

# Negative Economic Shocks and the Compliance to Social Norms\*

Francesco Bogliacino<sup>1</sup>, Rafael Charris<sup>2</sup>, Camilo Gomez<sup>3</sup>, and Felipe Montealegre<sup>4</sup>

<sup>1</sup>Corresponding Author. Dipartimento di Scienze Economiche, University of Bergamo, Bergamo, [francesco.bogliacino@unibg.it](mailto:francesco.bogliacino@unibg.it)

<sup>2</sup>NYU, Grossman School of Medicine, Department of Population Health, Epidemiology Division, New York City, USA

<sup>3</sup>CERGE-EI, Prague, Czech Republic

<sup>4</sup>University of Bologna, Bologna, Italy

October 2023

This is a previous version of the paper accepted at Judgment and Decision-Making.

<https://doi.org/10.1017/jdm.2024.1>

## Abstract

We study why suffering a Negative Economic Shock (NES), i.e. a significant loss, may trigger a change in other-regarding behaviour. 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 result can be formally proved when preferences are norm-dependent. We study this question experimentally: After administering losses on the earnings from a Real Effort Task, we analyze

---

\*Francesco acknowledges financial support from Open Evidence via contracts 18810 and 29047 with Universidad Nacional de Colombia. Ethical approval for the experiments of this projects has been issued by the Comite de Etica de la Universidad Nacional-Facultad de Medicina (Acta 012-159-19, Acta 016-163, Acta 020-189). 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, Eugenio Verrina, for comments and suggestions. Douglas Rodriguez and Sebastian Ramirez 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. The remaining errors are ours.

choices in prosocial and antisocial settings. To derive our predictions, we elicit social norms for each context analyzed in the experiments. We find robust evidence that shock increases deviations from norms.

**Keywords:** Negative Economic Shocks; Social Norms; Norm compliance; Anti Social Behavior; Cooperation.

**JEL Codes:** C91; C92; D90; D91.

# 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, injuries) or traumatic events (violence, disasters). The literature has studied NES extensively to grasp the effect of poverty (Mani et al., 2013; Haushofer and Fehr, 2014; Carvalho et al., 2016). 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 negative economic shocks have tended to support extreme candidates at the elections, which has furthered the interest in how suffering an NES shapes behaviour (Guiso et al., 2017; Inglehart and Norris, 2016; Rodrik, 2021; Benczes and Szabó, 2022).

When figuring out how NES affects social behaviour, 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 (Szekely et al., 2021; Roos et al., 2015; Gelfand et al., 2017; Heinrich, 2020; Prediger et al., 2014). For the aim of this article, a social norm is defined as a rule of behaviour which is contingent and supported by expectations of other people's behaviour 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 (Cassar et al., 2017; Botchway and Filippin, 2021). 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 (Kimbrough and Vostroknutov, 2018; Krupka and Weber, 2013; Bicchieri, 2006; 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 decision-maker 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 behavioural model. Assuming that norms enter the utility function and participants are heterogeneous in their psychological cost of compliance, we analyze optimal behaviour in several binary decision problems ( $d \in \{0, 1\}$ ) where a substantive norm applies, considering both anti-social and pro-social 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 three 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 (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 (2009)'s methodology. Participants provide their personal normative beliefs (PNBs) over the action space of the decision-maker. 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 behaviour (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 (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 four and five, the number with the highest payoff. The effect is equivalent to more than one-fourth of a sd. When we look at the JoD, the *decrease* in retaliation is as large as 50% of a sd, again supporting the prediction of the model. However, we cannot conclude that NES increases

defection in the prisoner’s dilemma.

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, (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 manuscript, 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 anti-social behaviour (Dube and Vargas, 2013; Bignon et al., 2017; Cortés et al., 2016; Boonmanunt et al., 2020). More importantly, it outperforms alternative explanations for the same findings.

This article makes a contribution to the literature on the consequences of shocks. We studied individual-level shocks, whereas the bulk of the literature focuses on aggregate shocks (Szekely et al., 2021; Roos et al., 2015; Gelfand et al., 2017; Heinrich, 2020; Prediger et al., 2014). Our contribution is complementary: experiencing the shock has a different effect from exposure to the threat. This differential behavioural effect of experience v. information has been documented in works unrelated to social norms (Malmendier, 2021a,b; D’Acunto et al., 2021). By focusing on random shock, we also distinguish ourselves from the papers on framed shocks (e.g. natural disaster) or intentional shocks (Bauer et al., 2016; Bogliacino et al., 2020; Eckel et al., 2022).

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 (Bicchieri et al., 2021; Andreoni and Bernheim, 2009; Bicchieri, 2010; 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; Cassar et al., 2017; Bogliacino et al., 2021). When it comes to social preferences, there exist several contributions on trust (Castillo and Hernandez, 2023; Bejarano et al., 2018, 2021; 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 Online Materials (SOM).

## **2 The model**

### **2.1 Intuition and the main prediction**

We study the problem of a decision-maker (DM) facing a binary choice involving a social norm. The action is harmful: it induces a loss to the counterpart or prevents her from enjoying a gain. The social norm is a rule of behaviour that may recommend abstaining or carrying out the action, depending on the setting. The DM derives utility from income, including assets and the monetary payoff from her choice, but has a conditional preference to follow the social norm. Acting in violation of a norm results in a psychological cost. DMs are indexed by their norm propensity ( $\theta$ ), to capture heterogeneity in norm sensibility.

What happens when a DM suffers an NES? Since an NES represents a reduction in the DM's asset position, suffering an NES increases the concern for money. This may happen because of decreasing marginal utility of income or because of loss aversion. This leads to more norm violations in equilibrium.

Consider the following settings: stealing, cheating, JoD, and PD. These are well known experimental games. In stealing, participants decide whether to “take” other people's money. In a cheating task, they may lie to their benefit. In JoD, they may harm the counterpart at a

| Setting            | Social Norm  | Prediction               |
|--------------------|--------------|--------------------------|
| Stealing           | Do Not Steal | $S(NES) > S(C)$          |
| Cheating           | Do Not Cheat | $C(NES) > C(C)$          |
| Joy of Destruction | Retaliation  | $D(NES) < D(C) P(D) = 1$ |
| PD                 | Tit-for-Tat  | $C(NES) < C(C) P(C) = 1$ |

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

cost. And in PD, cooperating is not an equilibrium but it is collectively more efficient. In these settings, we assume that the norms are: *thou shalt not steal*, *thou shalt not lie*, retaliate, and tit-for-tat (conditional cooperation). While we analyze the problem theoretically, these norms are assumed to hold, but in the next Section, we will discuss how we validated them in a norm elicitation experiment. In all these cases, norm compliance is costly and the NES should induce norm violation. This claim translates into specific predictions for each setting. The predictions are summarized in Table 1 below.

