerability to shocks can drive the abandonment of norms and the establishment of new rules of behavior, the heterogeneity of such vulnerability has important implications for the evolution of cultural norms.

Finally, this article has implications for the discussion around welfare state reforms. Markets *discipline*, but downward adjustment of price and earnings may lead to unintended consequences (Dix-Carneiro et al., 2018). Recent evidence from the United Kingdom confirmed that this might be the case (d'Este and Harvey, 2022).

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/jdm.2024.1.

Data availability statement. This project includes an OSF page with the DoFile to generate the analysis, all data, and the Supplementary Materials (https://osf.io/rwxpt/).

Acknowledgments. The authors acknowledge helpful comments, constructive critiques, and suggestions from the Editor and two anonymous referees. A special thanks to Roberto Galbiati, Zachary Grossman, Marcela Ibanez, Erin Krupka, Cesar Mantilla, Pietro Ortoleva, Daniel Parra, Leonardo Pejsachowicz, Gerhard Riener, Nuria Tolsa-Caballero, and Eugenio Verrina for comments and suggestions. Douglas Rodríguez and Sebastian Ramírez helped during data collection. We are very grateful to the participants at various seminars and conferences: BEBES, Universidad de Los Andes, Universidad del Valle, Cournot Seminar at Strasbourg, AMSE Seminar in Aix-en-Provence, Development Seminar in Goettingen, Warsaw Economic Seminar, GREDEG Seminar in Sophia Antipolis, Economics and Psychology Seminar at Paris I-PSE, ESA Conference 2021, LACEA Brain 2022, BEEC 2022, and NOBEC 2022.

Author contributions. Conceptualization, data curation, and methodology: all authors. Data visualization: F.B., R.C. Writing—original draft: F.B. All authors approved the final submitted draft.

Funding statement. This research was supported by Open Evidence via contracts 18810 and 29047 with Universidad Nacional de Colombia.

Competing interests. The authors declare no competing interests.

Ethical standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country. Ethical approval for the experiments of this projects has been issued by the Comité de Ética de la Universidad Nacional-Facultad de Medicina (Acta 012-159-19, Acta 016-163, and Acta 020-189).

References

Abbink, K., & Herrmann, B. (2011). The moral costs of nastiness. *Economic Inquiry*, *49*(2), 631–633. https://doi.org/10.1111/j.1465-7295.2010.00309.x.

Aksoy, B., & Palma, M. A. (2019). The effects of scarcity on cheating and in-group favoritism. *Journal of Economic Behavior and Organization*, *165*, 100–117. https://doi.org/10.1016/j.jebo.2019.06.024.

Alós-Ferrer, C., & Garagnani, M. (2020). The cognitive foundations of cooperation. *Journal of Economic Behavior and Organization*, *175*, 71–85. https://doi.org/10.1016/j.jebo.2020.04.019.

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607–1636. https://doi.org/10.3982/ECTA7384.

Ayton, P., Bernile, G., Bucciol, A., & Zarri, L. (2020). The impact of life experiences on risk taking. *Journal of Economic Psychology*, *79*(April), 1–4. https://doi.org/10.1016/j.joep.2020.102274.

Battigalli, P., & Dufwenberg, M. (2022). Belief-dependent motivations and psychological game theory. *Journal of Economic Literature*, *60*, 833–882. https://doi.org/10.1257/jel.20201378.

Bauer, M., Blattman, C., Chytilová, J., Henrich, J., Miguel, E., & Mitts, T. (2016). Can war foster cooperation? *Journal of Economic Perspectives*, *30*(3), 249–274. https://doi.org/10.1257/jep.30.3.249.

Bejarano, H., Gillet, J., & Rodriguez-Lara, I. (2018). Do negative random shocks affect trust and trustworthiness? *Southern Economic Journal*, *85*(2), 563–579.

Bejarano, H., Gillet, J., & Rodriguez-Lara, I. (2021). Trust and trustworthiness after negative random shocks. *Journal of Economic Psychology*, *86*, 102422. https://doi.org/10.1016/J.JOEP.2021.102422.

Benabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, *96*(5), 1652–1678.

