

How effective are social norm interventions?

Evidence from a laboratory experiment on managerial honesty

Rajna Gibson, Carmen Tanner, and Alexander F. Wagner*

June 11, 2016

Abstract

Social norms can act as safeguards against corporate misconduct, but can also foster undesirable behavior. We conduct a laboratory experiment where we expose participants (in the role of CEOs) to social norms approving or disapproving of earnings management. There are systematic differences among individuals' reactions to the situational pressure. Specifically, individuals with strong preferences for truthfulness react less to both kinds of social norms. Self-signaling provides a convincing explanation of individual behavior. These findings have implications for the empirical analysis of managerial behavior and for the use of social norms as steering tools for corporate governance.

JEL Classification: G02, G30, C91, M14.

Keywords: Behavioral finance, honesty, social norm interventions, laboratory experiments, corporate governance.

* Gibson (rajna.gibson@unige.ch) is at the Swiss Finance Institute - University of Geneva. Tanner (carmen.tanner@zu.de) is at the Leadership Excellence Institute Zeppelin at Zeppelin University and at the Department of Banking and Finance, University of Zurich. Wagner (alexander.wagner@bf.uzh.ch) is at the Swiss Finance Institute - University of Zurich, CEPR, and ECGI. We thank the Research Priority Program Finance and Financial Markets of the University of Zurich, the NCCR FINRISK, the Swiss Finance Institute, the Swiss National Science Foundation (Grant PP001-102845), and the ERC (FP7/2007-2013, grant agreement 249415 RMAC) for financial support. We thank Nicolas Berkowitsch and Elisabeth Proehl who both provided excellent research assistance, and Marcel Morf, who programmed the first version of the code for the experiments. We thank conference participants at the European Finance Association Meetings 2014 and the Montpellier International Doctoral Colloquium as well as seminar participants at the University of Cambridge, UC San Diego, the University of Hamburg, the University of Lugano, the University of Vienna, and the University of Zurich for comments. We are grateful to Björn Bartling, Michael Brennan, Mariassunta Giannetti, Uri Gneezy, Richard Holton, Michel Maréchal, Abraham Ravid, Jean-Charles Rochet, Joel Sobel, Janna Ter Meer, and Roberto Weber for comments, and to Peter Bossaerts, Chiddi Chidambaran, Alain Cohn, François Degeorge, Ernst Fehr, Gerlinde Fellner, Susanne Neckermann, Mei Wang, and Rainer Winkelmann for discussions on the general topic of this paper.

1 Introduction

How can corporate misconduct and deception be avoided? To foster appropriate employee behavior and to support the implementation of codes of conduct and values statements, increasing attention within corporations has centered on the role of “social norms.” Can corporate decision-makers effectively intervene by employing normative influence? Much like medicinal prescriptions nowadays are mostly based on experiments (“clinical trials”), this paper draws on data from a controlled, laboratory experiment to shed light on whether actively changing social norms has the desired effect on managerial behavior.

Understanding the sources of corporate misconduct and deceptive practices in business is important. In his Presidential Address, Zingales (2015) argues that in the financial sector “fraud has become a feature and not a bug.” Half of all hedge fund failures are due to operational risks, and the vast majority of those involve fraud (Capco, 2003). But it is not just financial institutions that are in the spotlight, as recent cases such as Volkswagen show. Dyck, Morse, and Zingales (2014) report the stunning result that the probability of a company engaging in a fraud in any given year is 14.5%, with substantial costs. 70% of respondents to the EY (2009) European Fraud Survey stated that management are likely to cut corners to meet targets when economic times are tough. Several papers study the impact of fraud and the importance of trust in financial markets (Guiso, Sapienza, and Zingales, 2008; Karpoff, Lee, and Martin, 2008; Dyck, Morse, and Zingales, 2010; Gurun, Stoffman, and Yonker, 2015).

One approach to address these problems is to incentivize managers to behave as desired. Nowadays corporations include “values and behavior” as elements of balanced scorecards that determine bonus payments. “Clawbacks” of bonuses obtained by cheating can, in some circumstances, provide appropriate ex-ante incentives. Formal enforcement of corporate guidelines can also include the threat of termination of employment.