Notice that for JoD and PD, the prediction is conditional on the belief that the counterpart will carry out the action. In fact, the trade-off exists only under that belief. On the contrary, in the JoD, if the DM expects the counterpart will abstain, norms and incentives recommend the same action. The same applies to the PD. Beliefs are elicited in the experiments.

In the following subsections, we will provide a formal justification.

## 2.2 The optimal choice

Formally, a DM ( $i$ ) should choose  $d_i \in \{0, 1\}$ . If the setting is strategic, she will be interacting with  $j$ . By convention,  $d = 1$  is the harmful action, defined as the action that causes a loss to the counterpart or prevents her from enjoying a gain. Preferences include two terms. The first is the utility of income: an additively separable utility function  $u(e + w(d_i, d_j))$ , where  $w(\cdot)$  is the monetary payoff, and  $e$  is the initial endowment. The second term is  $\mathbb{1}_{d_i \neq n} c(\theta_i)$ , the psychological cost of deviating from a social norm  $n$ . The cost increases in  $\theta_i$ , the propensity to comply. We have  $\theta_i \in [0, 1]$ , with Cumulative Density Function  $F(\cdot)$ . The preferences are similar to Krupka and Weber (2013), Kimbrough and Vostroknutov (2018), and Levitt and List (2007). Models of preferences with social image have a similar framework, but the social image is endogenous (Andreoni and Bernheim, 2009; Benabou and Tirole, 2006).

The problem can be written as follows

$$\max_{d_i \in \{0,1\}} u(e + w(d_i, d_j)) - \mathbf{1}_{d_i \neq n} c(\theta_i) \quad (1)$$

In some situations, like stealing,  $d = 1$  transgresses a social norm (i.e.  $n = 0$ ), in others, like punishment, it is prescribed by the norm ( $n = 1$ ). If the norm is conditional, as in tit-for-tat, we will use the notation  $n = d_j$ .

An NES is modelled as  $de < 0$ .

The following assumptions hold:

**Assumption 1.**  $u(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$

$$u'(\cdot) > 0$$

$$u''(\cdot) < 0$$

**Assumption 2.**  $c(\cdot) : [0, 1] \rightarrow \mathbb{R}$

$$c'(\theta) > 0, \quad c''(\theta) > 0$$

Assumption 1 is the standard decreasing marginal utility of income. Assumption 2 formalizes the utility cost of norm violation and the dependence on the psychological parameter  $\theta$ .

To understand the logic of the argument, consider a non strategic choice where a fairness norm is in place ( $n = 0$ ) and  $d = 1$  is a transgression. An agent of parameter  $\theta$  chooses  $d = 1$  if  $u(e + w(1)) - u(e + w(0)) \geq c(\theta)$ . The term  $u(e + w(1)) - u(e + w(0))$  captures the benefit  $B$  accruing from transgressing the norm, constant across agents. The cost is increasing in  $\theta$ . In Figure 1a, we plot the optimal choice as a function of  $\theta$ : there is a threshold  $\bar{\theta} = \theta_1$  below which agents will transgress, and above which they will comply.

What happens when a DM suffers an NES? Due to the concavity of the utility function, the marginal utility of transgression increases, leading to more norm violations. In Figure 1a, for the new benefit curve, more people choose to carry out  $d = 1$ , i.e.  $\bar{\theta}$  moves to the right, from  $\theta_1$  to  $\theta_2$ .

Consider also the opposite situation where  $d = 1$  is costly and recommended by the norm (i.e.



$n = 1$ ). An agent of parameter  $\theta$  chooses  $d = 1$  if  $u(e + w(1)) - u(e + w(0)) \geq -c(\theta)$ . The left-hand side is the utility loss from punishment, and the right-hand side is the utility cost of norm violation. In presence of an NES, concavity implies that  $\frac{\partial u(e+w(1))-u(e+w(0))}{\partial e} > 0$ , the utility loss from following the norm increases and less people will choose  $d = 1$ . This is illustrated in Figure 1b.

In settings with interaction, we need to introduce strategic uncertainty: the DM will now maximizes  $E[u(e+w(d_i, d_j)) - \mathbb{1}_{d_i \neq n} c(\theta)]$ . Define  $p$  to be the expected likelihood that  $d_j$  chooses 1. There are three cases, either  $n = 0$ ,  $n = 1$ , or  $n = p$  (tit-for-tat). We can write the expression in a compact form as  $p(u(e+w(1,1)) - u(e+w(0,1))) + (1-p)(u(e+w(1,0)) - u(e+w(0,0))) \geq (1-2n)c(\theta)$ .

Consider when the norm is tit-for-tat. The DM chooses 1 if  $p(u(e+w(1,1)) - u(e+w(0,1))) + (1-p)(u(e+w(1,0)) - u(e+w(0,0))) \geq (1-2p)c(\theta)$ . There are three terms:  $u(e+w(1,1)) - u(e+w(0,1))$  is the utility loss from retaliation,  $u(e+w(1,0)) - u(e+w(0,0))$  is the benefit of defection, and  $(1-2p)c(\theta)$  is the (expected) psychological cost.

We will derive our predictions in the two extreme cases,  $p = 0$  and  $p = 1$ . These predictions are testable once the beliefs are elicited in an experiment. Conditioning on a degenerate belief also represents a plausible description of decision-making in one-shot interactions. In SOM, Section B, we will show that the conclusions are supported in equilibrium by a formal comparative statics result.

Under  $p = 1$ , there is a cost of retaliation if  $u(e+w(0,1)) - u(e+w(1,1)) > 0$ . When the latter condition holds, the DM chooses  $d = 1$  only if the cost of transgression is larger than the cost of retaliation. Since an NES raises the cost of retaliation, the share of DMs who chooses  $d = 1$  decreases. If retaliation is not costly, everybody will make the same choice, regardless of the shock.

Under  $p = 0$ , there is a benefit from defection if  $u(e+w(1,0)) - u(e+w(0,0)) > 0$ . The optimal choice is determined by whether  $u(e+w(1,0)) - u(e+w(0,0)) \geq c(\theta)$ . Since an NES increases the benefit from defection, the share of DM who chooses  $d = 1$  increases. If defection is not profitable, everybody will comply, regardless of the shock.

The reasoning for  $n = 0$  and  $n = 1$  are special cases of the tit-for-tat.

