CESifo WORKING PAPERS

7568
2019

Original Version: March 2019
This Version: July 2019

Partial Norms

Giovanna D'Adda, Martin Dufwenberg, Francesco Passarelli, Guido Tabellini

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7568*
Category 13: Behavioural Economics

Partial norms

Abstract

We consider an expanded notion of social norms that renders them *belief-dependent* and *partial*, formulate a series of related testable predictions, and design an experiment based on a variant of the dictator game that tests for empirical relevance. Main results: Normative beliefs influence generosity, as predicted. Degree of partiality leads to more dispersion in giving behavior, as predicted.

JEL-Codes: C910, D910.

Keywords: social norms, partial norms, normative expectations, consensus, experiment.

Giovanna d'Adda
Università degli Studi di Milano
Department of Economics, Management and Quantitative Methods (DEMM)
Milan / Italy
giovanna.dadda@unimi.it

Martin Dufwenberg
University of Arizona
Department of Economics
1130 E. Helen Street
USA - 85721-0108 Tucson AZ
martind@eller.arizona.edu

Francesco Passarelli
Università degli Studi di Torino
Department of Economics ESOMAS
Corso Unione Sovietica, 218
Italy - 10134 Torino
francesco.passarelli@unito.it

Guido Tabellini
Universitá Bocconi
IGIER
Via Roentgen 1
Italy - 20136 Milano
guido.tabellini@unibocconi.it

July 22, 2019
We are grateful to PRIN grant 2015PJL3XP for financial support, and to Giacomo Lemoli, Francesca Miserocchi, and Tommaso Cohen for excellent research assistance.

1. Introduction

Despite a very large literature, we still do not know much about how and why social norms influence individual behavior, and how social norms change over time. The reason for this lack of understanding is probably twofold. On the one hand, there is not a single and generally agreed upon conceptual framework to study social norms. On the other hand, measurement of social norms is difficult and available evidence is not easy to interpret.

A clear definition of a social norm is provided by Fehr and Schurtenberger (2018), namely a commonly known standard of behavior that is based on widely shared views of how individual group members ought to behave in a given situation (see also Elster 1989, Bicchieri 2005). Thus, a social norm establishes a normative standard of behavior, which is not only based on the expectation of how others actually behave, but also on how others think one ought to behave. Moreover, this standard of behavior is commonly known and widely shared. In other words, social norms entail an aspect of conformity with the normative standards of others. In this sense, social norms are distinct from individual values or preferences, that instead do not logically require to conform to the expectations of others.

In this paper we adopt this definition of a social norm, and ask two empirical questions. First, consider an exogenous change in the perceived contents of a social norm, defined as a change in individual perceptions of the normative standards of others. Does this change influence individual behavior? Second, is the influence of the norm on individual behavior affected by the degree of perceived consensus around the contents of the norm?

Knowing the answer to these questions is important to understand the mechanisms through which social norms influence behavior, and also to explain changes in social norms. The first question is a check on the relevance of the conceptual framework suggested by Fehr and Schurtenberger (2018) and others. How important are the normative standards of others in shaping individual behavior? If they are found to be important, then explaining changes in social norms amounts to explaining changes in such normative standards. The second question is perhaps even more important. The ideal of a widely shared social norm is logically clear, but it does not often correspond to real social situations. Often there is not unanimous consensus on the prescriptive content of a social norm. Moreover, not all social norms are equally influential. Is the norm less powerful if it is perceived as vaguer or less widely shared?

To guide our empirical analysis, we formulate a simple theoretical model in which individuals trade off their material interest against their normative concerns. Individuals disagree on what is the right course of action, but this disagreement is common knowledge and they dislike deviating from the average normative standards of others (the social norm). The influence of the social norm is stronger if it is more widely shared. The model yields two predictions. First, individual behavior moves in the direction of the perceived social norm. Second, the perceived normative consensus also matters, in subtle ways. A stronger consensus reduces the dispersion in individual actions, and influences behavior in a direction that depends on the contrast between subjective values and the perceived social norm.

We then explore these predictions in a version of the dictator game, where the amount given to the other party is doubled by the experimenter (i.e., if the dictator gives 3 Euros out of 10 to the receiver, then the dictator retains 7 Euros while the receiver gets 6 Euros). This doubling of the amount given adds some ambiguity over the contents of the social norm, since in a simple dictator game the equal split (5 and 5) is a natural focal point. We change perceptions over the normative standards of others through an informational treatment. Before playing the game and before knowing whether they will be dictators or receivers, participants in the experiment are told that we previously asked other individuals what the socially appropriate behavior of the dictator in such a game would be. Different players are shown different distribution of answers to this question. We compare a default distribution of answers with two informational treatments: First, a distribution that differs from the default in that it has a lower average amount given as the right thing to do, but the same variance – we call this the Low Average treatment. Second, a distribution that differs from the default in that it has a higher variance of answers (on what is the right thing to do), but has the same mean as the default distribution – we call this the High Variance treatment.

We obtain two robust results, consistent with the predictions of the model. First, individuals who received the Low Average informational treatment give less as dictators, compared to individuals exposed to the default distribution. Second, individuals who received the High Variance treatment on average do not donate a different amount than the default group, but exhibit a higher variance in the amount given, as a group.

We also verified that the informational treatments change (incentivized) beliefs in the expected direction. We elicited three set of beliefs: 1. on the normative standards of others (i.e. what others regard as socially appropriate); 2. on positive expectations of others (i.e. what others

will actually do); 3. on values (i.e. opinions on what is the right-thing-to-do, as opposed as to what others regard the right-thing-to-do). We use a novel elicitation method for 1 and 2, which allows us to observe the full distribution of normative and positive expectations over the entire range of possible actions. All the beliefs that we elicit are affected by informational treatments in the expected direction, although the effect is strongest on the beliefs over the normative standards of others. Namely, individuals exposed to the Low Average treatment perceive others as having the normative expectation of a lower amount given, while those exposed to the High Variance treatment have a more dispersed perceived distribution of the normative expectation of others.[3] Moreover, adapting to our setting the methodology of Krupka and Weber (2013) of estimating a conditional logit model, we show that these beliefs are correlated with actions in the expected manner.

Overall, therefore, these results confirm that social norms influence actions through individual perceptions of the normative standards of others, and that such beliefs react to available information. Moreover, the consensus around the social norm, measured by the variance of perceived normative beliefs, also matters. A more partial norm (i.e. one on which consensus is weaker) is associated with more dispersion in individual actions, suggesting that the social norm is less influential.

Sections 2-7 contain, respectively, a discussion of related literature, theory, experimental design, implementation, results, and a discussion.

2. Related Literature

Our paper contributes to the vast literature studying how social norms shape the behavior of individuals (see Fehr and Schurtenberger 2018 for an excellent review). Specifically, we investigate how behavior is influenced by a change in the perception of the social norm or by a change in the consensus around the norm. We do it in the laboratory, identifying the normative standards of individuals and generating exogenous variations through informational treatments. Other recent papers have addressed similar questions or adopted analogous methodology. Below we explain how we differ from some of them.

[3] As explained below, to avoid introducing a confounding effect on actions, we elicited beliefs and studied how beliefs are affected by informational treatments in a different experiment and on a different set of players.

The idea that individuals tend to conform to the normative expectation of others is familiar in sociology and social psychology. The psychological foundation of such behavior is related to the notion of social identity. Identification with a social group entails an element of depersonalization. The individual slants his positive and normative beliefs towards those that are distinctive of the group with which it identifies, because this is how a typical group member behaves (Hogg and Abrams 1998). In our experimental setting there is no group conflict, so that implicitly all players share a common identity, but the influence of the social norm could reflect the same mechanism.

Krupka and Weber (KW-2013) propose an ingenious method to identify social norms in the laboratory. They elicit a profile of appropriateness ratings over all actions available to the decision-maker, which they call "social norm". The elicitation technique consists in a coordination game in which participants are incentivized to guess the modal rating of each action. People are not directly asked their opinions about the appropriateness of an action. They are asked to guess modal appropriateness ratings. The authors point out that this method might lead participants to coordinate over ratings that do not reflect their true opinions, but if telling the truth is focal, doing so would make for a natural equilibrium.

We differ from KW in three respects. First, we allow for norms to be partial, reflecting incomplete consensus. A partial social norm allows for a non-degenerate distribution of individuals' opinions about the right thing to do. Second, KW's method of using a coordination game is not amenable to measuring the degree of partiality of a norm, so we use a different technique. We ask participants to guess the actual distribution of individual values, and we pay them if their guesses are correct. Thus, their answers should reflect their true guesses about others' normative opinions. Third, KW's treatment variation consists in changing some contextual features of the dictator game. Our treatment consists in providing different information about the normative beliefs of others.

Bursztyn, Gonzalez and Yanagizawa-Drott (BGY-2018) use an informational treatment to study norms about female labor force participation in Saudi Arabia (FLFP). On average participants (all men) underestimate the share of people approving FLFP. The treated group is provided information about the true approval share, while the control group is provided no information. Once beliefs are corrected, the members of the treated group are more likely to let their wives join the labor force. There are two substantial differences between our paper and BGY.

First, men's actions in BGY are publicly observable. In our paper, dictators make totally anonymous choices. This means that BGY studies the effect of social norms on individuals' social-image, while we study the possible impact of social norms on self-image management. Second, in BGY only the treated group receives information. The effect of "correcting beliefs" is actually determined by the composition of two different factors: a) providing information *per se*; b) the specific content of information. BGY's design does not allow disentangling a) from b).[4] In our design, all groups receive information about the social norm. The treatment consists in providing different information to different groups. Thus, effects are driven only by b), so we do not have to worry about possible interactions with a).

Bicchieri and Xiao (BX-2009) study the effect of information regarding what the majority of other people actually does ("empirical" social norm) and what the majority thinks ought be done ("normative" social norm). When the two types of information are in conflict, only the former type is found to influence behavior. BX are the first to propose an incentivized method to elicit individuals' expectations about the empirical and normative social norm. Differently from us, BX only provide information on what a given majority thinks or does. Thus, their setting is not suitable to study partial consensus regarding the norm. Bicchieri and Dimant (2018) explore self-serving manipulation of social norms, as defined by BX, to justify selfish behavior. Their treatment variation consists in changing the beneficiary of norm-violation (self vs. other), or changing the type of social norm (normative vs. empirical), or changing participants' information about the upcoming opportunity to violate the norm.

Tankard and Paluck (2016) show that people use other individuals' public behavior, summary information about a group, and institutional signals to shape their own perception of social norms. Changes in each of these sources of information can influence norm perception and related behavior (see Tankard and Paluck 2016 for a literature review). Byrne et al. (2018) show the key role of beliefs on average behavior in determining the sign of the effect of social information.[5] Danilov et al (2018) focus on the difference between aversion to violate others'

[4] In a paper studying protest participation, Cantoni, Yang, Yuchtman, and Zhang (CYYZ-2018) adopt an informational treatment that is similar to BGY's, but also shares some similarities with our experiment. They elicit individuals' priors about participation rates. Then they provide a random group truthful information about planned participation rates. Thus, as we do in one of our experiments (cf. experiment B&A in section 4.2), they elicit posteriors and observe actions. Note that also CYYZ's design does not allow disentangling a) from b).

[5] Our paper is also related to a large literature in economics demonstrating how information about what others do or think can influence behavior in a variety of settings, from energy and water consumption (Allcott 2011;

expectations (i.e., guilt aversion) and aversion to act differently than others (i.e., compliance with descriptive social norms). Information on descriptive norms monotonically affects behavior in a dictator game, while too high expectations may even lead to a reduction of donations.

3. Theory

We next (i) discuss norms & conformity, (ii) introduce partiality, (iii) restrict attention to one setting for which we specify a formal model and (iv) state precise predictions to be tested.

(i) *Values, conformity and social norms*

Despite a very large literature, there isn't a single unambiguous and commonly accepted definition of social norms. Thus, we start by clarifying what we mean by social norm. In principle, deviations of individual behavior from narrowly defined self-interest can be motivated by three sets of variables: (i) *Values*, namely personal conceptions of what is the right thing to do in a specific circumstance, such as internalized moral values. (ii) *Conformity*, namely a desire to follow the actual behavior of others, such as with fashions. (iii) *Compliance with social norms*, that we define as a desire to comply with the perceived values of others. In other words, we reserve the term "norm" for something normative, prescribing what one "should" do. In this sense, social norms differ from fashions, that are purely motivated by a desire to conform, but with no normative content. Many scholars call also such behavior norm-driven, but we do not. Social norms also differ from individual values in the sense that adhering to a norm is the *socially* "right-thing-to-do." In other words, social norms also entail an aspect of conformity, that is conceptually absent from pure individual values. However, social norms and individual values are related because a certain behavior is perceived as a social norm only if there is sufficient consensus that it is the right thing to do.