Another approach that corporations use to ensure honesty is to search for the “right” agents by trying to assess the “integrity” of candidates for executive and board member positions. Indeed, a large array of research has argued that some individuals possess preferences for truthfulness and, therefore, experience intrinsic costs of lying.¹

The focus of this paper is on a third approach: the role of “social norms.” This paper specifically considers *situational social norms*.² It has long been known that situational norms guide human action in direct and meaningful ways (Sherif, 1936; Asch, 1956; Milgram, 1974; Cialdini, Kallgren, and Reno, 1991). This power has also been used actively to steer behavior in several domains. For example, norm-based interventions have been shown to support pro-environmental behavior (e.g., Nolan, Schultz, Cialdini, Goldstein, and Griskevicius, 2008; Allcott and Rodgers, 2014) or responsible alcohol consumption. In the corporate world, too, there is increasing recognition of the potential of social norms, though there is little direct evidence regarding how norm-based interventions work. For example, survey evidence by Adam and Rachman-Moore (2004) suggests that when it comes to implementing codes of conduct, informal factors such as social norms are perceived to be far more effective than formal methods such as training programs and formal enforcement.³

Even when corporations recognize that social norms can support the implementation of an organization’s values or codes of conduct, it is usually (implicitly) assumed either that

¹See, among several others, Evans, Hannan, Krishnan, and Moser (2001), Gneezy (2005), Sánchez-Pagés and Vorsatz (2007), Lundquist, Ellingsen, Gribbe, and Johannesson (2009), Erat and Gneezy (2012), Gibson, Tanner, and Wagner (2013), and López-Pérez and Spiegelmann (2013).

²Social norms refer to injunctions on behavior that are sustained by the threat of social disapproval or penalties (Elster, 1989). Situational norms are only temporarily salient. They are more transitory than long-standing internalized norms, which are uniformly in force at all times and in all situations. Situational norms do not need to be internalized in order to be effective. We use the terms *social norms*, *situational social norms* and *situational pressure* interchangeably. Other terms could be used to describe this force: *normative influence* (Deutsch and Gerhard, 1955), or *exhortations*.

³In their analysis of 812 employees of an Israeli subsidiary of a US multinational company, 57.4% of respondents rated social norms as the most influential method in the process of implementation of organizational ethical values and rules. Only 1.5% considered means of enforcement as most influential. See Adam and Rachman-Moore (2004) for additional studies that question the ability of formal enforcement to foster ethical behavior and Kaptein and Schwartz (2007) for an overview of the use of codes of conduct.

agents will conform equally to social norms, or that responses may vary but are largely unrelated to personal values.

In this paper, we challenge this assumption and explore the notion that the use of norm-based interventions has to be mindful of the targeted individuals' characteristics. Moreover, corporations need to know who is least susceptible to “bad” social norms.⁴ As discussed in Section 2, the direction of a possible interaction of social norms and intrinsic preferences is not clear a priori. They may interact negatively (that is, as substitutes), or social norms may enhance the power of intrinsic preferences.

How can one assess the potentially heterogeneous effectiveness of “good” social norm interventions and the potentially also heterogeneous responses of individuals to “bad” norms? In reality, social norms could be endogenous, and there are likely other confounding factors. An analogy may be helpful here: Clinical trials seek to determine the effectiveness of new types of medicine against diseases. Similarly, a laboratory experiment can provide insights into how companies might address the “disease of dishonesty.”⁵

Our experiment is set in a concrete context, accounting earnings management, referred to in this paper as *earnings management*. In this experiment, earnings management is designed to be a form of lying, which is defined as making “a statement that one knows to be false” (Grover, 2005).⁶ While the experiment thus covers a setting in which participants as CEOs report to the outside market, a similar situation arises, for example, when a division head can choose how to manage information reported to the corporate center.

⁴Situational pressure has been shown to potentially induce contagion effects both within and among groups (e.g., Gino, Ayal, and Ariely, 2009; Innes and Mitra, 2013). Fischer and Huddart (2008) note that besides fostering productive effort, social norms also can support potentially destructive behavior. Their leading example focuses on earnings management.

⁵We discuss external validity in Section 6.

⁶Accounting earnings management occurs “when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that depend on reported accounting numbers” (Healy and Wahlen, 1999). In practice, there are also other motives for managing earnings, and competitive pressures among firms also play a role (Shleifer, 2004). We can exclude these factors in our experiment.