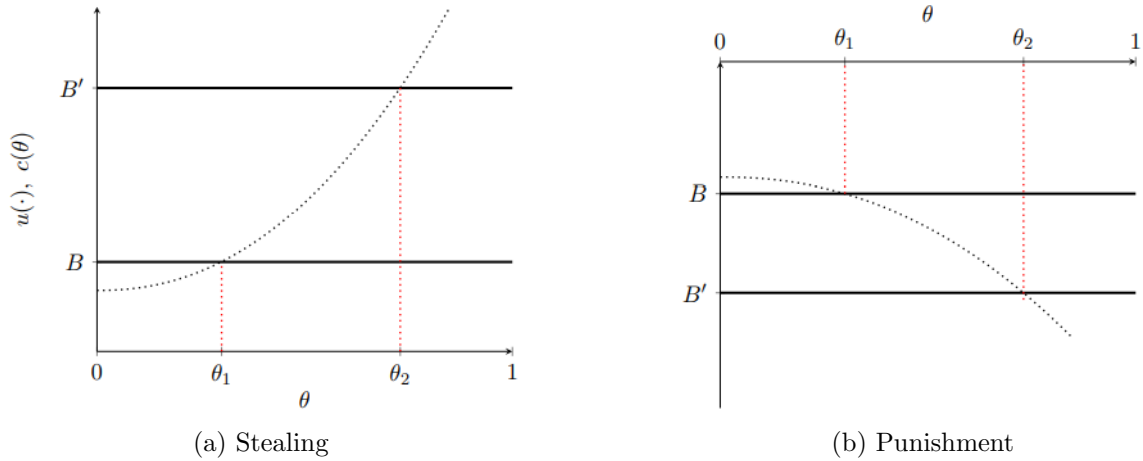


Figure 1: The optimal choice

### 2.3 Settings

Consider our four settings: cheating, stealing, Joy of Destruction (JoD), and cooperation in prisoner's dilemma (PD).

In the cheating and stealing task, the payoffs for the DM are  $w(1) > w(0)$  and the norm is  $n = 0$ .

The JoD is a simultaneous interaction where  $d = 1$  is costly and harmful.  $d = 1$  is called money burning. 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. More generally, it must hold that  $w(0, 0) > w(1, 0) > w(0, 1) > w(1, 1)$ . The social norm is to *retaliate*, which accounts for the evidence of costly money burning (Abbink and Herrmann, 2011). Retaliate means burning when expecting the counterpart to burn.

The prisoner's dilemma is a symmetric simultaneous game where  $w(1, 0) > w(0, 0) > w(1, 1) > w(0, 1)$ . We assume that the relevant social norm is conditional cooperation (Gächter, 2007).

### 2.4 Theoretical Predictions

As discussed in Section 2.2, when there is a trade-off between income and norm compliance, an NES makes people more attentive to income, leading to more transgression. For a trade-off to exist, following the norm should be costly in terms of payoff. This is the case for cheating and stealing, where the cost of following the norm is the loss of income that would accrue from

choosing  $d = 1$ .

For the JoD, the trade-off exists when a DM expects the counterpart to burn because burning is costly but the retaliation is prescribed by the norm. The prisoners' dilemma has a clear trade-off: under tit-for-tat, conditional cooperation is costly because defection is profitable. An NES generates more norm violations.

Table 1 summarizes these predictions. For stealing (cheating), under the norm of do not steal (lie), we expect an increase in stealing (cheating) due to NES. In JoD, under the norm of retaliation, we expect subjects to retaliate less due to NES. Finally, in the PD under the norm of conditional cooperation, we expect subjects to defect more (cooperate less when expecting the counterpart to cooperate) due to NES. A formal discussion is in the SOM, Section A.

## 2.5 Alternative Microfoundation: Loss Aversion

We used concavity to derive the prediction. Under Assumption 1, an NES is just a wealth effect. On average, individuals are risk averse (Camerer, 1995; Starmer, 2000), thus concavity seems a reasonable assumption. However, an implication of concavity is that differences in assets and not just shocks to assets would increase transgression. Notice that this interpretation would be coherent with the claim that shocks represent a plausible variation to study the causal effect of poverty (Mani et al., 2013; Haushofer and Fehr, 2014; Boonmanunt et al., 2020).

We can separate the cases of wealth effect from that of shock by assuming Loss aversion, i.e. the standard tenet that losses loom larger than gains (Kahneman and Tversky, 1979). When we condition on the belief, the problem of the DM can be reduced to one of the two cases, where  $d = 1$  is either costly but recommended or profitable but forbidden. As a result, we can prove the general argument without strategic interaction. Assume *a fortiori* that the utility function is linear (risk neutrality) but with loss aversion. The problem of the DM becomes:

$$\max_{d \in \{0, 1\}} e' + w(d) - v^l(\max\{0, e - e' - w(d)\}) - \mathbf{1}_{d_i \neq n} c(\theta) \quad (2)$$

with  $v^l(\cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and increasing, and  $e'$  is the current endowment, either equal to  $e$ , in the control, or lower than  $e$  in case of NES. In the formulation of the  $v^l$  function,  $e$  is the reference point.

In the control, the DM chooses 1 if  $w(1) - w(0) \geq (1 - 2n)c(\theta)$ , in presence of a (large enough) shock, and defining  $\Delta e = e - e'$ , if  $w(1) - w(0) + v^l(\Delta e - w(0)) - v^l(\Delta e - w(1)) \geq (1 - 2n)c(\theta)$ . Monotonicity of  $v$  ensures that this condition is met.

With loss aversion, the model generates the same predictions as in Table 1, but without reducing an NES to a wealth effect. For instance, a positive shock would be void of consequences in this case, whereas under concavity the shock effect would be symmetrical.

## 2.6 Equilibrium and Comparative Statics: general results

Table 1 presents the predictions under  $p = 0$  or  $p = 1$ . These are testable given that the beliefs are elicited in the experiments. They are also plausible as a description of how a DM interacts in a one-shot decision. They can be generalized as a formal equilibrium prediction if  $\theta_j$  belongs to  $i$ 's information set.

Alternatively, Assumption 3 states that the distribution  $F(\cdot)$  of the norm sensibility parameter is common knowledge. We can show that the direction of the effect of the NES is maintained.

**Assumption 3.**  $F(\theta)$  is common knowledge.

This is the definition of equilibrium:

**Definition 1.** Given a symmetric simultaneous  $2 \times 2$  game, with preferences  $u(e + w(d_i, d_j)) - \mathbb{1}_{d_i \neq n}c(\theta_i)$ , with randomly drawn players  $i, j$ , finite payoffs functions  $w(d_i, d_j)$ , an equilibrium with social norm  $n$  is a distribution of choices for the population such that each DM maximizes her utility and expectations are mutually consistent.

We apply the refinement that the equilibrium is stable. Here, stability means that small perturbations induce incentives that drive behaviour towards equilibrium.

The following proposition holds (the proof is in the SOM, Section B).