Ayres et al. 2013; Ferraro et al. 2011; Ferraro and Price 2013), contributions to charitable causes (Frey and Meier 2004; Shang and Croson 2009), voting (Gerber and Rogers 2009), and financial decisions (Beshears et al. 2015). Role models, such as those portrayed by the media or political leaders, have also been shown to have an impact on norm perception and norm-related behavior concerning, for instance, women's status (Beaman et al. 2012; Jensen and Oster 2009), fertility (La Ferrara et al. 2012), or dissent and cooperation (Paluck 2009). Similarly, laboratory experiments show how normative information conveyed by leaders' actions (d'Adda et al. 2017; Drouvelis and Nosenzo 2013) or by the institutional environment (Peysakhovich and Rand 2015) shapes participants' behavior.

In practice it may be difficult to disentangle these three motivations of individual behavior. Moreover, there may be interaction effects between preferences for adhering to norms, preferences for conformity and values. Nevertheless, for our purposes it is useful to keep them conceptually distinct.

Departing from a social norm entails an element of disappointing the expectations of others, and we explore the idea that decision makers are averse to doing so. In this regard, the motivation we look at resembles guilt aversion (see Battigalli and Dufwenberg 2007 for a general model), a belief-dependent sentiment the modeling of which requires the framework of psychological game theory (Geanakoplos et al. 1989; Battigalli and Dufwenberg 2009). However, we consider expectations regarding how one "ought to behave", not how one will actually behave, which marks a way that our approach is not formally captured by psychological games as formulated in the papers we cited.

(ii) *Partial norms*

Definitions of norms tend to require consensus. Most accounts require that if *r* (for "right-thing-to-do") is a norm then everyone's *r* is the same, everyone believes everyone's *r* is the same, … et cetera ad infinitum. Call this an *ideal* norm.

Beyond philosophical discussion, that concept is implausible. Hardly anything qualifies. Consider tipping in the US, a norm which influences behavior. But is there an ideal norm *r*, in which there is common belief? No. Some people say tipping 15% is a norm, but some say 15-25%, and some add riders as regards how the norm is conditional on "good service." There is no one *r*, and no common belief in any *r*. However, approximate common belief in some (small) set of *r*-values may be good enough to meaningfully talk about a norm. This is what we have in mind when we speak of a *partial* norm.[6]

Once one allows for partial norms it becomes natural to quantify degree of partiality and to explore its behavioral implications. It then seems natural to conjecture that the closer a partial norm is to an ideal one, the stronger is its normative pull on an individual's choice.

[6] In principle, a norm can be partial for two reasons: incomplete social consensus, or subjective doubts about the contents of the norm, irrespective of social consensus. We focus on incomplete social consensus while noting that this involves an abstraction.

(iii) *Our setting*

We consider the following version of a dictator game: There are two players. Player A is given 10€ and must indicate how much to give to player B, with the understanding that whatever is given will be value-doubled: If A gives x then B gets $2x$ while A gets $10-x$.

This setting has several virtues: First, it is simple. Second, while in principle norms may prescribe behavior in complicated situations, having multiple active players makes it difficult to infer how norms affect behavior.[7] Suppose a norm (somehow) changes, and suppose player i's behavior changes. Is this due to the shift in norm, or is it because i believes others' behavior changed? It is hard to tell. A dictator game avoids this problem.

Third, as mentioned earlier, the value-doubling feature (player B gets $2x$, not x) introduces some ambiguity over the content of the norm (without doubling the amount given, a 50/50 split would be an obvious norm): A norm of maximizing the total surplus from the experiment would lead to a norm of x=10. A norm of minimizing inequality in final payoffs would lead to a norm of x=3. And if the norm is that A gets to keep half his endowments, then it would imply that the norm is x=5. We exploit this ambiguity to explore how actions and beliefs react to informational treatments (we return to this point in section 4).

We assume that the player trades off two concerns. On the one hand, his material incentives are such that other things being equal he prefers to keep as much as possible for himself. However, other things being equal he would also like to behave in a way that is normatively appropriate. If those two goals are incompatible, he has to find a personally optimal choice, which strikes a balance.

What do we mean by "normatively appropriate"? We provide a full answer below, where we also consider norm-partiality as well as a concern for conformity. However, before we go there, let us first consider a benchmark: the case where everyone agrees what the norms says is the right thing to do. That is, all individuals have the same personal values which then coincide with the social norm. Moreover, there is no separate concern for conformity. In this case we say that there

[7] As the example indicates, in general games a norm may need to specify not just a choice but arguably a full strategy. In fact, since a norm may regulate the behavior of many, presumably it should specify a full strategy profile. Few scholars seem to have attempted to tackle the task of developing a theory of social norms that is applicable to any game; a noticeable exception is López-Pérez (2008).

is an *ideal* social norm, which is a number $N \in \{0,1,\ldots,10\}$. Each individual then donates the amount x that minimizes the following quadratic loss function:[8]

$$W = x + (N - x)^2/2\theta \quad (1)$$

The first term captures the player's material incentives, the second term captures his normative concern: to stay as close as possible to the ideal social norm. The parameter $\theta > 0$ reflects the steepness of this tradeoff. If θ tends to infinity we get selfishness as a limiting case, and the lower is θ the more the player cares about following the norm.

[[{}

Taking the first order condition of the above minimization problem, one sees that the optimal amount given is the number in $\{0,1,\ldots,10\}$ which is closest to:

$$x = Max\{0, N - \theta\} \quad (2)$$

Trading off the two concerns that motivate him the player gives less than N, and how much less depends on the parameter θ.[9]

Let us now introduce the additional features that are central to our approach. Let $r \in \{0,1,\ldots,10\}$ be a player's view of the "right-thing-to-do" – what above we called *values*. In our analysis r is a primitive notion, underlying much of our analysis. Let $E(r)$ be a subject's expectation of everyone's r – our notion of a *social norm*. Let $V(r)$ be a subject's variance in beliefs of everyone's r – our notion of *partiality* of a social norm. Also actual choices and corresponding beliefs will be important, for modelling players' concerns for conformity. As before, let $x \in \{0,1,\ldots,10\}$ be a player's choice of how much to give. Let $E(x)$ be a subject's expectation of how much others give. Let $V(x)$ be a subject's variance of how much others give. Sometimes we will refer to $E(r)$ and $E(x)$ as a subject's normative and positive expectations, respectively.

---

[8] The choice of loss function is non-obvious and substantial. See Michaeli and Spiro (MS-2015) for a critical discussion analyzing loss of deviating from a social norm in terms of "curvature of social pressure" in different "societies." They argue that "strict societies are those emphasizing full adherence to the social norm, and hence they utilize concave social pressure; liberal societies are those allowing freedom of expression as long as it is not too extreme, and hence they utilize convex social pressure" (pp. 51-52.) On the presumption that Italy is a "liberal society," MS arguments help justify our specification (1).

[9] Note that the typical choice involves some shading relative to N, a typical feature in "liberal societies" following the thoughts of MS (see footnote 8). This marks a contrast to MS's "strict societies,'' and also to models where agents care about their social image and (in equilibrium, for signaling reasons) end up conforming to each others' choices (see Bernheim 1994, Bénabou and Tirole 2006, Andreoni and Bernheim 2009) or to each others' opinions (see Burstzyn et al. 2019).

Again, we think of a player as trading off material interests and normative concerns. However, from now on *N*, which was previously the same given number for all, should be seen, for each individual, as the value of a function, and different values of the arguments may be plugged in for different individuals. More precisely, *N* reflects the three forces we have previously hinted at: values, social norms and conformity. Specifically, we assume that *N* is a weighted average of three variables:

$$N = r + \alpha[E(r) - r] + \beta[E(x) - r] \quad (3)$$

where $\alpha,\beta > 0$ and $\alpha+\beta < 1$. This implies that α, β, and 1-α-β are, respectively, the relative weights given to the social norm, $E(r)$, to the expectations of others, $E(x)$ (capturing a conformity concern), and to individual values, r. We have, however, presented *N* in the (equivalent) form given by (3) in order to emphasize a comparative static that will be put to crucial use below: If $E(r) > (<)\ r$, then *N* is increasing (decreasing) in α, the relative weight on the social norm *E(r)*.

For simplicity of exposition, we assume throughout that players differ only in their values, i.e. they have different realization of the random variable *r* (and hence *N*) but they have the same preferences and the same parameters α, β, θ. Note that if α tends to 1 and if $V(r) = 0$, then for all individuals it holds that $r = E(r)$ and their *N*'s would be the same.

Under rational expectations and common knowledge:

$$E(N) = E(r)(1 - \beta) + \beta(E(x)) \quad (4)$$

We solve for the full rational expectations equilibrium under the assumption that the non-negativity constraint on *x* is never binding, for all realizations of *r* and for any value of $E(x) \geq 0$. This assumption is satisfied if the weight β on the expectations of others is not too large, specifically if

$$\beta < Min\{1 - \frac{\theta}{E(r)}, \frac{(1-\alpha)\underline{r} + \alpha E(r) - \theta}{\underline{r}}\}$$

where $\underline{r} < E(r)$ denotes the smallest possible realization of the random variable *r*. Under this condition and under common beliefs, the equilibrium is unique and by (1),

$$E(x) = E(N) - \theta > 0 \quad (5)$$

The full rational expectations equilibrium is obtained by solving the system of linear equations (2)-(5). The equilibrium amount given by each player, as a function of his values and of the social norm is[10]:

$$x = r + (\alpha + \beta)[E(r) - r] - \frac{\theta}{1-\beta} \quad (6)$$

We thus have that in equilibrium the amount given by each player is an increasing function of his expectations of the normative beliefs of other players, $E(r)$, and of his own values, r. Moreover, if individuals place more weight on the expected normative standards of others (i.e. if α increases), then the amount given rises (falls) if $E(r) - r > 0$ (< 0). Intuitively, if $E(r) - r > 0$ then the individual expects others to be more generous than he is. Putting more weight on the normative standards of others induces him to give a greater amount, but the opposite happens if he perceives others to be less generous than he is. Similarly, a higher weight β on the expected actual behavior of others has an ambiguous effect on the amount given, that depends on the sign of $E(r) - r$ but also on the size of θ (the weight given to selfish utility as opposed to normative concerns).

We re-emphasize that the only source of variation across individuals is in their values, r. Specifically, the variance in the amounts donated across individuals is $V(x) = (1 - \alpha - \beta)^2 V(r)$ where $V(r)$ denotes the variance of r. Thus, putting more weight on normative and/or positive expectations (i.e. increasing α and/or β) also reduces the variance $V(x)$ of the amounts given by different players. Intuitively, higher weights on expectations reduce the weight on the only idiosyncratic variable that varies across individuals.

Finally, we address the key issue of partiality of norms. Although we have written α and β as parameters, they are likely to reflect other features of the environment. In particular, it is reasonable to conjecture that the size of α is also affected by how much consensus there is around the value system captured by the random variable r. The more consensus there is, the more relevant is the desire to conform to what others regard as good behavior, and the higher is the relevant weight α. In the limit, if everyone shares the same value system, then there is an ideal norm equal to $E(r)$. As hinted at before, in this case $E(r)=r$ since the player himself is included in "everyone." Individual values and social norms coincide. If instead individuals have very different value

[10] If the condition on β is not satisfied, then a rational expectations equilibrium still exists and also admits $x = 0$, although it need not be unique and to obtain a closed form solution we need to impose a specific functional form on the distribution of r.

systems, then the subject would no longer believe that an ideal norm existed. Rather, the norm would be partial, still equal to $E(r)$, but less likely to influence behavior. Thus, the relative weight α would be smaller. More formally, we posit that the weight α on the social norm is a decreasing function of *V(r)*, the variance of the normative standards of others. As argued above, therefore, a higher variance *V(r)* increases the dispersion in the amount given, *V(x)*, both directly and indirectly, and increases (decreases) the amount given if $E(r) - r > 0$ (< 0).[11]

The result that a more partial (and hence less influential) social norm does not always lead to more selfish behavior may seem puzzling. The intuition for this result is that altruistic behavior is not only driven by the perceived social norm, but also by subjective values (the variable *r*). If the norm is more partial, and hence less influential, behavior is driven to a larger extent by individual values, that can be more or less demanding than the social norm. Hence, a weaker social norm certainly leads to less conformity in behavior, but its effect on donations is heterogeneous and depends on whether the individual is more or less altruistic than the average.