Section 3 describes the experimental method and design. Participants are cast in the role of CEOs and are told the truthful level of earnings. However, they are informed that they can legally announce higher earnings and receive higher bonuses. In this setting, we study the effects of injunctive norms, describing specific kinds of behaviors that meet with real or perceived social approval or disapproval. We randomly inform some participants that society disapproves of earnings management, while others are informed that society approves of earnings management.⁷ As a second situational feature, we also vary the economic incentives in favor of earnings management.

Our primary empirical contribution consists of evidence demonstrating how the effects of social norms vary with individuals' preferences. Using established scales (Tanner, Ryf, and Hanselmann, 2009), we measure the strength of individual commitments to honesty as a protected value. We mainly focus on "protected values: reactions to violations of honesty" (PRV), a measure of the degree to which individuals experience affective reactions when the value of honesty is violated or when the possibility of such a violation becomes salient (Tetlock, Kristel, Elson, Green, and Lerner, 2000).

We find that individuals with weak PRV conform more to social norms, whether those are dishonesty-approving or dishonesty-disapproving; individuals with strong PRV are more steadfast and less influenced by *both* types of norms. These findings are noteworthy. A simple alternative prediction would be that there is more room for honesty-loss among high-PRV individuals. Thus, the dishonesty-approving norm might have been expected to have stronger detrimental effects for those who are initially more likely to report the truth

⁷In contrast to injunctive norms, descriptive norms describe the percentage of individuals choosing a particular behavior; see Cialdini, Reno, and Kallgren (1990). The usage of descriptive norms appears to be rare inside organizations. In our experiment, following social norms does not bring instrumental benefits. Injunctive norms entailing explicit penalties can be even stronger (though such penalties are not required for a social norm to be effective). Krupka and Weber (2013) elicit injunctive social norms from one group of people and then use this as a common predictor for the behavior of another group of people. Our analysis differs in two respects: We exogenously vary the prevailing norm, and we are interested in variation in sensitivity of different people to injunctive norms.

and, thus, start from a higher level of inclination to tell the truth. Our results instead show the opposite: Stronger resistance to dishonesty-approving norms precisely among those who could have changed their behavior the most. Conversely, because at higher costs even high-PRV individuals frequently lie, they could have increased their truth-telling in response to the dishonesty-disapproving norm, but they did not. Similar results hold for pro-social concern (PSC) as a proxy of ethical motivations. We also document that experiment-participation effects are unlikely to be driving the results.

In Section 5, we propose a self-signaling model as an explanation of our findings. The model is based on the premise that individuals interpret their actions as signals to themselves of their own preferences for truthfulness (Bem, 1972; Bodner and Prelec, 2002; Bénabou and Tirole, 2004, 2006). Additional tests – for example, taking into account the beliefs that the agents have regarding the consequences of their actions – further support the self-signaling hypothesis. Still, we cannot rule out that part of the observed resistance against social norms is arguably also due to attempts of participants as CEOs to signal to the market.

The central empirical implication of our findings is that, due to the positive correlation of intrinsic preferences for truthfulness and resistance to extrinsic influences, the degree of steadfastness in the face of varying economic situations and social norms is an important source of information for market participants regarding the likely ethical preferences of managers. Our paper also has important implications for attempts to leverage the power of social norms in corporate governance. “Good” norms can be discouraging for otherwise strongly ethically motivated individuals, but on the bright side these individuals resist “bad” norms as well. Our analysis shows that, rather than designing social norms, calibrating incentives, and making hiring decisions separately, it is important to consider their interaction. We develop these implications further in Section 6.

2 Theoretical framework and hypotheses

We offer a fairly generic framework to develop the hypotheses. Our experiment is then set in a concrete (earnings management) context, but the framework could be applied to other corporate decisions as well.

2.1 Framework

Consider an individual i who decides whether to tell the truth, $T = 1$, or to lie, $T = 0$. For the time being, suppose the agent takes two factors into account in this decision: the agent-specific costs of lying, and the prevalent social norm. In the experiment, each participant faces the decision to announce the true earnings or to legally announce higher earnings, and she/he is exposed to one or the other social norm regarding this choice. (Moreover, in the experiment, we also introduce economic consequences of lying.)