**Proposition 1.** Under assumptions 1, 2 and 3, the following comparative statics hold in equilibrium: a) in the JoD,  $\frac{\partial P(d=1)}{\partial e} > 0$ ; a) in the PD,  $\frac{\partial P(d=1)}{\partial e} < 0$ .

## 2.7 Testing the predictions: methods

The assessment of the predictions requires two steps. First, since the predictions rely on the assumptions that norms  $x$  or  $y$  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. They 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 are contingent (Bicchieri, 2006), 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 a RET for money. After receiving feedback on their performance, subjects were assigned randomly to either a treatment condition, where they suffered an NES, or a control condition, where they maintained what was previously earned. Following this phase, participants perform one or several tasks among stealing, cheating, JoD, and PD. Sections 4-6 present the design and results of this set of experiments.

Finally, after assessing the prediction, we perform 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 introduce additional treatments where we manipulate incentives ex ante, instead of ex post with shocks, to identify the difference between wealth effects and shocks. Section 7 describes 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 two main methods to elicit normative expectations: the coordination game by Krupka and Weber (2013, 2009) and the two steps 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 two steps elicitation method by Bicchieri and Xiao (2009) recovers first-order and second-order beliefs. Subjects report their Personal Normative Beliefs (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 SOM, Section D-IV). Data were gathered in February 2021.

Participants went through two 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 main task from the experiment conducted by Bogliacino et al. (2020). The sequence of questions came in random order. In total, participants evaluated six decisions, three antisocial and three 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 twelve predictions, and one was randomly selected for payment at the end. A correct guess was paid 25000 COP. The show-up fee was 10000 COP. The average time of completion was 35 mins. We collected 109 observations. On average, participants earned 21000 COP (6 USD). Participants did not make decisions in the settings or participate in the experiments described in the survey. This ensures the elicitation of normative expectations separately from behaviour (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 three numbers unconditionally, reporting four or five unconditionally, reporting six unconditionally, misreporting the drawn number plus or minus a maximum of two to own advantage, misreporting the drawn number plus or minus a maximum of two 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 2. 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 two 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 2 plots the most frequent response to the normative expectations. In SOM, Section C, we report the tabulation of responses for each task.

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

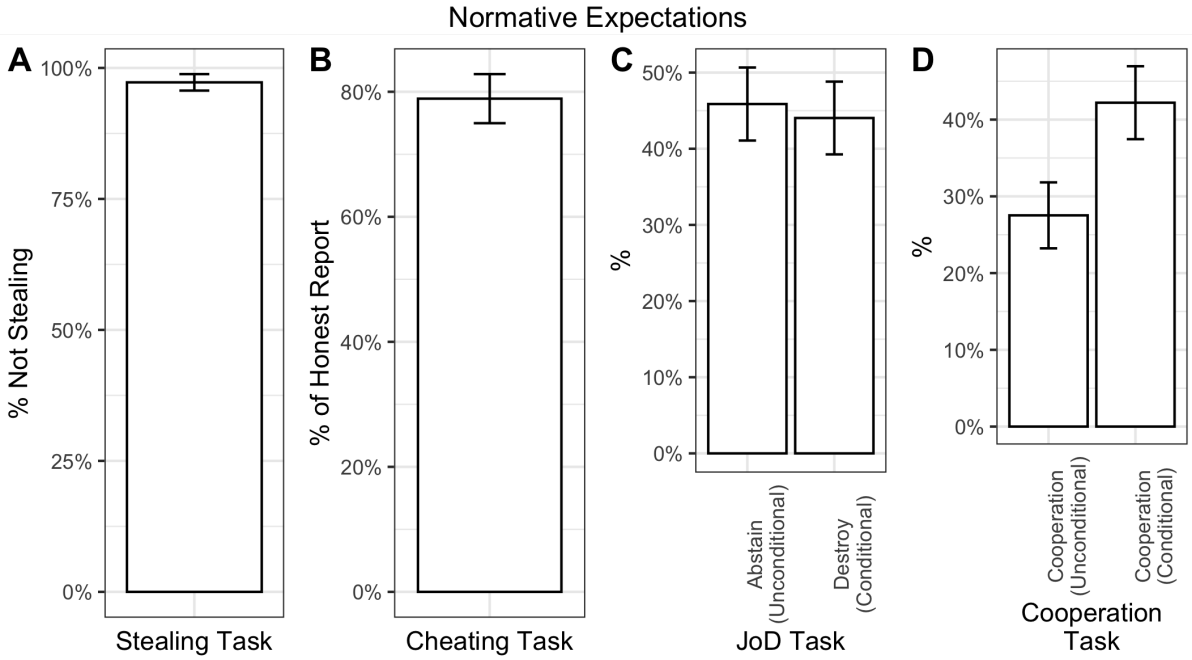


Figure 2: The elicited normative expectations

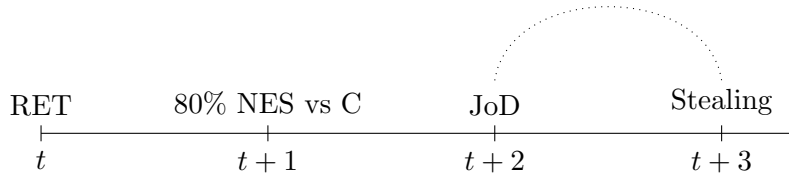


Figure 3: The timeline of Experiment I

## 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 earning from a Real Effort Task (RET), experienced with a 50% probability. The probability was common information. The control is the set of those who did not experienced the shock. The RET was the Niederle and Vesterlund (2007)'s task of summing sequences of two-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 induce possible retaliatory behaviour changing the nature of the task with respect to Section 2.3. Additionally, two 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 one 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 of the experiment is reported in Figure 3. The procedures were as follows. 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 nine sessions in total. The exchange rates were 1000 COP per ECU. On average each participant earned 17000 ( $\pm 5200$ ) COP (approximately USD