(iv) *Predictions*

Summing up the analysis of section 3(iii), we suggest that individual behavior can be influenced by five distinct dimensions:

- a value-per-se dimension: individual beliefs of what is the right thing to do, i.e. *r*;
- a norm-compliance dimension: individual beliefs about others' *r*'s, i.e. $E(r)$;
- a conformity dimension: individual beliefs about *x*, i.e. $E(x)$;
- the degree of consensus around the value system, reflected in the weights α and β.

We make the following key predictions:

- The amount donated, *x*, is increasing in the first three dimensions, *r*, $E(r)$ and $E(x)$
- Less consensus on the right thing to do, measured by a higher variance *V(r)*:
  - (i) Increases the variability across individuals of the amounts donated
  - (ii) Decreases (increases) the amount given if $E(r) - r > 0$ (< 0).

Finally, norms may influence not only behavior but also beliefs. In particular, the higher is $E(r)$ the higher is $E(x)$ and/or the higher is *r*, *ceteris paribus*.

[11] This formalization assumes that the tradeoff between selfish concerns and social concerns is not affected by consensus; i.e., parameter θ does not depend of *V(r)*. If it did and the relationship was negative, there would be an additional implication that individuals are more selfish if *V(r)* is higher.

In the next section we describe the experiment that we have designed in order to test these predictions. Many of them have design counterparts that create relevant *exogenous variation*. Where the theory calls for $E(r)$ or degree of consensus to influence x, our design includes some feature that induces exogenous (by treatment) variation in $E(r)$ or in the degree of consensus.

4. Experimental design & procedures

Our design adopts the dictator game described in section 3(iii). We randomly form pairs of subjects, one of which is randomly assigned the role of "Individual A" while the other one plays in the role of "Individual B". The former is the Dictator (a term that we never use in the instructions). He/she is given €10 and has to choose how much to transfer to individual B, knowing that the transfer will be doubled.

As pointed out earlier, it is not obvious what the right thing to do is in such game. The possibility of alternative normative benchmarks represents a virtue of this design, since it should yield a substantial degree of variation in individuals' opinions, r, of what ought to be done.[12]

Our design achieves exogenous variation in individuals' beliefs through three *Information Disclosure Treatments*. Information is about what other people think of the right-think-to-do. Since they may have different opinions, and their opinions might occur with different frequency, information is disclosed in the form of a distribution (see below). The three informational treatments consisted in showing the following three different distributions before subjects make any choice:

1. **Baseline**. The baseline distribution is in the left-hand diagram of Figure 1: its mean is equal to 4.32, its mode is equal to 5 and its variance is equal to 0.78.
2. **Low Average**. Subjects are shown the distribution in the center diagram of Figure 1. It has the same variance as the Baseline distribution, but a lower mean, equal to 0.63.

[12] We conducted a preliminary test with 400 subjects, recruited on the online platform Prolific Academic to test the distribution of r in different variants of the dictator game (DG): a standard DG, a DG where the recipient gets $2x$ as described in section 3(iii), a DG such that if player A gives x then player B gets $0.5x$, and a DG with a taking option. The results from this test indicated that the DG of section 3(iii) generated the largest variance in individuals' opinions. Full results are available upon request.

3. **High Variance**. We show subjects the distribution in the right-hand diagram in Figure 1. It has the same mean as the Baseline distribution but the variance is significantly larger.[13]

Figure 1. Treatment distributions

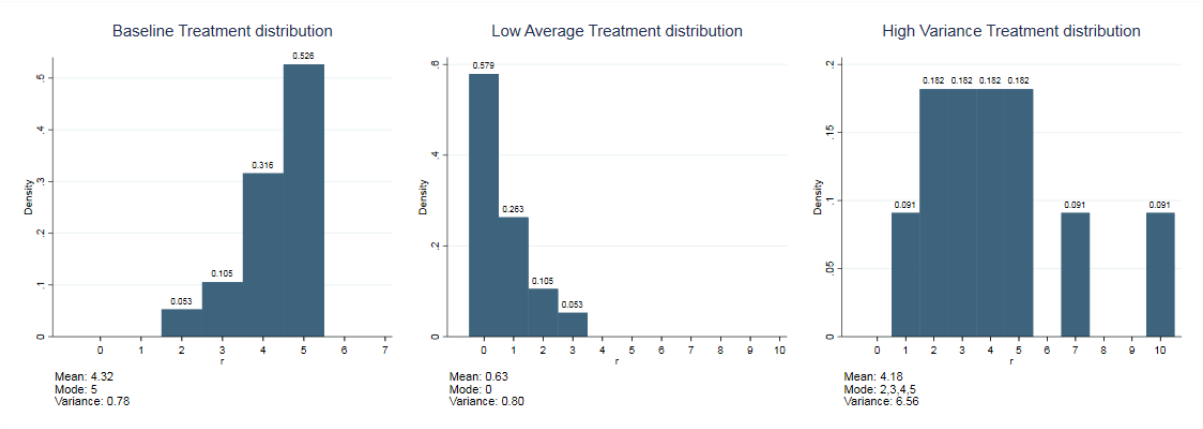


Table 1 summarizes the average, the mode, and the variance of each distribution.

Table 1. Treatment distributions

Treatment	Mean	Mode	Variance
Baseline	4.32[a]	5	0.78[b]
Low Average	0.63	0	0.80[b]
High Variance	4.18[a]	2,3,4,5	6.56

[a] The means of Baseline and High Variance distributions are not significantly different from each other (p-value: 0.87)
[b] The variances of Baseline and Low Average distributions are not significantly different from each other (p-value: 0.96)

As we disclose information we explicitly tell participants that we collected other subjects' opinions in previous sessions of a similar experiment. Therefore, although participants are unaware of the specific selection criteria we applied in generating the distributions, there is no deception involved in our experiment. Besides the above differences in the distributions, the three treatments

[13] We also ran one other informational treatment with a milder contrast to the Baseline. Specifically, we presented subjects with a distribution such that the Low Average Treatment had an average value for $E(r)$ = 3.23, which was significantly lower than that in the Baseline treatment. The variance was $V(r)$ = 0.71, not statistically different from that in the Baseline treatment.

are equal in all other respects. Once individuals have been shown a distribution, we let them make their choices.

We run two different experiments: an experiment with Actions (hereafter, Experiment A), and one with Belief elicitation & Actions (hereafter, Experiment B&A).[14] The rest of this section explains and motivates.

4.1 Experiment A

In Experiment A, dictators choose their donations immediately after information disclosure. This design cleanly and directly tests how information about the social norm affects actions. A higher amount donated in the Baseline treatment than in the Low Average treatment would imply that subjects' generosity is *causally* affected by the generosity of the social norm. Thus, the comparison between the amounts donated in the two treatments represents a clean test of the prediction that the amount donated, x, is increasing in $E(r)$.

Moreover, a higher dispersion of the amounts given in the High Variance treatment compared to Baseline would support the theoretical prediction that norm partiality also influences behavior. However, based on data gathered in Experiment A, we cannot test the more precise prediction that the effect of $V(r)$ on the amount given by each individual depends on the contrast between his own values r and the social norm $E(r)$, because we do not observe these individual beliefs. We can do it instead using beliefs elicited in Experiment B&A, although this raises other issues discussed below.

The identification of treatment effects on actions comes at a cost, as this design only tests a reduced form of our theory of norm compliance and conformity. The design of Experiment A enables us to observe whether $E(r)$ or $V(r)$ affect behavior, but we cannot pin down the channel. For instance, it is plausible that x is affected by $E(r)$ directly, or it is affected indirectly, because a change in $E(r)$ may lead to a change in $E(x)$ or in r. As for $V(r)$, there might be multiple channels as well. A higher $V(r)$ may plausibly lower the desire to comply with $E(r)$ – a lower value of parameter α, as postulated in our theoretical model. It may also lower the desire to conform with $E(x)$ – as captured by a lower β. Moreover, we cannot exclude that a change in $V(r)$ may also lead

[14] Experimental instructions are reported in Appendix B.

to a change in $E(r)$ or $E(x)$ and then affect behavior through a change in subjects' beliefs about others.

4.2 Experiment B&A

In order to shed more light on the channel through which treatments affect behavior, in Experiment B&A we prompt the beliefs of all participants (see next section for the incentivizing scheme). We do it before participants are revealed their role in the game. Then, we assign roles and ask dictators to make their donations. Our treatments provide subjects with different information about others' opinions of the right-thing-to-do, *r*. Thus, experiment B&A allows us to study whether subjects' beliefs are *causally* affected by the social norm.[15]

A decrease in subjects' perception of what is right-thing-to-do, *r*, in the Low Average treatment would support the hypothesis that people think they should be less generous if more people think the same. In other words, *r* is positively affected by $E(r)$. We can also test the hypothesis that positive beliefs, $E(x)$, are positively affected by $E(r)$. In this case the Low Average treatment should induce a drop in $E(x)$.

The treatment High Variance is intended to exogenously increase a subject's beliefs about the *partiality* of a norm. We can test whether this treatment induces higher dispersion in normative or positive beliefs, $V(r)$ and $V(x)$, respectively.

The design of Experiment B&A could reveal the mechanism through which norms affect individuals' donations, by showing how both actions and beliefs respond to the information treatments. In practice, however, we may worry about our ability to detect changes in dictators' actions through this design for a number of reasons. First, belief elicitation induces players to focus their mind on the subject matters asked in the questions, and in particular to reason on what the-right-thing-to-do is (before knowing whether they will act as dictator or recipient). Second, the design introduces a long time span between the informational treatment and the choice of actions. Third, belief elicitation requires a considerable amount of cognitive effort and causes cognitive fatigue, which might independently affect later choices. These features of the design may dampen

[15] In Experiment B&A we elicit three different beliefs: a) a subject's belief about the right-thing-to-do, *r*; b) his/her beliefs about others' beliefs; c) his/her beliefs about the amount donated by others. Note that beliefs in b) and c) are elicited in the form of two distributions. For each of the eleven possible values of *r* and *x* we ask subjects to guess how many participants choose each of those values. Thus subjects have 22 numbers to guess.

the effect of the information on dictators' actions. In particular, reasoning in the abstract on "what is the right thing to do" and on how others respond to this question may lead participants to give more weight to the intrinsic merit of the alternatives faced by the dictator. This could lead them to act based on moral or value criteria (i.e. consistently with what they said is "the right thing to do"), discounting social conventions and hence reducing the effect of information about how others perceive the social norm. For these reasons, in what follows we use Experiment A to test our predictions on the treatment effects on actual behavior, and we use Experiment B&A only to assess the treatment effects on beliefs.

5. Implementation

The sessions were conducted at BELSS *(Bocconi Experimental Laboratory in Social Sciences)* in Milan, during the period September 2017-February 2018. We recruited 686 Bocconi students (or exchange students at Bocconi).[16] No subject was recruited more than once. We ran 15 sessions of Experiment B&A (3 treatments; 5 sessions per treatment), and 16 sessions (=5+5+6) of Experiment A.[17] The average number of participants in a session was 19. Sessions lasted on average an hour for Experiment B&A and 30 minutes for Experiment A. The average payment was €9.54 including the show-up fee.

Table 2 summarizes the experimental conditions and reports the number of subjects and sessions in each condition.

[16] Subjects were recruited from the Laboratory's sign-up list using ORSEE (Greiner, 2015).

[17] We also ran 5 sessions of a Pilot Experiment, that we describe in section 5.3 below. We did it in order to generate data to run the Baseline informational treatment.

Table 2. Experimental conditions and number of subjects per condition

	Experiment B&A		Experiment A	
	Number of subjects	Number of sessions	Number of subjects	Number of sessions
Baseline	98	5	98	5
Low Average	96	5	92	5
High Variance	96	5	114	6

Note: the table reports the total number of subjects and sessions per condition. Within each cell, data on DG allocations are available only for the subjects assigned to the role of dictators, while we have data on beliefs for all subjects in the B&A treatments.

The experiment and payment protocols were designed to ensure the highest degree of anonymity and minimize the possibility that subjects' choices were driven by reputational concerns or experimenter demand effects. Upon arrival, *n* subjects entered the experimental laboratory one by one and were randomly assigned an isolated seat with a computer terminal. The number of participants in a session, *n*, was always an even number between 16 and 20.[18] Participants could read instructions on the computer screen. At the beginning of each session, an assistant read aloud the General Instructions (see below) and checked that participants correctly understood it. The experiment was conducted with real money.