On the one hand, individuals have intrinsic preferences for truthfulness. That is, they experience psychologically or morally driven costs of lying which are proportional to θ_i , the agent's ethical type (henceforth just "type"). More ethical types value truthfulness more.

On the other hand, situational social norms regarding honesty cause *situational norm-driven costs of stating the truth*, $SNCOST_s$, where s denotes the type of social norm. A dishonesty-disapproving social norm operates against dishonesty; thus, $SNCOST_{DISAPP} < 0$. Our analysis also applies to a dishonesty-approving social norm, implying $SNCOST_{APP} > 0$. A key element of these situational norm-driven costs of lying is that they are non-monetary: individuals may conform to norms because of threats of emotional penalties for norm violations and of disapproval, such as accusations that the agent does not understand the "rules of the game" (Cialdini and Goldstein, 2004).

The key ingredient of our analysis is that agents may differ in their susceptibility to social norms, that is, in their responses to $SNCOST_s$. Let ρ_i indicate how much agent

i resists social norms. The consequences the agent perceives when telling the truth are $SNCOST_s(1 - \rho_i)$. In sum, the global utility function is

$$V_{is}(T) = \begin{cases} -SNCOST_s(1 - \rho_i) & \text{if } T = 1 \\ -\theta_i & \text{if } T = 0. \end{cases} \quad (1)$$

We are in particular interested in whether the resistance parameter ρ_i is systematically associated with the agent’s type θ_i . For simplicity, suppose that $\rho_i(\theta_i) = r\theta_i$. Depending on whether $r > 0$ or $r < 0$ or $r = 0$, resistance to social norms is increasing in, decreasing in, or independent of the ethical type. A natural benchmark is to posit that for some agents, truthfulness is a Kantian (deontological) imperative, a “taboo value,” or a “sacred value,” meaning that the highest types, $\bar{\theta}$, are people who endorse values for truthfulness so completely that they resist all trade-offs (Baron and Spranca, 1997; Fiske and Tetlock, 1997). Formally, when $r = 1/\bar{\theta}$, the highest type does not react at all to social norms; when r is smaller (but still positive), all types resist to some extent to social norms.

2.2 Hypotheses

Our empirical approach is to collect data on proxies for the (unobservable true) type and then to see whether these proxies are both correlated with truth-telling (as they should be, if they proxy for the ethical type) *and* with resistance against social norms and economic incentives.

Our primary proxy is a measure of the extent to which agents suffer negative emotional consequences when the value of honesty is or may be violated, called *protected values for truthfulness associated with reactions to violations of honesty* (PRV).⁸

We now state the model in terms of the parameters we will be able to identify empiri-

⁸See Section 3 for more details on this concept and related concepts. Below, we also report results for other proxies, including a measure of pro-social concern (PSC).

cally. For simplicity, posit $\theta_i = PRV_i$. Thus, $\rho_i = r\theta_i = rPRV_i$.

From equation (1), combining and renaming coefficients, we can express the difference in utility between telling the truth and lying for individual i in economic situation e under situational social norm s as:

$$Y_{is}^* = b_0 + b_P PRV_i + b_{SNCO} SNCO_{sT} + b_{PSNCO} PRV_i SNCO_{sT}. \quad (2)$$

Expanding equation (2) by allowing dishonesty-approving ($SNCO_{sT} > 0$) and dishonesty-disapproving ($SNCO_{sT} < 0$) social norms to have different effects,

$$\begin{aligned} Y_{is}^* = & b_0 + b_P PRV_i + b_{ASN} SNCO_{sT} 1_{\{SNCO_{sT} > 0\}} + b_{DSN} SNCO_{sT} 1_{\{SNCO_{sT} < 0\}} \quad (3) \\ & + b_{PASN} PRV_i SNCO_{sT} 1_{\{SNCO_{sT} > 0\}} + b_{PDSN} PRV_i SNCO_{sT} 1_{\{SNCO_{sT} < 0\}}, \end{aligned}$$

where 1_{Ω} is an indicator term, indicating whether a dishonesty-approving social norm (abbreviated as ASN in the index) or a dishonesty-disapproving social norm (DSN), respectively, is in place.

Finally, in the experiment we will also introduce direct *economic costs of stating the truth*, $ECOST_e \geq 0$, where the index e denotes the ECOST situation. Because wealth effects are unlikely in our experiment, we posit linear utility for simplicity.