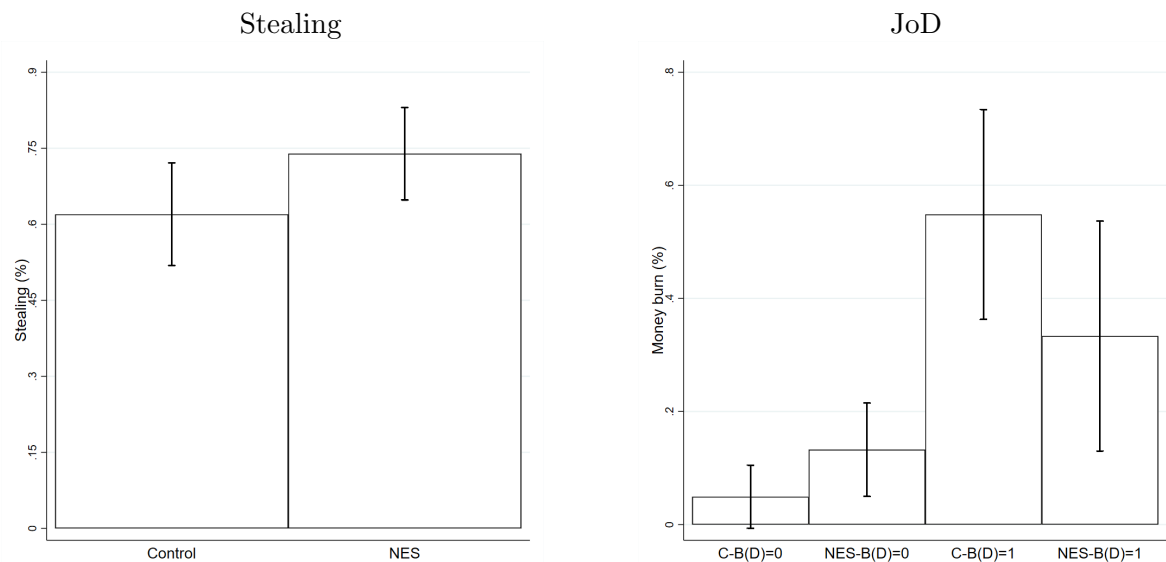


Figure 4: The impact of NES in the Stealing and JoD tasks

5 at the time). The experiment is programmed in oTree (Chen et al., 2016) and the English version of the protocol is available in the SOM, Section D-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 4 (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.

To assess the prediction that stealing increases with NES, we run an OLS regression, controlling for order, and compute a one sided test, since we are postulating a direction for the alternative hypothesis. Table 2, Column (1), reports the results. The increase in stealing is both economically relevant, around 25% of a standard deviation of the outcome in the control condition, and statistically significant ( $F(1, 181) = 3.07$ ,  $p = 0.040$  one-sided).

The behavior in the JoD is shown in Figure 4 (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 zero. 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 2 and replicate the standard results in JoD experiments (Abbink and Herrmann, 2011). However,

Table 2: OLS estimates of effect of NES on Stealing and JoD

|            | (1) Stealing        | (2) JoD             |
|------------|---------------------|---------------------|
| NES        | 0.120<br>(0.068)    |                     |
| Order      | 0.126<br>(0.070)    | 0.073<br>(0.052)    |
| NES-B(D)=0 |                     | 0.080<br>(0.049)    |
| C-B(D)=1   |                     | 0.486***<br>(0.096) |
| NES-B(D)=1 |                     | 0.277**<br>(0.102)  |
| Constant   | 0.545***<br>(0.067) | 0.011<br>(0.037)    |
| N          | 184                 | 184                 |
| R2         | 0.03                | 0.21                |
| F-test     | 3.00                | 8.95***             |

Robust standard errors shown in parenthesis. \*  $p < 0.05$ ,  
\*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

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 (51.85% of a sd computed for the control condition) and statistically significant ( $F(1, 179) = 2.44$ ,  $p = 0.05$  one-sided).

The supporting regression is reported in Table 2, Column (2). We run an OLS regressions with three dummies: an indicator for the condition NES and  $B(D) = 0$ , an indicator for the condition Control and  $B(D) = 1$ , and an indicator for the condition NES and  $B(D) = 1$ . The omitted category is control and  $B(D) = 0$ .

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 two 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 a 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

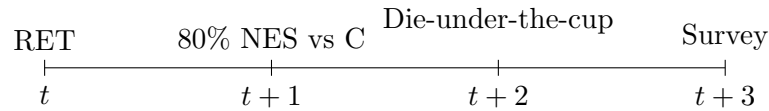


Figure 5: The timeline of Experiment II

control, the RET was changed with respect to Experiment I. Logging in from home, some participants may access a calculator, changing the nature of the RET task. Instead, we adopted a 4 minutes 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 (2013)’s die-under-the-cup. In this task, participants rolled a dice privately and reported their results. In the instructions, an online dice was made available, highlighting that the dice was 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 2000 COP times the reported number (from one to five) and zero 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. The timeline is depicted in Figure 5.

Participants received a show up fee, the earnings from the RET (to make the treatment salient) and the dice. In total, we recruited 158 participants. Data collection occurred in June 2020, on average each participant earned around 25000 COP (around 7 USD).

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

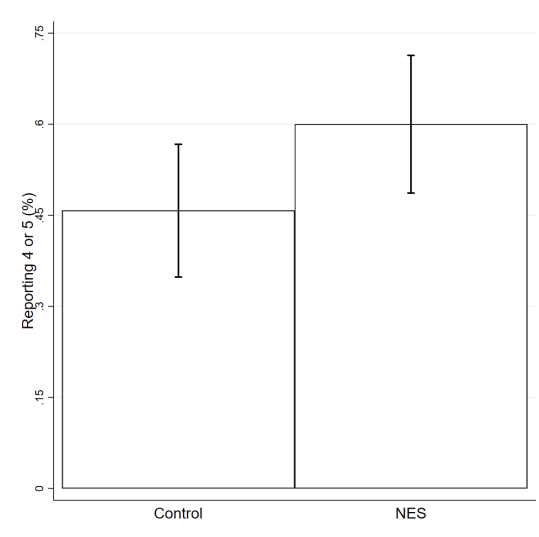


Figure 6: The impact of NES in the die-under-the-cup task

## 5.2 Results

On average, participants tried  $11.91 \pm 3.78$  transcriptions, completing successfully  $9.17 \pm 4.15$  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).

There was a considerable amount of cheating: we neatly reject the null hypothesis that the observed data comes 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, but we can measure how the likelihood of reporting the numbers with the highest payoff (four or five) differs between treatment and control, as in Bogliacino and Monteleagre (2020).

Figure 6 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 28.36% of a sd of the outcome in the control condition and is statistically significant ( $t = -1.79$ ,  $p = 0.037$  one sided, controlling for unequal variance).

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

Table 3: Experiment III: the stage game

## 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  $2 \times 1$  between-subject design. In the first part, participants performed a RET (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 two experiments.

Table 3 show the normal form of the PD. Strategies C and NC were labelled green and blue. The second innovation with respect to the standard design consisted in the introduction of this procedure: a) participants declared their strategy conditional on the belief that the counterpart was playing C and NC; b) then they declare which scenario was more likely among the counterpart playing C, NC, or choosing randomly. Subjects knew that the answer determine their behaviour. Eliciting the belief contingent decisions represents the cleanest test of our prediction, and overcomes the endogeneity of the belief.

Since we asked four 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 SOM, Section E-II).