5.1 Experiment A

Phase 1: General instructions

Subjects were informed about the number of participants *n*. We told them that another participant in the room had been randomly paired with them. Thus *n*/2 pairs had been formed. One subject in the pair would soon be randomly assigned the role of "Individual A" while the other one would be assigned the role of "Individual B". All subjects found a carton box on their desk. We asked them not to open it until instructed to do so. We informed them that the box of Individual A

[18] In order to reduce the risk of having too few people in a session, we recruited 22 people. If for instance 22 participants showed up, we randomly selected the 21th and the 22nd participant to exclude from the session. If say 17 people (an odd number) showed up we randomly chose the 17th to exclude. All excluded persons were paid an increased show-up fee of €5 and were allowed to sign up for another session in the future. In no case we had less than 16 people showing up in a session.

contained two small bubble cushioned envelopes, one white and one yellow. The former was marked "Money for Individual A", and it contained €10 in one-euro coins. The latter was marked "Money for Individual B", and it was empty. The subject playing in the role of Individual A would be able to choose how many coins x to transfer to Individual B, by simply putting them in the yellow envelope ($x \in \{0,1,\ldots,10\}$). Individual A would keep the remaining coins while the experimenters would take care of the transfer to Individual B. They would also match the transfer with additional money. Thus there was common understanding that, by the end of the experiment, individual B would receive $2x$. All participants were informed that, whoever Individual A was, no other person would ever know his/her identity and choice. Hereafter participants could read instructions on the computer screen.

Phase 2: Information disclosure

Before participants were assigned their role, the computer showed them a distribution of previous participants' opinions about the most appropriate donation amount.[19] Participants were informed that those opinions had been previously gathered in sessions similar to the one they were in. We had three treatments as described in section 4: Baseline - Low Average - High Variance (cf. Figure 1). This is when our information manipulation eventually kicked in.

Phase 3: Role assignment, actions and transfers

Participants could now read on screen the role they were randomly assigned in the pair. They could open the box. Those who were assigned the role of Individual B found it empty. Then they were instructed by the computer to remain silent. Those who played in the role of Individual A found the two envelopes. The computer instructed them to silently transfer their donation x from the white envelope to the yellow envelope for Individual B. They kept the remaining coins. Their actions could not be seen or heard by anyone. They were asked to leave the two envelopes in the box and record their choice on the computer. Then experimenters collected all of the boxes (also those of the receivers) and took them to another room, while all subjects remained seated and

[19] Specifically, participants were shown the distribution of other subjects' answers to the question: *In your opinion, which is the most socially appropriate action that Individual A should take* (cf. Figure 1). We made it clear that by "socially most appropriate" we meant behavior that they considered the "correct" or "ethical" thing to do.

silent. In the other room, yellow envelopes containing transfers were actually transferred from each A's box to his/her paired B's box. All of the boxes were returned to the participants. Receivers could then keep the amount received and check if it was the correct amount on the computer.[20] They were told that that amount would be doubled at the end of the experiment. This procedure ensured complete anonymity between subjects. No subject in the room could infer the identity of his/her paired subject or whether any other participant was an Individual A or an Individual B.

Phase 4: Payments

Participants could read on the screen the total amount they earned, including the participation fee. Then subjects were called one by one outside of the room by their seat number. They received an anonymous envelope with their seat number marked on it. The envelope contained the money they earned in addition to the coins they already got during the experiment. This procedure was designed to ensure the maximum degree of anonymity during the experiment, and to minimize the risk that dictators made their choices to please the experimenter.

5.2 Experiment B&A

Experiment B&A represents a "longer version" of Experiment A. The difference is in the fact that it includes an additional Phase 2a, in which we elicit beliefs, as follows. All other phases, are the same as in Experiment A.

Phase 2a: Belief elicitation

We ran Phases 1 and Phase 2, as in Experiment A. After disclosing information, we asked three questions. First, "Which is the most appropriate action that Individual A should take?" In other words, we asked subjects to tell us their "*r*'s". They could tick one of the eleven boxes in a table containing the eleven possible transfers. Second, we asked them to guess the distribution of the answers that the *n* participants in the same session gave to the first question. They had to guess eleven frequencies, which we incentivized by paying €0.2 for each right guess. Third, we asked

[20] At this point, we also asked each Receiver for her feelings with respect to the Dictator's allocation. Namely, Receivers could say by how much they would be willing to reduce Individual As' earnings: this decision was completely hypothetical and unincentivized. Receivers knew that their payment reduction decision would not be implemented, thus their expressions of approval or disapproval had no real consequence at all on Dictators.

them to guess the distribution of the actions that the *n/2* Individual A's would take in the same session. Also in this case they had to guess eleven frequencies, which we incentivized with €0.2 for each correct guess.

In synthesis, for each participant we elicited three different beliefs. First, his/her individual value, *r*. Second, his/her beliefs of the distribution of others' values. From that distribution we could compute each participant's *E(r)* and *V(r)*. Third, we elicited each participant's beliefs of the distribution of what dictators would actually do. We used this distribution to compute each participant's *E(x)* and *V(x)*. The remaining part to the experiment is the same as in Experiment A (see Phases 3-6 above).[21]

5.3 Pilot

We ran also a Pilot experiment to test the design of the Experiment B&A. We used the data on individuals' opinions about *r*, the right-thing-to-do to selectively build the distribution that we used in the Baseline treatment of both Experiment B&A and Experiment A. The incentivizing scheme and the payment procedures are the same as in Experiment B&A.

6. Main results

6.1. Treatment effect on actions

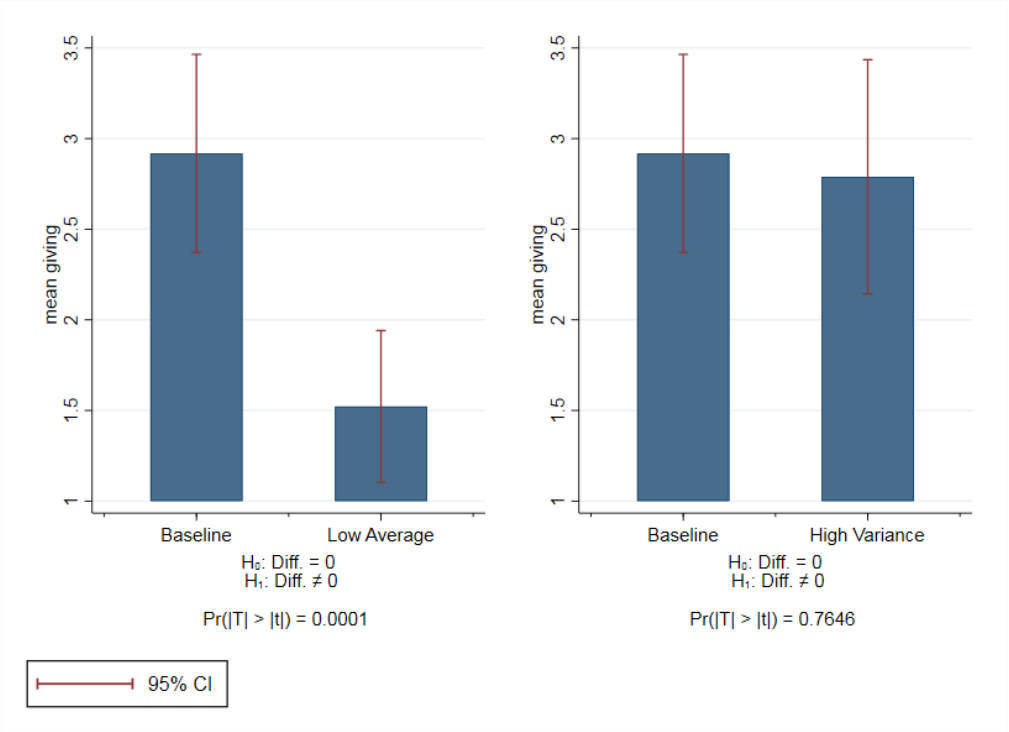
We begin by focusing on experiment A, and explore how actions are influenced by the informational treatments. We comment on experiment B&A in the next subsection. Table A1 in the appendix provides summary statistics of average dictator giving in Experiment A and B&A, by treatment (the overall amount given on average is € 2.29).

Figure 2 shows the average amounts given by the dictator, in a pairwise comparison, by treatment. As shown in the left hand panel of Figure 2, the difference in the average amount donated between the Baseline and the Low Average is about €1.4. This corresponds to almost 50% of the average donation in the Baseline treatment, and to about 40% of the difference in *E(r)* across

[21] In phase 4 of Experiment B&A (Payment), besides other payments, participants could read on the screen also rewards from guessing other participants' answers or behavior.

the two treatments (in the Baseline treatment participants are shown a distribution that has $E(r)$ = 4.36, while in the Low Average treatment they are shown a distribution with $E(r)$ = 0.63). As shown on the right hand panel of Figure 2, there is no significant difference in the amount donated between the Baseline and the High Variance treatments.

Figure 2. Dictator giving in Experiment A, mean-comparison test



In order to assess whether the treatments also affected the variability of donations, we turn to the distribution of dictators' choices. Figure 3 illustrates the histograms of the actions chosen by the dictator under the three treatments. The three distributions look different, particularly the first two. Compared to the Baseline, the distribution of actions in the Low Average treatment is shifted towards 0, while the distribution in the High Variance Treatment seems more spread out. Pearson's Chi-square distribution tests reveal that the distribution of giving in the Baseline treatment is significantly different from that in the Low Average treatments (p = 0.001), but not from the one of the High Variance treatment.

Figure 3. Distribution of dictator giving in Experiment A, by treatment

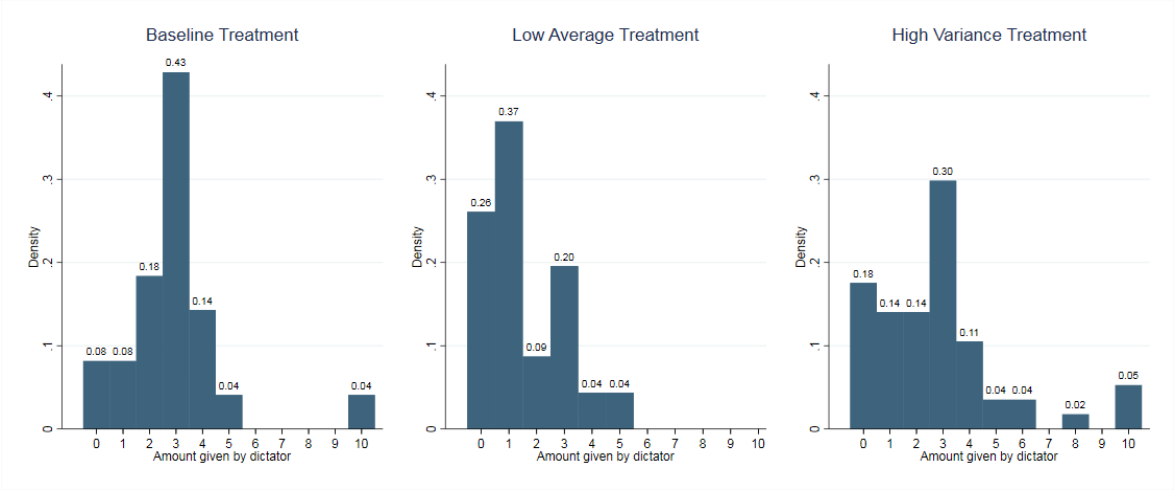
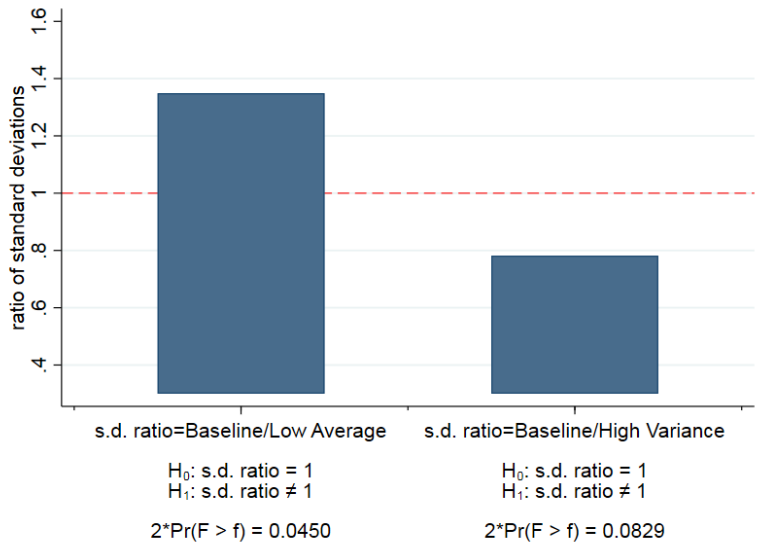


Figure 4 compares the variance of the distribution of the amounts given by the dictators in each treatment. Here the ratio of the two variances (rather than the difference between variances) is compared to the null hypothesis that the ratio equals one. The High Variance treatment indeed yields a significantly higher variance of the amounts given than the Baseline treatment, while the Low Average treatment has lower variance than the Baseline treatment.