Benczes, I., & Szabó, K. (2023). An economic understanding of populism: A conceptual framework of the demand and the supply side of populism. *Political Studies Review*, *21*(4), 680–696. https://doi.org/10.1177/14789299221109449.
Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.
Bicchieri, C. (2010). Norms, preferences, and conditional behavior. *Politics,Philosophy & Economics*, *9*(3), 297–313. https://doi.org/10.1177/1470594X10369276.
Bicchieri, C., Dimant, E., & Xiao, E. (2021). Deviant or wrong? The effects of norm information on the efficacy of punishment. *Journal of Economic Behavior & Organization*, *188*, 209–235.
Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making*, *22*(2), 191–208. https://doi.org/10.1002/bdm.
Bignon, V., Caroli, E., & Galbiati, R. (2017). Stealing to survive? Crime and income shocks in nineteenth-century France. *Economic Journal*, *127*(599), 19–49. https://doi.org/10.1111/ecoj.12270.
Blanco, M., Engelmann, D., Koch, A. K., & Normann, H. T. (2010). Belief elicitation in experiments: Is there a hedging problem? *Experimental Economics*, *13*(4), 412–438. https://doi.org/10.1007/s10683-010-9249-1.
Bogliacino, F., Codagnone, C., Montealegre, F., Folkvord, F., Gómez, C., Charris, R., Liva, G., Lupiáñez-Villanueva, F., & Veltri, G. A. (2021). Negative shocks predict change in cognitive function and preferences: Assessing the negative affect and stress hypothesis. *Scientific Reports*, *11*(1), 1–10. https://doi.org/10.1038/s41598-021-83089-0.
Bogliacino, F., Gómez, C., & Grimalda, G. (2023). Crime-related exposure to violence and prosocial behavior: Experimental evidence from Colombia. *Journal of Behavioral and Experimental Economics*, *104*, 102023. https://doi.org/10.1016/j.socec.2023.102023.
Bogliacino, F., & Montealegre, F. (2020). Do negative economic shocks affect cognitive function, adherence to social norms and loss aversion? *Journal of the Economic Science Association*, *6*(1), 57–67. https://doi.org/10.1007/s40881-020-00091-4.
Boonmanunt, S., Kajackaite, A., & Meier, S. (2020). Does poverty negate the impact of social norms on cheating? *Games and Economic Behavior*, *124*, 569–578. https://doi.org/10.1016/j.geb.2020.09.009.
Botchway, E., & Filippin, A. (2021). Cooperation as adaptation to persistent risk of natural disasters: Evidence from a natural experiment.
Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, *50*(2), 368–425.
Camerer, C. (1995). Individual decision making. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 587–704). Princeton, NJ: Princeton University Press. https://doi.org/10.2307/j.ctvzsmff5.12.
Carvalho, L., Meier, S., & Wand, S. (2016). Poverty and economic decision-making: Evidence from changes in financial resources at payday. *American Economic Review*, *42*(2), 407–420. https://doi.org/10.1002/jmri.24785.Free-Breathing.
Cassar, A., Healy, A., & von Kessler, C. (2017). Trust, risk, and time preferences after a natural disaster: Experimental evidence from Thailand. *World Development*, *94*, 90–105. https://doi.org/10.1016/j.worlddev.2016.12.042.
Castillo, J. G., & Hernandez, M. A. (2023). The unintended consequences of confinement: Evidence from the rural area in Guatemala. *Journal of Economic Psychology*, *95*(May 2022), 102587. https://doi.org/10.1016/j.joep.2022.102587.
Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., & Camerer, C.(2023). Econographics. *Journal of Political Economy Microeconomics*, *1*(1), 115–161. https://doi.org/10.1086/723044.
Chavez, A. K., & Bicchieri, C. (2013). Third-party sanctioning and compensation behavior: Findings from the ultimatum game. *Journal of Economic Psychology*, *39*, 268–277. https://doi.org/10.1016/j.joep.2013.09.004.
Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree – An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001.
Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431–457. https://doi.org/10.1257/aer.99.1.431.
Chuang, Y., & Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics*, *117*, 151–170. https://doi.org/10.1016/j.jdeveco.2015.07.008.
Cortés, D., Santamaría, J., & Vargas, J. F. (2016). Economic shocks and crime: Evidence from the crash of Ponzi schemes. *Journal of Economic Behavior and Organization*, *131*, 263–275. https://doi.org/10.1016/j.jebo.2016.07.024.
D'Acunto, F., Malmendier, U., Ospina, J., & Weber, M. (2021). Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, *129*(5), 1615–1639.
d'Este, R., & Harvey, A. (2022). The Unintended Consequences of Welfare Reforms: Universal Credit, Financial Insecurity, and Crime. *The Journal of Law, Economics, and Organization*, ewac009, https://doi.org/10.1093/jleo/ewac009.
Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z.
Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: Murray. https://doi.org/10.1038/011305a0.
de Bruijn, E. J., & Antonides, G. (2022). Poverty and economic decision making: A review of scarcity theory. *Theory and Decision*, *92*, 5–37. https://doi.org/10.1007/s11238-021-09802-7.
Dix-Carneiro, R., Soares, R. R., & Ulyssea, G. (2018). Economic shocks and crime: Evidence from the Brazilian Trade Liberalization. *American Economic Journal: Applied Economics*, *10*(4), 158–195. https://doi.org/10.1257/app.20170080.
Dube, O., & Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *Review of Economic Studies*, *80*(4), 1384–1421. https://doi.org/10.1093/restud/rdt009.