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

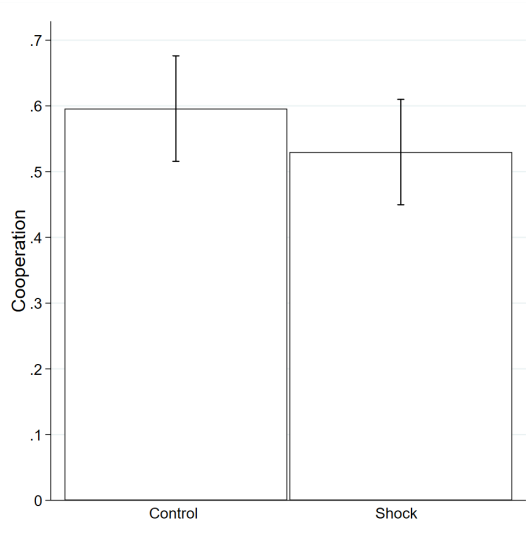


Figure 7: The impact of NES in the Prisoners' Dilemma

## 6.2 Results

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

These are the main results. 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 ( $F(1, 295) = 0.01$ ,  $p = 0.462$  one 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 ( $F(1, 295) = 1.32$ ,  $p = 0.126$  one sided) and represents around 15% of a 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 three different experiments, NES induce more norm violations. The evidence is robust for anti-social tasks and less so for the prisoner's dilemma. To provide a robustness check, we designed an additional experiment. This experiment improves on the previous design in three ways: a) it separately controls for a wealth effect, b) it imposes the rule to follow instead of

relying on elicited social norm, c) it allows for an intensive margin to increase statistical power.

Participants faced a RET under four possible conditions: (1) each correct transcription is paid one point, the potential 80% loss is announced but not administered (Control); (2) each correct transcription is paid one point, the potential shock is announced and administered (Shock treatment); (3) each correct transcription is paid one 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 two phases, where the payment could change between phases one and two. Treatments (3) and (4) have no second phase.

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 five traffic lights. Endowed with 30 seconds, each second worth one 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 five seconds for a traffic light to change from red to green. The protocol is programmed in oTree (Chen et al., 2016) and can be found in the SOM-Section D-V.

This is a between-subject incentivized survey. 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 SOM, Section E-I). 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 is much lower than expected.

These are the results. The performance in the RET was not 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 ( $\pm 6.80$ ). In treatment (2) (the NES condition in Experiment I-III) the outcome was 4.95 ( $\pm 2.73$ ). A (one sided) t-test controlling for unequal variance returns  $t = 2.46$  (one sided  $p = 0.008$ ). In treatment (3), the outcome was 5.89 ( $\pm 4.34$ ), whereas in treatment (4) it was



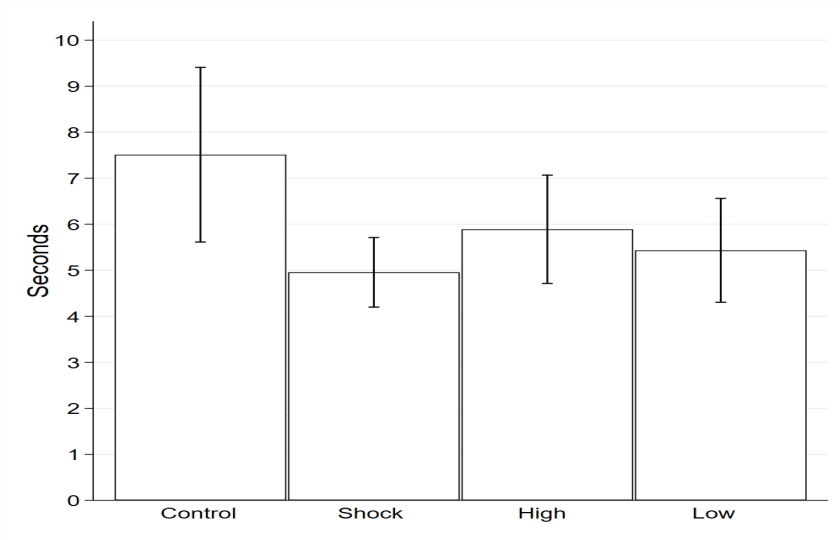


Figure 8: Results from the rule following task

5.43 ( $\pm 4.28$ ). A t-test controlling for unequal variance returns  $t = 0.55$  (one sided  $p = 0.290$ ). In other words, the wealth effect did not generate the same result as the shock. Results are plotted in Figure 8.

We run an OLS regression using the seconds as the outcome variable, with robust standard errors (Table 4). The difference between the shock effect (dummy for treatment 2) and the wealth effect (the difference in performance between treatment 3 and 4) is 2.94 seconds and is statistically significant ( $t = -2.27$ ,  $p = 0.024$ ).

Table 4: OLS estimates of effect of NES on Rule Following

| <b>(1) Rule Following</b> |                     |
|---------------------------|---------------------|
| NES                       | -2.556*<br>(1.036)  |
| High                      | -1.621<br>(1.133)   |
| Low                       | -2.079<br>(1.120)   |
| Constant                  | 7.511***<br>(0.963) |
| N                         | 210                 |
| R2                        | 0.04                |
| F-test                    | 2.27                |

Robust standard errors shown in parenthesis.  
\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## 8 Discussion and Concluding remarks

What happens to norm compliance when subjects face an NES? We explore this question experimentally. Norm-dependent utility predicts that subjects will deviate more from social norms. To assess this prediction, we randomly administer shocks on the earnings from a RET and elicit the behaviour of participants in several tasks where there exists a rule of behavior grounded on normative expectations. The experiments show how experiencing wealth losses can be a source of norm violation.

The main results match several stylized facts. The fact that NES may generate antisocial behaviour, particularly crimes against property, has not gone unnoticed. Compelling quasi-experimental evidence document a positive relationship between negative economic shocks and antisocial behaviour. For example, Dube and Vargas (2013) use the change in coffee prices to study variations in crime in communities highly dependent on harvest income. Cortés et al. (2016) use the collapse of the Ponzi scheme in Colombia to detect variation in a portfolio of criminal activities. Bignon et al. (2017) exploit the regional variation in the exposure to phylloxera in wine-producing regions in France to identify the increase in property crime. Dix-Carneiro et al. (2018) use the trade liberalization shock in Brazil to estimate the causal impact of the shock on criminal activity. Weather shocks have also led to an increase in property crimes (Mehlum et al., 2006). Cheating has been less studied. Aksoy and Palma (2019) look at cheating under “scarcity” - the shock around paycheck variation - but could not detect any significant variation. Bogliacino and Montealegre (2020) also look at the effect of NES in the die-under-the-cup task, finding no effect, but the presence of four tasks may have diluted the incentives. Additionally, Boonmanunt et al. (2020) document that people under scarcity are less responsive to a social norm intervention, in line with our main argument.