Figure 4. Dictator giving in Experiment A, variance-comparison test



The findings illustrated in Figure 2 are confirmed also by simple linear regressions. In Table 3, we regress the amount given on a dummy variable for the Low Average treatment, and a

dummy variable for the High Variance treatment. The omitted category is the Baseline treatment. Column 1 only adds an intercept, while the remaining columns add fixed effects for the time of day or for the day, respectively. Only the Low Average treatment is statistically significant, and the estimated coefficients are quite stable and similar to the effects displayed in Figure 2.[22]

Table 3. Treatment effect on dictator giving, Experiment A

	Amount given by dictator		
	(1)	(2)	(3)
Low Average	-1.397***	-1.723***	-1.348***
	(0.342)	(0.558)	(0.446)
High Variance	-0.129	-0.463	-0.181
	(0.422)	(0.601)	(0.437)
Constant	2.918***	3.027***	2.940***
	(0.272)	(0.375)	(0.345)
Time Fixed Effects	No	yes	no
Day Fixed Effects	No	no	yes
Observations	152	152	152
R-squared	0.087	0.104	0.108

Note: OLS estimates. Regressions in columns (1) to (3) include observations from Baseline, Low Average and High Variance treatments (sessions A). Robust standard errors in parentheses. *** p<0.01, **<0.05, * p<0.1

Table 4 repeats the same analysis for different properties of the distribution of the amounts given. In Columns 1 to 3, the dependent variable is a dummy variable that equals 1 if a positive amount is donated, and 0 otherwise. Individuals assigned to the Low Average treatment are less likely to give a positive sum. The High Variance treatment also decreases the probability of giving a positive amount, although here statistical significance is lower. In Columns 4 to 6, the dependent variable is the amount donated, but the sample is restricted to subjects who give a positive amount. Conditional on giving, the amount given is smaller in the Low Average treatment. The High Variance dummy, instead, is not statistically significant.

[22] We obtain qualitatively similar results when we compare average donations in the Baseline versus the milder Low Average Treatment (see footnote 13): average donations in the milder Low Average Treatment are smaller than in the Baseline, but the difference is not statistically significant (p = 0.1688). We interpret this finding as evidence that, in order to influence actions, the informational treatments need to exhibit a sufficiently stark contrast.

Table 4. Treatment effect on dictator giving, Experiment A: extensive and intensive margins

	(1)	(2)	(3)	(4)	(5)	(6)
	Give>0	Give>0	Give>0	Amount given| give>0	Amount given| give>0	Amount given| give>0
Low Average	-0.179**	-0.250**	-0.186**	-1.119***	-1.219**	-0.955**
	(0.0764)	(0.0991)	(0.0913)	(0.339)	(0.563)	(0.443)
High Variance	-0.0938	-0.179*	-0.101*	0.205	0.0833	0.191
	(0.0644)	(0.103)	(0.0607)	(0.423)	(0.612)	(0.440)
Constant	0.918***	1.066***	1.001***	3.178***	2.875***	2.909***
	(0.0395)	(0.0546)	(0.0489)	(0.263)	(0.355)	(0.331)
Time Fixed Effects	no	yes	no	no	yes	no
Day Fixed Effects	no	no	yes	no	no	yes
Observations	152	152	152	126	126	126
R-squared	0.035	0.095	0.114	0.082	0.109	0.098

Note: OLS estimates. Regressions in columns (1) to (6) include observations from Baseline, Low Average and High Variance treatments (sessions A). In columns from (1) to (3), the dependent variable is a dummy equal to 1 if the dictator has given 0. In columns (4) to (6), the dependent variable is the amount given by dictators, in the restricted sample of those giving more than 0. Notes: Robust standard errors in parentheses. Robust standard errors in parentheses. *** p<0.01, **<0.05, * p<0.1.

Overall, these results suggest that being exposed to information about the normative beliefs of others affects individual behavior. On average the dictator is more generous, both on the extensive and on the intensive margin, if he is told that more people consider that a more generous behavior is socially appropriate. More dispersion in the normative beliefs of others increases the dispersion of individual donations and increases the fraction of individuals who give 0. We find no effect on the average amount given, i.e. it reduces giving only on the extensive margin.[23] This effect on the extensive margin is consistent with the idea that individuals become more selfish when there is less consensus about the norm (see also footnote 11 above).

These findings are consistent with our theoretical priors. Being exposed to information about what others regard as socially appropriate changes behavior in the direction of the average perceived social norm. Moreover, being informed that there is more disagreement over the contents of the social norm leads to more dispersion in individual donation and marginally increases the

[23] We find no significant treatment effects on our hypothetical measure of Receivers' aggrievement (see footnote 20). Results are available upon request.

frequency of very selfish behavior. The average amount donated is the same irrespective of the perceived consensus on the normative standards; the theory here has no predictions on aggregate behavior, and with these data we cannot test the more specific predictions on the heterogeneous effects of the high variance treatment.

We now explore the mechanisms behind these effects.

6.2. Treatment effect on beliefs

We use data from Experiment B&A to study how informational treatments affect individual beliefs, focusing in particular on: (i) individual assessments of what is the-right-thing-to-do, *r;* (ii) individual beliefs of what others regard as the-right-thing-to-do on average, $E(r)$; (iii) individual beliefs of what others will actually do on average, $E(x)$. Since for (ii) and (iii) we observe the whole distribution of beliefs, we can also study the treatment effect on the variance of normative and positive beliefs of each respondent, namely $V(r)$ and $V(x)$. Recall that this sample of respondents is different from that analyzed above in section 5.1 (i.e., Experiment A), where play of the dictator game was preceded by the informational treatment but there was no belief elicitation.

Table 5 presents descriptive statistics of the three beliefs by treatment. Consider first the comparison over the contents of beliefs, i.e., compare the elements of each row. For all treatments, $E(x)$ is always smaller than both r and $E(r)$ (all p-values < 0.0001, two-sided t.-tests). In other words, on average individuals expects that dictators will donate less than what they deem the right thing to do, and of the perceived social norm.

Next, consider the treatment effects (i.e. compare the element of each column). Compared to the Baseline, the Low Average treatment displays significantly lower beliefs in all three dimensions, r, $E(r)$ and $E(x)$ and the differences are statistically significant at 1% or lower. The effect of the Low Average Treatment is particularly strong on $E(r)$ and $E(x)$. In the High Variance treatment, instead, beliefs are not statistically different from the Baseline, except for $E(x)$ which is only marginally higher in the High Variance treatment ($p = 0.0972$). Note also that the standard deviation of r is highest in the High Variance treatment, and higher in the Low Average treatment than in the Baseline.

Table 5. Average beliefs in Experiment B&A, by treatment

	N	*r*	*E(r)*	*E(x)*
Baseline	98	3.643	4.020	2.347
		(1.151)	(0.916)	(1.525)
Low Average	96	3.094	1.897	1.264
		(1.693)	(1.287)	(0.967)
High Variance	96	3.698	4.003	2.725
		(1.795)	(1.214)	(1.635)

Note: the table reports means and standard deviations, in parentheses, of subjects' elicited beliefs in Experiment B&A.

The fact that different components of beliefs react to treatments in the same direction reflects a strong positive correlation between such components. Table 6 shows that a subject generally believes that, on average, others share his/her own assessment of the right-thing-to-do, as shown by the correlation coefficient between r and $E(r)$ ($\rho = 0.448$, $p<0.001$). He/she also expects that others' actions will be consistent with what she regards as the right-thing-to-do, as captured by the correlation between r and $E(x)$ ($\rho = 0.306$, $p<0.001$). Finally, a subject expects that others' actions will be consistent with what they regard as the right-thing-to do: the correlation coefficient between $E(r)$ and $E(x)$ is $\rho = 0.570$ ($p<0.001$).

Table 6. Correlation between *r*, *E(r)* and *E(x)*

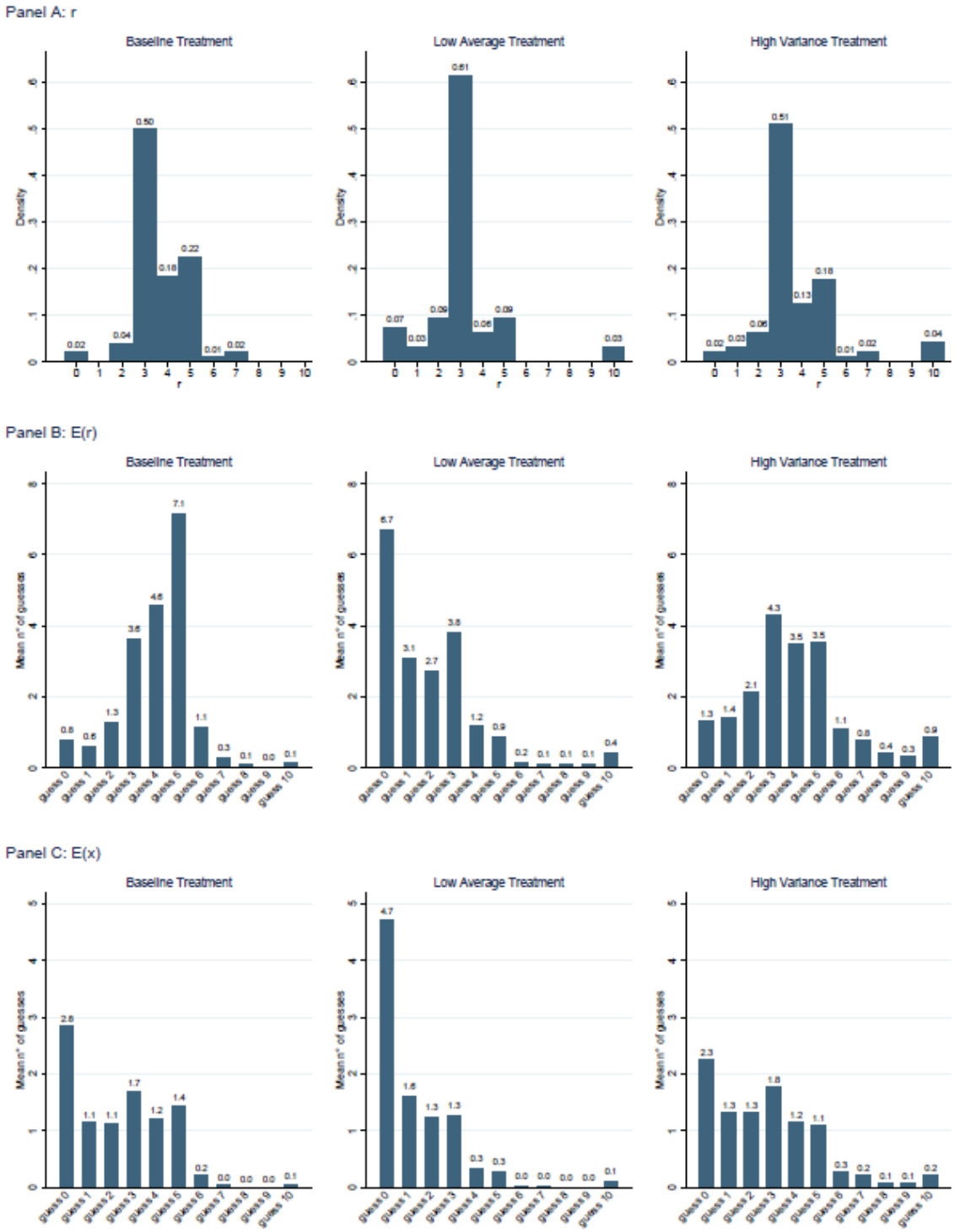
	r	*E(r)*	*E(x)*
r	1		
E(r)	0.448***	1	
E(x)	0.306***	0.570***	1

Note: Pearson correlation coefficients. Baseline, Low Average, and High Variance treatments are included. *** p<0.01, ** p<0.05, * p<0.1

We now exploit our rich data on beliefs to examine their distributions. Figure 5 depicts the distribution of r in Panel A, of $E(r)$ in Panel B and of $E(x)$ in Panel C. The distributions look quite different across treatments, in particular the Low Average treatment is associated with a shift of the distribution of beliefs to the left, towards lower values. This effect is particularly pronounced on the statistics $E(r)$, namely beliefs of what others regard as the right-thing-to-do on average.