Eckel, C. C., Fatas, E., & Kass, M. (2022). Sacrifice: An experiment on the political economy of extreme intergroup punishment. *Journal of Economic Psychology*, *90*(January), 102486. https://doi.org/10.1016/j.joep.2022.102486.
Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise – An experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547. https://doi.org/10.1111/jeea.12014.
Frank, R. H. (1988 *). Passions within reason: The strategic role of emotions*. New York: Norton.
Friehe, T., & Marcus, J. (2021). Lost job, lost trust? On the effect of involuntary job loss on trust. *Journal of Economic Psychology*, *84*(February), 102369. https://doi.org/10.1016/j.joep.2021.102369.
Gachter, S. (2007). Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In B. Frey & A. Stutzer (Eds.), *Economics and psychology: A promising new cross-disciplinary field* (pp. 19–50). Cambridge, MA: MIT Press.
Gelfand, M. J., Harrington, J. R., & Jackson, J. C. J. C. (2017). The strength of social norms across human groups. *Perspectives on Psychological Science*, *12*(5), 800–809. https://doi.org/10.1177/1745691617708631.
Gintis, H. (2007). A framework for the unification of the behavioral sciences. *Behavioral and Brain Sciences*, *30*(1), 1–16. https://doi.org/10.1017/S0140525X07000581.
Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.
Grossman, Z. (2015). Self-signaling and social-signaling in giving. *Journal of Economic Behavior and Organization*, *117*, 26–39. https://doi.org/10.1016/j.jebo.2015.05.008.
Grossman, Z., & van der Weele, J. J. (2017). Self-image and willful ignorance in social decisions. *Journal of the European Economic Association*, *15*, 173–217. https://doi.org/10.1093/jeea/jvw001.
Guiso, L., Herrera, H., Morelli, M., & Sonno, T. (2017). *Demand and supply of populism* [EIEF Working Papers Series 1703]. Einaudi Institute for Economics and Finance.
Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *Science*, *344*(6186), 862–867. https://doi.org/10.1126/science.1232491.
Heinrich, J. (2020). *The WEIRDest people in the World*. New York: Farra, Straus and Giroux.
Heinrich, J., Heine, S. J., & Norenzayan, A. (2010). Weird people. *Behavioral and Brain Sciences*, *33*(2–3), 61–135.
Inglehart, R., & Norris, P. (2016). *Trump, Brexit, and the rise of populism: Economic have-nots and cultural backlash* [HKS Working Paper No. RWP16-026]. Harvard Kennedy School.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 293–298. https://doi.org/10.2307/j.ctv1kr4n03.21.
Kimbrough, E. O., & Vostroknutov, A. (2018). A portable method of eliciting respect for social norms. *Economics Letters*, *168*, 147–150. https://doi.org/10.1016/J.ECONLET.2018.04.030.
Krupka, E. L., & 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. https://doi.org/10.1111/jeea.12006.
Kvarven, A., Strømland, E., Wollbrant, C., Andersson, D., Johannesson, M., Tinghög, G., Västfjäll, D., & Myrseth, K. O. R. (2020). The intuitive cooperation hypothesis revisited: A meta-analytic examination of effect size and between-study heterogeneity. *Journal of the Economic Science Association*, *6*(1), 26–42. https://doi.org/10.1007/s40881-020-00084-3.
Levitt, S. J., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*(2), 153–174.
List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, *115*(3), 482–493.
List, J. A., Berrens, R. P., Bohara, A. K., & Kerkvliet, J. (2004). Examining the role of social isolation on stated preferences. *American Economic Review*, *94*(3), 741–752. https://doi.org/10.1257/0002828041464614.
Malmendier, U. (2021a). Experience effects in finance: Foundations, applications, and future directions. *Review of Finance*, *25*(5), 1339–1363. https://doi.org/10.1093/rof/rfab020.
Malmendier, U. (2021b). Why personal histories matter in economics. *Journal of the European Economic Association*, *19*(6), 2857–2894.
Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science*, *341*(6149), 976–980. https://doi.org/10.1126/science.1238041.
Mehlum, H., Miguel, E., & Torvik, R. (2006). Poverty and crime in 19th century Germany. *Journal of Urban Economics*, *59*(3), 370–388. https://doi.org/10.1016/j.jue.2005.09.007.
Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *American Economic Review*, *84*(3), 658–673.
Merton, R. K. (1938). Social structure and anomie. *American Sociological Review*, *3*(5), 672–682. https://doi.org/10.2307/2084686.
Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101, https://doi.org/10.1162/qjec.122.3.1067
Pickup, M. A., Kimbrough, E. O., & de Rooij, E. A. (2020). Identity and the self-reinforcing effects of norm compliance. *Southern Economic Journal*, *86*, 1222–1240. https://doi.org/10.1002/soej.12410.
Prediger, S., Vollan, B., & Herrmann, B. (2014). Resource scarcity and antisocial behavior. *Journal of Public Economics*, *119*, 1–9. https://doi.org/10.1016/j.jpubeco.2014.07.007.