Our theoretical framework also outperformed other theories. In the JoD, the norm of retaliation generates a trade-off between compliance and income (money burning is costly). Although grounded on the same reasoning as in stealing and cheating, the model predicts *less* antisocial activity and is consistent with our controlled evidence. 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). Psychological

theories of NES cannot fully account for the experimental evidence: NES induce cognitive load (Mani et al., 2013; Bogliacino and Montealegre, 2020) but there is no consensus on the relationship between cognitive load and social preferences (Alós-Ferrer and Garagnani, 2020; Rand, 2016; Kvarven et al., 2020). Haushofer and Fehr claim that people experiencing shocks may face increasing risk aversion (Haushofer and Fehr, 2014) 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).

This article has 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 -Acronym of 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 UK confirmed that this might be the case (d'Este and Harvey, 2022).

## References

- Abbink, Klaus and Benedikt Herrmann, “The moral costs of nastiness,” *Economic Inquiry*, 2011, 49 (2), 631–633.
- Aksoy, Billur and Marco A. Palma, “The effects of scarcity on cheating and in-group favoritism,” *Journal of Economic Behavior and Organization*, 2019, 165, 100–117.
- Alós-Ferrer, Carlos and Michele Garagnani, “The cognitive foundations of cooperation,” *Journal of Economic Behavior and Organization*, 2020, 175, 71–85.
- Andreoni, James and B Douglas Bernheim, “Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects,” *Econometrica*, 2009, 77 (5), 1607–1636.
- Ayton, Peter, Gennaro Bernile, Alessandro Bucciol, and Luca Zarri, “The impact of life experiences on risk taking,” *Journal of Economic Psychology*, 2020, 79 (April), 1–4.
- Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts, “Can War Foster Cooperation?,” *Journal of Economic Perspectives*, 2016, 30 (3), 249–274.
- Bejarano, H., J. Gillet, and I. Rodriguez-Lara, “Do negative random shocks affect trust and trustworthiness?,” *Southern Economic Journal*, 2018, 85 (2), 563–579.
- Bejarano, Hernán, Joris Gillet, and Ismael Rodriguez-Lara, “Trust and trustworthiness after negative random shocks,” *Journal of Economic Psychology*, 2021, 86, 102422.
- Benabou, R and J Tirole, “Incentives and Prosocial Behavior,” *American Economic Review*, 2006, 96 (5), 1652–1678.
- Benczes, István and Krisztina Szabó, “An Economic Understanding of Populism: A Conceptual Framework of the Demand and the Supply Side of Populism,” *Political Studies Review*, 2022, 0 (0), 14789299221109449.
- Bicchieri, Cristina, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge University Press, 2006.
- , “Norms, preferences, and conditional behavior,” *Politics, Philosophy & Economics*, 2010, 9 (3), 297–313.
- and Erte Xiao, “Do the Right Thing: But Only if Others Do So,” *Journal of Behavioral Decision Making*, 2009, 22 (October 2008), 191–208.
- , Eugen Dimant, and Erte Xiao, “Deviant or wrong? The effects of norm information on the efficacy of punishment,” *Journal of Economic Behavior & Organization*, 2021, 188, 209–235.
- Bignon, Vincent, Eve Caroli, and Roberto Galbiati, “Stealing to Survive? Crime and Income Shocks in Nineteenth Century France,” *Economic Journal*, 2017, 127 (599), 19–49.
- Blanco, Mariana, Dirk Engelmann, Alexander K. Koch, and Hans Theo Normann, “Belief elicitation in experiments: Is there a hedging problem?,” *Experimental Economics*, 2010, 13 (4), 412–438.

- Bogliacino, Francesco and Felipe Montealegre**, “Do negative economic shocks affect cognitive function, adherence to social norms and loss aversion?,” *Journal of the Economic Science Association*, 2020, 6 (1), 57–67.
- , **Camilo E Gomez, and Gianluca Grimalda**, “Crime-related Exposure to Violence and Social Preferences: Experimental Evidence from Bogota,” 2020.
- , **Cristiano Codagnone, Felipe Montealegre, Frans Folkvord, Camilo Gómez, Rafael Charris, Giovanni Liva, Francisco Lupiáñez-Villanueva, and Giuseppe A. Veltri**, “Negative shocks predict change in cognitive function and preferences: assessing the negative affect and stress hypothesis,” *Scientific Reports*, 2021, 11 (1), 1–10.
- Boonmanunt, Suparee, Agne Kajackaite, and Stephan Meier**, “Does poverty negate the impact of social norms on cheating?,” *Games and Economic Behavior*, 2020, 124, 569–578.
- Botchway, Ebo and Antonio Filippin**, “Cooperation as Adaptation to Persistent Risk of Natural Disasters : Evidence from a Natural Experiment,” 2021.
- Camerer, Colin**, “Individual Decision Making,” in John H Kagel and Alvin E Roth, eds., *The Handbook of Experimental Economics*, Princeton University Press, 1995, pp. 587–704.
- Carvalho, Leandro, Stephan Meier, and Stephanie Wand**, “Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday,” *American Economic Review*, 2016, 42 (2), 407–420.
- Cassar, Alessandra, Andrew Healy, and Carl von Kessler**, “Trust, Risk, and Time Preferences After a Natural Disaster: Experimental Evidence from Thailand,” *World Development*, 2017, 94, 90–105.
- Castillo, Jose Gabriel and Manuel A. Hernandez**, “The unintended consequences of confinement: Evidence from the rural area in Guatemala,” *Journal of Economic Psychology*, 2023, 95 (May 2022), 102587.
- Chavez, Alex K. and Cristina Bicchieri**, “Third-party sanctioning and compensation behavior: Findings from the ultimatum game,” *Journal of Economic Psychology*, 2013, 39, 268–277.
- Chen, Daniel L., Martin Schonger, and Chris Wickens**, “oTree-An open-source platform for laboratory, online, and field experiments,” *Journal of Behavioral and Experimental Finance*, 2016, 9, 88–97.
- Chuang, Yating and Laura Schechter**, “Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results,” *Journal of Development Economics*, 2015, 117, 151–170.
- Cortés, Darwin, Julieth Santamaría, and Juan F Vargas**, “Economic shocks and crime: Evidence from the crash of Ponzi schemes,” *Journal of Economic Behavior and Organization*, 2016, 131, 263–275.
- D’Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber**, “Exposure to Grocery Prices and Inflation Expectations Ulrike Malmendier Juan Ospina,” *Journal of Political Economy*, 2021, 129 (5), 1615—1639.
- Dana, Jason, Roberto A. Weber, and Jason Xi Kuang**, “Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness,” *Economic Theory*, 2007, 33 (1), 67–80.