Pearson's Chi-square distribution tests reveal that the distributions of all types of beliefs in the Baseline and the Low Average treatments are significantly different ($p = 0.001$ for r, $p = 0.000$ for $E(r)$ and $p = 0.004$ for $E(x)$). There instead are no significant differences in the distributions of beliefs between the Baseline and the High Variance treatments.

Figure 5. Distribution of beliefs, Experiment B&A



We turn to regression analysis to further test these results. Table 7 regresses distribution statistics for individual beliefs on dummy variables for the Low Average and High Variance treatments, plus an intercept, thus again the omitted category is the Baseline treatment. Being

exposed to the Low Average treatment induces a drop in one's perception of what is right-thing-to-do, *r*, in beliefs of what others regard as right-thing-to-do on average, *E*(*r*), and in the positive beliefs of what others will actually do on average, *E*(*x*). The High Variance treatment increases the dispersion of normative and positive beliefs, *V*(*r*) and *V*(*x*), with no effect on average beliefs (except for a small increase in *E*(*x*) which is only significant at the 10% level).[24]

Table 7. Treatment effect on beliefs, Experiment B&A

	r	*E(r)*	*E(x)*	*V(r)*	*V(x)*
	(1)	(2)	(3)	(4)	(5)
Low Average	-0.549***	-2.123***	-1.084***	1.266***	0.258
	(0.208)	(0.161)	(0.183)	(0.400)	(0.367)
High Variance	0.0551	-0.0174	0.378*	2.667***	1.230***
	(0.217)	(0.155)	(0.227)	(0.382)	(0.397)
Constant	3.643***	4.020***	2.347***	1.568***	1.772***
	(0.116)	(0.0926)	(0.154)	(0.124)	(0.173)
Observations	290	290	290	290	290
R-squared	0.029	0.431	0.163	0.115	0.033

Note: OLS estimates. Regressions in columns (1) to (5) include observations from Baseline, Low Average and High Variance treatments (sessions B). Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Since all beliefs move in the same direction with the treatment, we cannot fully disentangle which beliefs are responsible for the observed change in behavior. Nevertheless, the effect of the Low Average treatment on *E*(*r*) is particularly large: the estimated coefficient of -2.1 is twice or four times as large in absolute value as the effects on *E*(*x*) and *r* respectively, and it is also larger in absolute value than the estimated effect of almost -1.4 on the average amount donated – cf. column 1 of Table 7. This suggests that the Low Average treatment affects behavior mainly through its impact on individual beliefs of what others regard as socially appropriate (i.e. on the content of the social norm, *E(r)*), rather than through an effect on individual values, *r*, or on

[24] We obtain similar results when comparing beliefs in the Baseline and in the milder Low Average treatment (see footnote 13). The milder Low Average treatment has the expected effect on beliefs on what others perceive to be the right-thing-to-do: *E(r)* is significantly lower (p = 0.000, two-sided t-test) and *V(r)* is significantly higher (p = 0.0764) then in the Baseline treatment. In the milder Low Average treatment beliefs on others' actions, *E(x)*, are also on average lower, but not significantly so (p = 0.4196), while, surprisingly, *r* is significantly higher on average (p = 0.0634).

expectations of actual behavior of others. This result is consistent with the mechanism postulated by our theory.

Tables A2 and A3 in the Appendix repeat the same exercise controlling for time of day or day fixed effects. Some of the estimated coefficients loose significance, but the effect of the Low Average treatment on $E(r)$ and the effect of the High Variance treatment on $V(r)$ and $V(x)$ are very robust and stable.

Finally, in the sessions with beliefs elicitation, we also asked participants to play the dictator game after expressing their belief. We can thus explore how the informational treatments influenced actions in this setting too, comparing each treatment to the Baseline. Figure A.1 in the appendix compares the average amounts given. Unlike in Figures 2 above, here neither treatment has any effect on the average amount given, compared to the Baseline. Figure A.2 compares the variance of the amount, also to the Baseline. The High Variance treatment increases the dispersion in the amount given compared to the Baseline, as in Figure 3, but the variance of amounts given is also larger in the Low Average treatment than in the Baseline. These results are confirmed by OLS estimates (Appendix Tables A4 and A5). Thus, as argued above in section 4.2, the belief elicitation stage dampens the treatment effects on dictator giving. When we add belief elicitation to the experiment, the evidence of a causal effect of information on the amount given disappears.

This also explains why we fail to find support for prediction (ii) on the heterogeneous effect of a change in *V(r)*. Recall that we predict that more consensus on the right thing to do increases (decreases) the amount donated if $E(r) - r > 0$ (< 0). Testing this prediction requires examining whether the impact of the high variance treatment on donation differs between subjects with $E(r) - r > 0$ and subjects with $E(r) - r < 0$. We thus can test this prediction only with data on both beliefs and actions for the same subjects, i.e. with data from Experiment B&A. However, the lack of effect of the experimental treatments on actions in Experiment B&A limits our ability to perform this analysis. Indeed, when we conduct these tests we fail to detect any statistically significant results.

6.3. Correlation between beliefs and actions

We now turn to examining the correlation between beliefs and behavior. We exploit data from the B&A treatments, where we observe both beliefs and actions of the sample of Dictators, to correlate each Dictator's choice of x with her own beliefs $E(r)$, $E(x)$ and r. Table 8 reports the

correlation coefficients between amount given and each of the three beliefs. We compute these correlations first pooling the Baseline, Low Average and High Variance treatments together, and then considering each treatment in isolation.[25] Our theoretical model predicts a positive correlation between players' beliefs about others' *r*, i.e. their *E(r)*, and their choice of *x*. Moreover, the theory predicts that such correlation should be weaker in the High Variance treatment than in the Baseline and Low Average treatments.

The results displayed in Table 8 are consistent with the model's predictions: Colum 2 reports the correlation between *x* and *E(r)*, which is positive when we pool all treatments together, as well as when we consider each of them separately. Moreover, such correlation is statistically significant only in the Baseline and Low Average treatments ($p = 0.0515$ and $p = 0.0245$, respectively), while it is smaller in magnitude and far from statistically significant in the High Variance treatment.

Other beliefs are also, overall, significantly and positively correlated with donation amounts. In particular, consistent with the argument made in section 4.2, that reasoning in the abstract on "the right thing to do" may lead participants to give more weight to the intrinsic merit of different choices and act based on moral or value criteria, rather than on social conventions, the correlation between *r* and the amount given by dictators is positive and statistically significant. Of course, these are only correlations, and cannot be interpreted as a causal effect of beliefs on actions.

[25] We opt to present these results through pairwise correlations, rather than through regressions of donation amount on beliefs, which would include simultaneously *E(r)*, *E(x)* and *r*, because of the strong multicollinearity between the different beliefs, as reported in Table 6.

Table 8. Correlation between actions and beliefs

	N	*E(r)*	*E(x)*	*r*
	(1)	(2)	(3)	(4)
All	290	0.0934	0.2410	0.2379
		(0.1125)	(0.0000)	(0.0000)
Baseline	96	0.1973	0.3313	0.2307
		(0.0515)	(0.0009)	(0.0223)
Low Average	98	0.2295	0.3245	0.3406
		(0.0245)	(0.0013)	(0.0007)
High Variance	96	0.0427	0.2622	0.1529
		(0.6796)	(0.0099)	(0.1369)

Note: each cell reports the Pearson correlation coefficient between the amount given by Dictators in the B&A treatments and their own beliefs *E(r)* (Column 2), *E(x)* (Column 3) and *r* (Column 4). Baseline, Low Average, and High Variance treatments are pooled in the first row, and considered separately in the remaining rows. P-values in parentheses.

7. Discussion

The experimental results in this paper support the idea that social norms, defined as perceived normative standards of others, react to available information and influence behavior. Informational treatments that shift the perceived social norm also shift behavior in the same direction. And informational treatments that reduce the perceived consensus on normative standards increase dispersion in individual behavior, consistently with the idea that the social norm is less influential if it is perceived as less widely shared.

An important question for future research is why individuals react to the perceived normative standards of others. The literature on social identity suggests an answer: because individuals who identify with a social group behave consistently with their perceptions of how a typical member of the group ought to behave. This leads to the conjecture that increasing the salience of group identification, and creating situations of group conflict, would also increase the influence of group norms (defined as normative standards shared within the group). Exploring the empirical validity of this conjecture is both feasible and interesting.

A second related question is whether individuals react differently to normative vs positive features of the group with which they identify. This question too lends itself to empirical analysis,

for instance manipulating perceptions of the normative standards vs the actual behavior of typical group members.

More generally, studying the link between social norms and social identities, and how both are influenced by perceptions of group features and of group leaders is an important and promising area of research.

References

Allcott, H. (2011), "Social Norms and Energy Conservation." *Journal of Public Economics* 95: 1082-1095.

Andreoni, J., and B. D. Bernheim (2009), "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica* 77: 1607-1636.

Ayres, I., Raseman, S. and Shih, A. (2013), "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage." *Journal of Law, Economics and Organizations* 29(5): 992-1022.

Battigalli, P. and Dufwenberg, M. (2007), "Guilt in Games." *American Economic Review: Papers & Proceedings* 97: 170-176.

Battigalli, P. and Dufwenberg, M. (2009), "Dynamic Psychological Games." *Journal of Economic Theory* 144: 1-35.

Beaman, L., Duflo, E., Pande, R. and Topalova, P. (2012), "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science* 335: 582-586.

Bénabou, R. and J. Tirole (2006), "Incentives and Prosocial Behavior." *American Economic Review* 96: 1652-78.

Bernheim, B. D. (1994). "A Theory of Conformity." *Journal of Political Economy* 102: 841-877.

Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C. and Milkman, K.L. (2015), "The Effect of Providing Peer Information on Retirement Savings Decisions." *The Journal of Finance* 70: 1161-1201.

Bicchieri, C. (2005), *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press, Cambridge, MA.

Bicchieri, C and Dimant, E. (2018), "It's Not A Lie If You Believe It. Lying and Belief Distortion Under Norm-Uncertainty." *PPE Working Papers 0012*, University of Pennsylvania.

Bicchieri, C. and Xiao E. (2009), "Do the right thing: But only if others do so." *Journal of Behavioral Decision Making*, 22(2): 191-208.

Bursztyn, L., Gonzalez, A. L. and Yanagizawa-Drott, D. (2018), "Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia." *NBER Working Paper* 24736.

Bursztyn, L., G. Egorov, G. and Fiorin, S. (2019), "From Extreme to Mainstream: How Social Norms Unravel", *mimeographed*.

Byrne, D.P., La Nauze, A. and Martin L.A. (2018), "Tell Me Something I Don't Already Know: Informedness and the Impact of Information Programs." *The Review of Economics and Statistics*, 100: 510-527.

Cantoni, D., Yang, D. Y., Yuchtman, N. and Zhang, Y. J. (2018), "Protests as Strategic Games: Experimental Evidence from Hong Kong's Anti-Authoritarian Movement." *Quarterly Journal of Economics*, forthcoming.

d'Adda, G., Darai, D., Pavanini, N. and Weber, R.A. (2017), "Do Leaders Affect Ethical Conduct?" *Journal of the European Economic Association* 15: 1177-1213.

Danilov, A., Khalmetski K. and Sliwka D. (2018). “Norms and Guilt.” *CESifo Working Paper* 6999.

Drouvelis, M. and Nosenzo, D. (2013), “Group Identity and Leading-by-example.” *Journal of Economic Psychology* 39: 414-425.

Elster, J. (1989), “Social Norms and Economic Theory.” *The Journal of Economic Perspectives* 3(4): 99-117.

Ferraro, P.J., Miranda, J.J. and Price, M.K. (2011), “The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment.” *The American Economic Review* 101: 318-322.

Ferraro, P.J. and Price, M.K. (2013), “Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment.” *Review of Economics and Statistics* 95: 64-73.

Fehr, E. and Schurtenberger, I. (2018), “Normative Foundations of Human Cooperation.” *Nature* 2: 458-468.

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

Geanakoplos, J., Pearce, D. and Stacchetti, E. (1989), “Psychological games and sequential rationality.” *Games and Economic Behavior* 1: 60-79.

Gerber, A.S. and Rogers, T. (2009), “Descriptive Social Norms and Motivation to Vote: Everybody’s Voting and So Should You.” *The Journal of Politics* 71: 178-191.