Incorporating ECOST into equation (3) and allowing for heterogeneous responses to economic incentives, we have

$$\begin{aligned} Y_{ise}^* = & b_0 + b_P PRV_i + b_{ASN} SNCO_{sT} 1_{\{SNCO_{sT} > 0\}} + b_{DSN} SNCO_{sT} 1_{\{SNCO_{sT} < 0\}} \quad (4) \\ & + b_{PASN} PRV_i SNCO_{sT} 1_{\{SNCO_{sT} > 0\}} + b_{PDSN} PRV_i SNCO_{sT} 1_{\{SNCO_{sT} < 0\}} \\ & + b_E ECOST_e + b_{PE} PRV_i ECOST_e. \end{aligned}$$

2.2.1 Direct impact of PRV

Naturally, we expect individuals with higher agent-specific costs of lying to perceive truthfulness as more attractive than lying. Thus, to the extent that PRV is a valid proxy for

the ethical type, we expect $b_P > 0$.

2.2.2 Hypothesis regarding direct responses to situational norms and economic incentives

A large literature (see the introduction) predicts that the pressure exerted by social norms leads individuals to conform to these norms. Social norms are hypothesized to trigger an internal mechanism by which truthfulness becomes more or less attractive. Accordingly:

Hypothesis CONFORM: *Truthfulness becomes less attractive under dishonesty-approving social norms and more attractive under dishonesty-disapproving social norms. Thus, $b_{ASN} < 0$ and $b_{DSN} > 0$.*

The alternative hypothesis, in line with Brehm's (1966) theory of reactance, suggests that agents tend to act in the opposite direction of what is suggested by the situational norm: $b_{ASN} > 0$ and $b_{DSN} < 0$.

Additionally, we expect that truthfulness becomes less attractive as the economic costs of truth-telling increase. Under our assumptions, $b_E = -1$.⁹

2.2.3 Hypothesis regarding heterogeneous responses to situational norms and economic incentives

One possible hypothesis is that the impact of social norms and economic incentives may be the same for everybody. Thus, social norms and incentives on the one hand and intrinsic preferences are separable. This would lead to $b_{PDSN} = 0$, $b_{PASN} = 0$ and $b_{PE} = 0$. Indeed, this is the implicit assumption under which many practitioners operate when they

⁹More generally, we expect b_E to be equal to minus the marginal utility of money. In addition, economic costs of truthfulness and the situational norm-driven costs may also affect preferences. If, for example, ECOST is positively related to costs of lying, there is a countervailing effect. We cannot identify these effects within our study. Taking this possibility into account, we expect the attractiveness of truthfulness not to increase in economic costs of stating the truth, that is, $b_E < 0$.

implement corporate behavioral rules.¹⁰

However, significant evidence suggests that economic incentives can undermine intrinsic preferences, though in several studies, in fact, complementarity occurred. See Bowles and Polanía-Reyes (2012) and Gneezy, Meier, and Rey-Biel (2011) for economic surveys and Deci, Koestner, and Ryan (1999) for a survey of the psychology literature.¹¹

In particular, extrinsic incentives, when perceived as *over-justification* of one’s actions, can reduce intrinsic motivations.¹² Bénabou and Tirole (2006) formalize this attributional theory in an economic model of self-signaling where agents discount their own pro-social motivations when incentives for pro-social behavior are in place. A related mechanism, proposed by motivational theories and also captured by the self-signaling framework, is the idea that individuals experience *control aversion*. We expect this to apply especially for those with strong PRV.¹³ Although the specific psychological channels of the attributional and motivational theories are different, they have similar implications for how changing extrinsic incentives leads to changes in behavior (Deci, Koestner, and Ryan, 1999). We hypothesize that analogous effects also apply to the differential susceptibility of individuals to social norms. Overall, these effects induce $b_{PDSN} < 0$.

¹⁰Note that the hypotheses are formulated in terms of predictions for the *coefficients* b_{PDSN} , b_{PASN} , and b_{PE} . Our empirical tests correspond to these quantities. In terms of *behavior* changes, for example, when a dishonesty-approving social norm is put in place, and if indeed $b_{PASN} = 0$, one would expect a bigger change towards dishonesty for those individuals starting with a tendency to tell the truth (which tend to be the high PRV types) than for individuals starting with a tendency to tell a lie (which tend to be the low PRV individuals).