- Dix-Carneiro, Rafael, Rodrigo R. Soares, and Gabriel Ulyssea**, “Economic Shocks and Crime: Evidence from the Brazilian Trade Liberalization,” *American Economic Journal: Applied Economics*, 2018, 10 (4), 158–195.
- Dube, Oeindrila and Juan F. Vargas**, “Commodity price shocks and civil conflict: Evidence from Colombia,” *Review of Economic Studies*, 2013, 80 (4), 1384–1421.
- d’Este, Rocco and Alex Harvey**, “The Unintended Consequences of Welfare Reforms: Universal Credit, Financial Insecurity, and Crime,” *Journal of Law, Economics, and Organization*, 2022, *Online First*, 521–565.
- Eckel, Catherine C., Enrique Fatas, and Malcolm Kass**, “Sacrifice: An experiment on the political economy of extreme intergroup punishment,” *Journal of Economic Psychology*, 2022, 90 (January), 102486.
- Fischbacher, Urs and Franziska Föllmi-Heusi**, “Lies in disguise—an experimental study on cheating,” *Journal of the European Economic Association*, 2013, 11 (3), 525–547.
- Friehe, Tim and Jan Marcus**, “Lost job, lost trust? On the effect of involuntary job loss on trust,” *Journal of Economic Psychology*, 2021, 84 (February), 102369.
- Gächter, S.**, “Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications,” in BS Frey and A Stutzer, eds., *Economics and psychology: A promising new cross-disciplinary field*, MIT Press, 2007, pp. 19–50.
- Gelfand, Michele J, Jesse R Harrington, and Joshua Conrad Josuah Conrad Jackson**, “The Strength of Social Norms Across Human Groups,” *Perspectives on Psychological Science*, 2017, 12 (5), 800–809.
- Gintis, Herbert**, “A framework for the unification of the behavioral sciences,” *Behavioral and Brain Sciences*, 2007, 30 (1), 1–16.
- Greiner, Ben**, “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, 2015, 1 (1), 114–125.
- Guiso, L, H Herrera, M Morelli, and T Sonno**, “Demand and Supply of Populism,” 2017.
- Haushofer, Johannes and Ernst Fehr**, “On the psychology of poverty,” *Science*, 2014, 344 (6186), 862–867.
- Heinrich, Joseph**, *The WEIRDest people in the World*, Farra, Straus and Giroux, 2020.
- , **Steven J Heine, and Ara Norenzayan**, “Weird people,” *Behavioral and Brain Sciences*, 2010, 33 (2-3), 61–135.
- Inglehart, Ronald and Pippa Norris**, “Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash,” 2016.
- Kahneman, Daniel and Amos Tversky**, “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 1979, 47 (2), 293–298.
- Kimbrough, Erik O. and Alexander Vostroknutov**, “A portable method of eliciting respect for social norms,” *Economics Letters*, 2018, 168, 147–150.
- Krupka, Erin L and Roberto A Weber**, “The focusing and informational effects of norms on pro-social behavior,” *Journal of Economic Psychology*, 2009, 30 (3), 307–320.

- **and** –, “Identifying social norms using coordination games: Why does dictator game sharing vary?,” *Journal of the European Economic Association*, 2013, 11 (3), 495–524.
- Kvarven, Amanda, Eirik Strømmand, Conny Wollbrant, David Andersson, Magnus Johannesson, Gustav Tinghög, Daniel Västfjäll, and Kristian Ove R Myrseth**, “The intuitive cooperation hypothesis revisited: a meta-analytic examination of effect size and between-study heterogeneity,” *Journal of the Economic Science Association*, 2020, 6 (1), 26–42.
- Levitt, Stephen J and John A. List**, “What do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?,” *Journal of Economic Perspectives*, 2007, 21 (2), 153–174.
- List, John A**, “On the Interpretation of Giving in Dictator Games,” *Journal of Political Economy*, 2007, 115 (3), 482–493.
- List, John A., Robert P. Berrens, Alok K. Bohara, and Joe Kerkvliet**, “Examining the role of social isolation on stated preferences,” *American Economic Review*, 2004, 94 (3), 741–752.
- Malmendier, Ulrike**, “Experience Effects in Finance: Foundations, Applications, and Future Directions,” *Review of Finance*, 2021, 25 (5), 1339–1363.
- , “Why personal histories matter in Economics,” *Journal of the European Economic Association*, 2021, 19 (6), 2857–2894.
- Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao**, “Poverty impedes cognitive function,” *Science*, 2013, 341 (6149), 976–980.
- Mehlum, Halvor, Edward Miguel, and Ragnar Torvik**, “Poverty and crime in 19th century Germany,” *Journal of Urban Economics*, 2006, 59 (3), 370–388.
- Mehta, Judith, Chris Starmer, and Robert Sugden**, “The Nature of Salience: An Experimental Investigation of Pure Coordination Games,” *American Economic Review*, 1994, 84 (3), 658–673.
- Merton, Robert K**, “Social Structure and Anomie,” *American Sociological Review*, 1938, 3 (5), 672.
- Niederle, Muriel and Lise Vesterlund**, “Do women shy away from competition? Do men compete too much?,” 2007.
- Prediger, Sebastian, Björn Vollan, and Benedikt Herrmann**, “Resource scarcity and antisocial behavior,” *Journal of Public Economics*, 2014, 119, 1–9.
- Rand, David G**, “Cooperation, Fast and Slow: Meta-Analytic Evidence for a Theory of Social Heuristics and Self-Interested Deliberation.,” *Psychological science*, 2016, 27 (9), 1192–1206.
- Rodrik, Dani**, “Why Does Globalization Fuel Populism? Economics, Culture, and the Rise of Right-Wing Populism,” *Annual Review of Economics*, 2021, 13 (1), 133–170.
- Roos, Patrick, Michele Gelfand, Dana Nau, and Janetta Lun**, “Societal threat and cultural variation in the strength of social norms: An evolutionary basis,” *Organizational Behavior and Human Decision Processes*, 2015, 129, 14–23.
- Smith, Adam**, *The Theory of Moral Sentiments*, London: Penguin, 1759.

- Starmer, Chris**, “Developments in nonexpected-utility theory: The hunt for a descriptive theory of choice under risk,” *Journal of Economic Literature*, 2000, 38 (2), 332–382.
- Szekely, Aron, Francesca Lipari, Alberto Antonioni, Mario Paolucci, Angel Sánchez, Luca Tummolini, and Giulia Andrighetto**, “Evidence from a long-term experiment that collective risks change social norms and promote cooperation,” *Nature Communications*, 2021, 12 (1), 1–7.