Greiner, B. (2015), “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE.” *Journal of the Economic Science Association* 1(1): 114-125.

Hogg, Michael A. and Dominic Abrams (1998), *Social Identification - A Social Psychology of Intergroup Relations and Group Processes.* Routledge, London.

Jensen, R. and Oster, E. (2009), “The Power of TV: Cable Television and Women’s Status in India.” *The Quarterly Journal of Economics* 124: 1057-1094.

Krupka, E. L. and Weber, R. (2013), “Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?” *Journal of the European Economic Association*, 11(3): 495-524.

La Ferrara, E., Chong, A. and Duryea, S. (2012), “Soap operas and fertility: Evidence from Brazil.” *American Economic Journal: Applied Economics* 4: 1-31.

López-Pérez, Raúl (2008) “Aversion to Norm-breaking: A Model.” *Games and Economic Behavior* 64: 237-267.

Michaeli, M. & D. Spiro (2015), “Norm Conformity across Societies.” *Journal of Public Economics* 132, 51-65.

Paluck, E.L. (2009), “What’s in a Norm? Sources and Processes of Norm Change.” *Journal of Personality and Social Psychology* 96(3): 594-600.

Peysakhovich, A. and Rand, D.G. (2015), “Habits of Virtue: Creating Norms of Cooperation and Defection in the Laboratory.” *Management Science* 62: 631-647.

Shang, J. and Croson, R. (2009), "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods." *The Economic Journal* 119: 1422-1439.

Tankard, M. E. and Paluck, E. L. (2016), "Norm Perception as a Vehicle for Social Change." *Social Issues and Policy Review*, 10: 181-211

APPENDIX A

Figures

Figure A1. Treatment effect on the mean of dictator giving, Experiment B&A

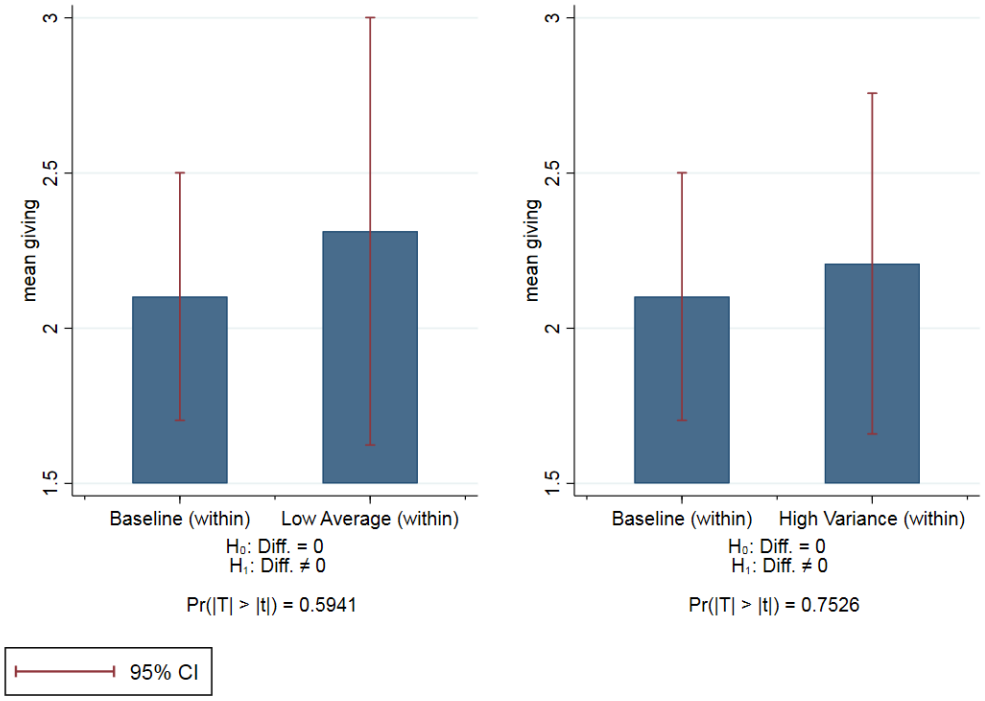
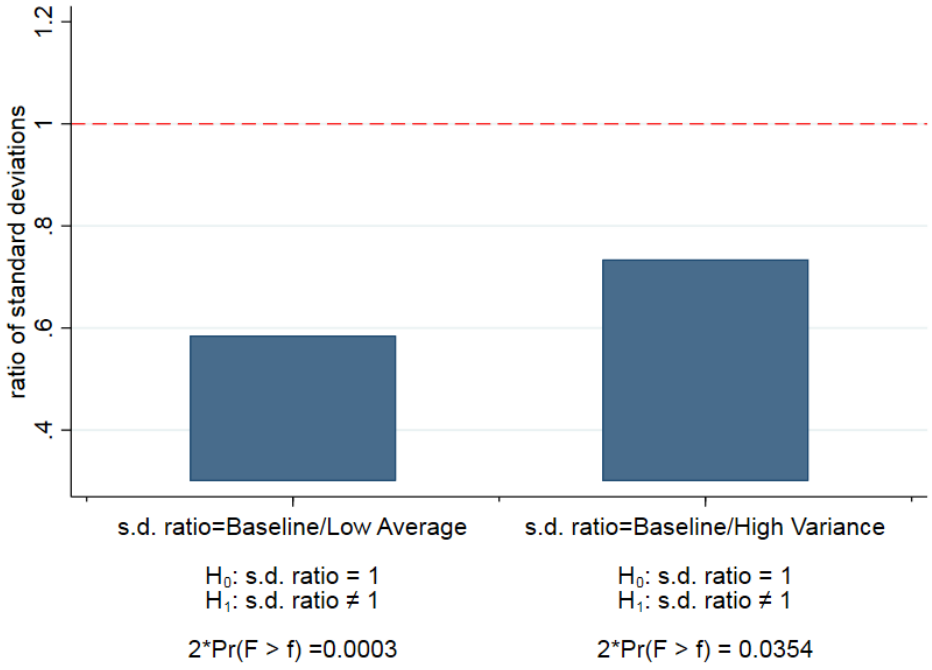


Figure A2. Treatment effect on the variance of dictator giving, Experiment B&A



Tables

Table A1. Average dictator giving, by experiment and treatment

	Amount given by Dictator			
	Experiment A		Experiment B&A	
	N	Mean (sd)	N	Mean (sd)
Baseline	49	2.918 (1.902)	49	2.102 (1.388)
Low Average	46	1.522 (1.410)	48	2.312 (2.371)
High Variance	57	2.789 (2.433)	48	2.208 (1.890)

Note: the table reports means and standard deviations of the amount given by the dictator.

Table A2. Treatment effect on beliefs, Experiment A, time Fixed Effects

	r	*E(r)*	*E(x)*	*V(r)*	*V(x)*
	(1)	(2)	(3)	(4)	(5)
Low Average	-0.771	-2.450***	-0.915**	-0.124	-0.302
	(0.509)	(0.339)	(0.357)	(0.527)	(0.447)
High Variance	0.211	0.101	0.403	2.826***	1.558***
	(0.257)	(0.192)	(0.307)	(0.463)	(0.434)
Constant	3.510***	3.811***	2.373***	1.307***	0.968***
	(0.236)	(0.187)	(0.328)	(0.379)	(0.348)
Time Fixed Effects	yes	yes	yes	yes	yes
Observations	290	290	290	290	290
R-squared	0.057	0.439	0.171	0.131	0.062

Note: OLS estimates. Regressions in columns (1) to (5) include observations from Baseline, Low Average and High Variance treatments (sessions B&A). Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Treatment effect on beliefs, Experiment A, date Fixed Effects

	r	*E(r)*	*E(x)*	*V(r)*	*V(x)*
	(1)	(2)	(3)	(4)	(5)
Low Average	-0.595	-1.909***	-0.359	1.738***	1.096**
	(0.366)	(0.252)	(0.327)	(0.536)	(0.476)
High Variance	0.0432	-0.00267	0.202	2.615***	1.277***
	(0.232)	(0.174)	(0.266)	(0.444)	(0.482)
Constant	3.700***	3.852***	1.625***	1.375***	1.140***
	(0.313)	(0.203)	(0.306)	(0.252)	(0.247)
Date Fixed Effects	Yes	yes	yes	yes	yes
Observations	290	290	290	290	290
R-squared	0.039	0.438	0.183	0.134	0.050

Regressions in columns (1) to (5) include observations from Baseline, Low Average and High Variance treatments (sessions B&A). Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Treatment effects on dictator giving, Experiment B&A

	Amount given by dictator		
	(1)	(2)	(3)
Low Average	0.210	0.886	0.411
	(0.396)	(0.966)	(0.625)
High Variance	0.106	0.0599	-0.000509
	(0.337)	(0.404)	(0.336)
Constant	2.102***	2.407***	1.800***
	(0.198)	(0.413)	(0.500)
Time Fixed Effects	no	yes	no
Date Fixed Effects	no	no	yes
Observations	145	145	145
R-squared	0.002	0.070	0.020

Note: OLS estimates. Regressions in columns (1) to (4) include observations from Baseline, Low Average and High Variance treatments (sessions B&A). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A5. Treatment effect on dictator giving, Experiment B&A: extensive and intensive margins

	(1)	(2)	(3)	(4)	(5)	(6)
	Give>0	Give>0	Give>0	Amount given| give>0	Amount given| give>0	Amount given| give>0
Disclosure Mode 0 within	-0.00383	0.185	0.0895	0.271	0.458	0.229
	(0.0798)	(0.157)	(0.164)	(0.408)	(0.992)	(0.619)
Distribution with high var	-0.0455	-0.0640	-0.0537	0.290	0.325	0.196
	(0.0830)	(0.106)	(0.0950)	(0.319)	(0.369)	(0.279)
Constant	0.816***	0.837***	0.700***	2.575***	2.875***	2.571***
	(0.0559)	(0.0983)	(0.149)	(0.168)	(0.360)	(0.461)
Time Fixed Effects	no	yes	no	no	yes	no
Day Fixed Effects	no	no	yes	no	no	yes
Observations	145	145	145	116	116	116
R-squared	0.003	0.083	0.024	0.006	0.042	0.022

OLS estimates. Regressions in columns (1) to (6) include observations from Baseline, Low Average and High Variance treatments (sessions B&A). In columns from (1) to (3), the dependent variable is a dummy equal to 1 if the dictator has given 0. In columns (4) to (6), the dependent variable is the amount given by dictators, in the restricted sample of those giving more than 0. Notes: Robust standard errors in parentheses. Notes: robust standard errors in parentheses. *** p<0.01 **<0.05, * p<0.1

Table A6. Conditional (fixed-effects) logit estimation of correlates of action across treatments, controlling for payoffs associated to each action

Action *a* is taken	(1)	(2)	(3)	(4)
r_a	5.405***			4.206***
	(0.734)			(0.900)
$E(r_a)$		3.883***		3.347**
		(0.631)		(1.355)
$E(x_a)$			1.193	-1.916
			(0.946)	(1.747)
Potential payoff	0.337***	0.276***	0.256***	0.392***
	(0.0397)	(0.0310)	(0.0465)	(0.0884)
Subjects	152	152	152	152
Observations	1,672	1,672	1,672	1,672
Log-Likelihood	-288.6	-299.5	-310.4	-285.7

Note: Conditional logit estimates. Baseline, Low Average, and High Variance treatments are included. ***p<0.01, ** p<0.05, * p<0.1

APPENDIX B

Experimental instructions[26]

Experiment A

PART I: INTRODUCTION

Instructions for everybody (Experimenter read aloud)

This is a study in decision making. For your participation, you will be paid a participation fee of €3. In addition, you may receive some additional money based on your choices and the choice of others during the experiment.

If you have any questions during the study, please raise your hand and wait for an experimenter to come to you. Please do not talk, exclaim, or try to communicate with other participants during the experiment.

All participants have a carton box on their desk. Please do not touch any material you find on your desk until you are instructed to do so.

Description

In the experiment you will randomly be paired with another subject in the room. Thus 10 pairs will be formed. The pairing is anonymous, meaning that neither individual will ever know the identity of the other individual with whom she or he is paired.

In each pair, one individual will be randomly selected to play in the role of "Individual A", while the other one will be "Individual B". This means that you have 50% probability to be either Individual A or Individual B.

Individual A will find two envelopes in his/her box. One envelope is white, and the other one is yellow.

- The white envelope is marked "Money for Individual A", and it contains €10 in coins.
- The yellow envelope is marked "Money for Individual B", and it contains no money.