¹¹Some papers refer to “*crowding-out*” and “*crowding-in*” in this context. Unfortunately, terminology varies. In Bowles and Polanía-Reyes (2012), *strong crowding-out* implies that the total effect of economic incentives, in fact, is opposite the direct effect. (Regular) *crowding-out* occurs already when economic incentives and intrinsic preferences interact such that the total effect of economic incentives is less than their direct effect (and, conversely, the total effect of intrinsic preferences is less than their direct effect). Other authors understand “*crowding-out*” to mean strong *crowding-out* only. Less ambiguously, *crowding-in* happens when the total effect of economic incentives is more than the direct effect. We avoid these semantic differences and speak of substitution and complementarity.

¹²This phenomenon was originally studied by Lepper, Greene, and Nisbett (1973).

¹³Control aversion was proposed by Deci (1975). The idea that heterogeneity in control aversion may be linked to PRV is consistent with individuals engaging in self-regulatory processes by which they control their behaviors so as to live up to their own intrinsic moral standards (Bandura, 1986; Aquino and Reed, 2002).

Importantly, our experiment also contains $SNCOST > 0$, and this allows us to test whether resistance occurs also against the dishonesty-approving situational norm. By resisting such a norm, an individual can prove her ethical values to herself.¹⁴ To the extent that symmetry holds, the substitution between intrinsic preferences for truthfulness and honesty norms actually has a bright side in that these intrinsic preferences will also make individuals less susceptible to dishonesty norms.

In sum, we investigate the following hypothesis:

Hypothesis SUBST: *High PRV types resist both types of situational norms. Thus, $b_{PDSN} < 0$ and $b_{PASN} > 0$.*

We also expect resistance against economic incentives, that is $b_{PE} > 0$.

As Bowles and Polanía-Reyes (2012) show, there are also studies that find complementarity of intrinsic preferences by extrinsic incentives. This would imply $b_{PE} < 0$. It is conceivable that social norms may also complement and strengthen intrinsic preferences. Thus, the alternative hypothesis is:

Hypothesis COMPL: *High PRV types conform more to both types of situational norms. Thus, $b_{PDSN} > 0$ and $b_{PASN} < 0$.*

3 Experimental method and design

We investigate the impact of changing social norms in a controlled experiment. We employ a laboratory experimental design used in prior work (Gibson, Tanner, and Wagner, 2013) (GTW). Thus, we start with data on participant choices in the absence of an explicit manipulation of situational norms. In this paper, we exploit data from an extension of the

¹⁴To see why a symmetric reaction is plausible, imagine that we measure “unethical” types and make lying the choice variable. Then, substitution would mean that the less ethical individuals would respond less to the dishonesty-approving situational norm. Therefore, when truth-telling is the choice variable, the sign flips, and we thus expect ethical types to respond less to the dishonesty-approving norm.

GTW experiment that manipulates social norms.

3.1 Participants and procedure

A total of 261 participants took part in this online experiment. We recruited participants from undergraduate classes at the University of Zurich (Switzerland). 50 percent of the participants were economics and finance students, 40 percent psychology students, and 10 percent students from other fields. 42 percent were women, and 58 percent were men (distributed across the various fields).

All participants were informed at the outset that their choices would remain unknown to the experimenters. Most participants received payment one week after the experiment. For this purpose, each participant received, before the experiment, a code, based on which the experimenter prepared an envelope containing the earnings. Participants received the sealed envelopes by indicating their personal codes.

The instructions are in the Supplementary Appendix. Participants were first asked to respond to a few demographic questions and to read some basic instructions. They were informed that they would individually receive a payment, CHF 8, for their completed participation in the study and an additional payment that depended on their decisions. After having demonstrated their understanding of the (unlabeled) tasks and of the rules of the experiment, the participants completed the four main parts of the experiment: 1) the earnings-management task (first without and then with social norms treatments), 2) the beliefs and manipulation check, 3) the effort task, and 4) the measurement of various controls and proxies for ethical types. The order of parts 3 and 4 was randomized. Finally, all the participants were paid. For simplicity, we describe the procedure for one of the randomized orders of tasks.