Rand, D. G. (2016). Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science*, *27*(9), 1192–1206. https://doi.org/10.1177/0956797616654455.

Rodrik, D. (2021). Why does globalization fuel populism? Economics, culture, and the rise of right-wing populism. *Annual Review of Economics*, *13*(1), 133–170. https://doi.org/10.1146/annurev-economics-070220-032416.

Roos, P., Gelfand, M., Nau, D., & Lun, J. (2015). Societal threat and cultural variation in the strength of social norms: An evolutionary basis. *Organizational Behavior and Human Decision Processes*, *129*, 14–23. https://doi.org/10.1016/j.obhdp.2015.01.003. SI: Social Norms and Cultural Dynamics.

Smith, A. (1759). *The theory of moral sentiments*. London: Penguin.

Starmer, C. (2000). Developments in nonexpected-utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, *38*(2), 332–382.

Sugden, R. (2018). *The community of advantage: A behavioural economist's defence of the market*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198825142.001.0001.

Szekely, A., Lipari, F., Antonioni, A., Paolucci, M., Sánchez, A., Tummolini, L., & Andrighetto, G. (2021). Evidence from a long-term experiment that collective risks change social norms and promote cooperation. *Nature Communications*, *12*(1), 1–7. https://doi.org/10.1038/s41467-021-25734-w.

Thöni, C., & Volk, S. (2018). Conditional cooperation: Review and refinement. *Economics Letters*, *171*, 37–40. https://doi.org/10.1016/j.econlet.2018.06.022.

Cite this article: Bogliacino, F., Charris, R., Gómez, C., and Montealegre, F. (2024). Negative economic shocks and the compliance to social norms. *Judgment and Decision Making*, e9. https://doi.org/10.1017/jdm.2024.1