The box of Individual B is empty.

Individual A must decide how much money to keep and how much to transfer to individual B. The

[26] Text in italics was not included in the instructions handed out to the participants. It contains information to help the reader understand how the experiment was conducted.

amount that Individual A decides to transfer to Individual B will be doubled by the experimenter. The table below reports the possible choices that Individual A can make, and the corresponding amount of money both Individual A and Individual B will receive after A's choice.

Table A

Individual A must choose one of the following actions										
Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
Resulting payment to A										
10 €	9 €	8 €	7 €	6 €	5 €	4 €	3 €	2 €	1 €	0 €
Resulting payment to B										
0 €	2 €	4 €	6 €	8 €	10 €	12 €	14 €	16 €	18 €	20 €

Once Individual A has made a choice, he/she will keep his/her money, while the experimenter will take care of the payment for Individual B at the end of the experiment. All participants will follow instructions provided on the screen or by the experimenter to ensure anonymity.

Every participant will be paid his/her participation fee and any additional earnings from the experiment before leaving the laboratory.
Individual A's choice is **anonymous**. No other subject will know about Individual A's decision.
Now please proceed to the next page, by clicking the Right arrow on the keyboard.

PART II: TREATMENT

On screen instructions for everybody

Before starting the experiment, we wish to let you know that we have run sessions similar to the one you are in before. Table B describes the proportion of choices that occurred in a sample of responses that previous subjects gave to this question.

Table B

(The content of Table B varies by treatment: Baseline, Low Average and High Variance)

"Which is the socially most appropriate action that Individual A should take?"											
	Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
Percentage of people who answered	..%	..%	..%	..%	..%	..%	..%	..%	..%	..%	..%

Now please follow the instructions on the screen. These instructions will reveal whether you have been selected to be Individual A or B.

PART III: MODIFIED DICTATOR GAME

On screen instructions for Dictators, p. 1

You have been paired to another participant, and you have been selected to play the role of

Individual A

Please open the box on your table. You will find two envelopes.

- The white envelope is marked "Money for Individual A". It contains €10.
- The yellow envelope is marked "Money for Individual B". It is empty.

You can transfer money from the white envelope to the yellow envelope. Please do it now, **paying attention that your action is not visible or audible to other persons in the room, so as to preserve anonymity**.

Please make your choice and record it in the Table below.

Table 4A

Please, make your choice. Tick the corresponding box in this table and put the corresponding amount in the yellow envelope										
Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Please insert the yellow envelope in the box. The white envelope is for you. You can put it in your pocket, or simply take the money and put it in your pocket, leaving the empty envelope in the box. Please remain seated. The experimenter will come to your table to collect the box.

On screen instructions for Receivers, p. 1

You have been paired to another participant. You have been selected to play the role of

Individual B

You can open the box on your table. It is empty. Please remain seated. The experimenter will collect it and hand it back to you with the money that your paired participant decided to transfer to you.

Instructions for everybody (Experimenter read aloud)

The experimenter will now collect all of the boxes and take them to the other room. There, he/she will transfer the yellow envelopes from each Individual A's box to the box of the Individual B matched with him/her. The experimenter will then return all boxes. Once boxes have been returned, please follow further instructions on the screen.

On screen Instructions for Dictators, p. 2

The experimenter hands you back the box after he/she has transferred the yellow envelope to your paired receiver. Now, please answer the question below.

Table 5A

If you could write a message to your paired Individual B, explaining your choice, what would it be?
…… ……

Now please go to the next page.

On screen Instructions for Receivers, p. 2

Once the experimenter hands you back the box, please check the yellow envelope in it. This is the money Individual A transferred to you. Recall that at the end of the experiment that money will be doubled by the experimenter. Now, please answer the questions below.

Table 5B

How much money did you receive in the envelope? **….. €.**										
If you could write a message to your paired Individual A expressing your feelings, what would it be? …… ……										
How much would you be willing to pay to have Individual A lose all his/her money? *(please tick one box below)*										
Pay 0 €	Pay 1 €	Pay 2 €	Pay 3 €	Pay 4 €	Pay 5 €	Pay 6 €	Pay 7 €	Pay 8 €	Pay 9 €	Pay 10 €
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

You can keep the money in the yellow envelope, it is yours. You can put the envelope in your pocket, or simply take the money and put it in your pocket, leaving the empty envelope in the box. Then please go to the next page.

PART IV: FINAL INSTRUCTIONS

Instructions for everybody (Experimenter read aloud)

We will now call you outside one by one by your ID (i.e. the number of the position where you are seated). We will give you your participation fee and the money you earned from guessing other participants' answers.

Experiment B&A

PART I: INTRODUCTION

Instructions for everybody (Experimenter read aloud)

This is a study in decision making. For your participation, you will be paid a participation fee of €3. In addition, you may receive some additional money based on your choices and the choice of others during the experiment.

If you have any questions during the study, please raise your hand and wait for an experimenter to come to you. Please do not talk, exclaim, or try to communicate with other participants during the experiment.

All participants have a carton box on their desk. Please do not touch any material you find on your desk until you are instructed to do so.

Description

In the experiment you will randomly be paired with another subject in the room. Thus 10 pairs will be formed. The pairing is anonymous, meaning that neither individual will ever know the identity of the other individual with whom she or he is paired.

In each pair, one individual will be randomly selected to play in the role of "Individual A", while the other one will be "Individual B". This means that you have 50% probability to be either Individual A or Individual B.

Individual A will find two envelopes in his/her box. One envelope is white, and the other one is yellow.

- The white envelope is marked "Money for Individual A", and it contains €10 in coins.
- The yellow envelope is marked "Money for Individual B", and it contains no money.

The box of Individual B is empty.

Individual A must decide how much money to keep and how much to transfer to individual B. The amount that Individual A decides to transfer to Individual B will be doubled by the experimenter. The table below reports the possible choices that Individual A can make, and the corresponding amount of money both Individual A and Individual B will receive after A's choice.

Table A

Individual A must choose one of the following actions										
Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
Resulting payment to A										
10 €	9 €	8 €	7 €	6 €	5 €	4 €	3 €	2 €	1 €	0 €
Resulting payment to B										
0 €	2 €	4 €	6 €	8 €	10 €	12 €	14 €	16 €	18 €	20 €

Once Individual A has made a choice, he/she will keep his/her money, while the experimenter will take care of the payment for Individual B at the end of the experiment. All participants will follow instructions provided on the screen or by the experimenter to ensure anonymity.

Every participant will be paid his/her participation fee and any additional earnings from the experiment before leaving the laboratory.
Individual A's choice is **anonymous**. No other subject will know about Individual A's decision.

Now please proceed to the next page, by clicking the Right arrow on the keyboard.

PART II: TREATMENT AND BELIEF ELICITATION

On screen instructions for everybody

Questions

Before starting the experiment, we would like to ask your opinion in a series of questions. We would like you to answer as truthfully as possible, based on your personal view. Please read carefully the questions in the next pages. You can receive additional money depending on your answers to these questions.

Question 1

In a few minutes you will be paired with another participant, and the experiment will start. One of you will be randomly selected to play in the role of Individual A. Please **use Table 1** to

answer the following question:

In your opinion, which is the socially most appropriate action that Individual A should take?

By "socially most appropriate", we mean behavior that you consider the "right" or "ethical" thing to do.

Before you answer, we wish to let you know that we have run sessions similar to the one you are in before. Table B describes the proportion of choices that occurred in a sample of responses that previous subjects gave to this question.

Table B

The content of Table B varies by treatment:
Baseline, Low Average and High Variance; cf. Figure 1 in the main text

"Which is the socially most appropriate action that Individual A should take?"											
	Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
Percentage of people who answered	..%	..%	..%	..%	..%	..%	..%	..%	..%	..%	..%

Now please answer our question:

Table 1

"Which is the socially most appropriate action that Individual A should take?" *Tick one box in the table below*										
Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Question 2

Consider the previous question: "*In your opinion, which is the socially most appropriate action that Individual A should take?*" We asked this question to all 20 participants in this session.

Now please,

guess the distribution of their answers.

Please use the last row of the following table. Write in each cell the number of people you think chose that action as the most appropriate for Individual A. For instance, if you think that, say, 3 people answered "Give €x", just write "3" in the cell under "Give €x". You will gain €0.2 for each correct guess.

Be careful, the numbers you write must sum up to 20, because this is the number of participants in the group. You have eleven numbers to guess, and they must sum up to 20.

Table 2

"Which is the socially most appropriate action that Individual A should take?" **For each action, please guess how many people answered that action** *Write the number of people in the last row. The numbers you write must sum up to 20*											
	Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
Number of people who answered	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

Question 3

Now we would like to ask you something different.

Please guess the distribution of the actions participants deciding as Individual A in this session will *actually* take

Please use the last row of following table. Write in each cell the number of "Individuals A" you think will take that action. For instance, if you think that, say, two Individuals A will choose "Give €x", just write "2" in the cell under "Give €x".

Remember that ten participants will be selected as "Individual A" in this session. Thus be careful, the numbers you write must sum up to 10, because this will be the number of individuals A in the group. You have eleven numbers to guess, and they must sum up to 10.

You will gain €0.2 for each correct guess. A guess is correct if it matches the number of Individuals A who will actually take that action in this session.

Table 3

"Please guess the distribution of the actions participants deciding as Individual A in this session will actually take" **For each action, please guess how many "Individuals A" will take that action**											
Write the number of people in the last row. The numbers you write must sum up to 10											
	Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
Number of Individuals A	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

Thank you for answering the three questions above.

Now please follow the instructions on the screen. These instructions will reveal whether you have been selected to be Individual A or B.

PART III: MODIFIED DICTATOR GAME

On screen Instructions for Dictators, p. 1

You have been paired to another participant, and you have been selected to play the role of

Individual A

Please open the box on your table. You will find two envelopes.

- The white envelope is marked "Money for Individual A". It contains €10.
- The yellow envelope is marked "Money for Individual B". It is empty.

You can transfer money from the white envelope to the yellow envelope. Please do it now, **paying attention that your action is not visible or audible to other persons in the room, so as to preserve anonymity**.

Please make your choice and record it in the Table below.

Table 4A

Please, make your choice. Tick the corresponding box in this table and put the corresponding amount in the yellow envelope										
Give nothing	Give 1 €	Give 2 €	Give 3 €	Give 4 €	Give 5 €	Give 6 €	Give 7 €	Give 8 €	Give 9 €	Give 10 €
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Please insert the yellow envelope in the box. The white envelope is for you. You can put it in your pocket, or simply take the money and put it in your pocket, leaving the empty envelope in the box. Please remain seated. The experimenter will come to your table to collect the box.

On screen Instructions for Receivers, p. 1

You have been paired to another participant. You have been selected to play the role of

Individual B

You can open the box on your table. It is empty. Please remain seated. The experimenter will collect it and hand it back to you with the money that your paired participant decided to transfer to you.

Instructions for everybody (Experimenter read aloud)

The experimenter will now collect all of the boxes and take them to the other room. There, he/she will transfer the yellow envelopes from each Individual A's box to the box of the Individual B matched with him/her. The experimenter will then return all boxes. Once boxes have been returned, please follow further instructions on the screen.

On screen Instructions for Dictators, p. 2

The experimenter hands you back the box after he/she has transferred the yellow envelope to your paired receiver. Now, please answer the question below.

Table 5A

If you could write a message to your paired Individual B, explaining your choice, what would it be?
…….
…….

Now please go to the next page.

On screen Instructions for Receivers, p. 2

Once the experimenter hands you back the box, please check the yellow envelope in it. This is the money Individual A transferred to you. Recall that at the end of the experiment that money will be doubled by the experimenter. Now, please answer the questions below.

Table 5B

How much money did you receive in the envelope? **….. €.**										
If you could write a message to your paired Individual A expressing your feelings, what would it be? ……. …….										
How much would you be willing to pay to have Individual A lose all his/her money? *(please tick one box below)*										
Pay 0 €	Pay 1 €	Pay 2 €	Pay 3 €	Pay 4 €	Pay 5 €	Pay 6 €	Pay 7 €	Pay 8 €	Pay 9 €	Pay 10 €
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

You can keep the money in the yellow envelope, it is yours. You can put the envelope in your pocket, or simply take the money and put it in your pocket, leaving the empty envelope in the box. Then please go to the next page.

PART IV: FINAL INSTRUCTIONS

Instructions for everybody (Experimenter read aloud)

We will now call you outside one by one by your ID (i.e. the number of the position where you are seated). We will give you your participation fee and the money you earned from guessing other participants' answers.