1) The earnings-management task consisted of two phases, in each of which the partici-

pants were given five choices. Specifically, the task consisted of the following steps: choices in the Neutral Phase (Phase N), social norm manipulation, and choices in the Social Norm Phase (Phase SN). The experiment is set as an anonymous decision-making situation, thereby excluding the complications that arise in sender-receiver (deception) games, such as differences in the strategic sophistication of players (Cai and Wang, 2006; Wang, Spezio, and Camerer, 2010).

In the earnings-management task within Phase N, each participant was placed in the situation of a CEO who had to announce earnings per share for the previous quarter. The variable component of the CEOs' salaries depended on the earnings they announced. Participants were also informed that the market currently anticipated the announcement of 35 cents per share as earnings, but that the true earnings were 31 cents per share. The participants were told that they could announce earnings of 35 cents per share while remaining within legal accounting limits and that the decision would be solely theirs. Therefore, risk preferences of participants did not matter, as their choices were not based on the trade-off between the expected benefits and costs of committing a crime. They were also informed that they would be paid an amount based on the CEO compensation (according to their decisions). This additional experimental payoff would be converted into real money at the rate of CHF 100,000 = CHF 0.5. Importantly, participants earned less when choosing to tell the truth. The participants were then told they would have to announce their financial statements that day.

The earnings-management task questionnaire follows, in five paired questions in randomized order, that is, participants had to choose between 31 and 35 cents per share:

Which earnings will you announce?

-- 31 cents per share – In this case, your compensation will be CHF 60,000 (CHF 0.30).

-- 35 cents per share – In this case, your compensation will be CHF 300,000 (CHF 1.50).

- 31 cents per share – In this case, your compensation will be CHF 120,000 (CHF 0.60).
- 35 cents per share – In this case, your compensation will be CHF 300,000 (CHF 1.50).
- 31 cents per share – In this case, your compensation will be CHF 180,000 (CHF 0.90).
- 35 cents per share – In this case, your compensation will be CHF 300,000 (CHF 1.50).
- 31 cents per share – In this case, your compensation will be CHF 240,000 (CHF 1.20).
- 35 cents per share – In this case, your compensation will be CHF 300,000 (CHF 1.50).
- 31 cents per share – In this case, your compensation will be CHF 300,000 (CHF 1.50).
- 35 cents per share – In this case, your compensation will be CHF 300,000 (CHF 1.50).

While these stakes may appear low, they are, in fact, substantial relative to the fixed fee and relative to the time needed to complete the task. Looking at all five dimensions in aggregate, the minimum, fixed payment out of the earnings-management task is CHF 4.5, and the maximum is CHF 7.5. In other words, the stakes of CHF 3 are 2/3 of the fixed compensation for the task, a substantial pay-behavior sensitivity (especially taking into account the fact that the task is short).

The novel feature of the present paper is that, in Phase SN, we analyze a manipulation that was introduced after Phase N, introducing social norms. We faced the following trade-off in the experimental design. On the one hand, social norms affect observed actions; this would support revealing participants' choices in the experiment. On the other hand, we wished to avoid experimenter-demand effects. We chose an intermediate approach that is consistent with the overall design of the experiment as a decision-making situation in a concrete context. Participants knew that the market was observing their actions as CEOs; however, participants knew the experimenters could not learn their individual choices as experimental subjects. Specifically, the participants were given a page to read that stated that their respective firms would likely be confronted with a good investment opportunity

the following year for the acquisition of another company. However, they would need the shareholders' approval for that project. At the shareholder meeting, they would have an opportunity to convince the shareholders of the soundness of this investment. These shareholders would be closely following the CEOs' earnings announcements as well as those of the competitors.

Then, the participants were randomly assigned to one of the following three groups, which were not labeled for the participants, and received the following group-specific information. This information was provided in a way which conveyed legitimacy and competence of the source of the information. In practice, companies refer to company-specific approved types of behavior, society's approval, or to peer approval. In ethical training programs reference is sometimes made to society's values because companies regard their reputation in society as a primary risk factor. Different framings may yield different average conformity. The three groups follow:

I. earnings-management-APPROVING social norm group:

"One evening, you are sitting with a friend of yours who is a financial analyst. He tells you that increasing reported earnings in order to meet market expectations meets with widespread societal approval."

II. earnings-management-DISAPPROVING social norm group:

"One evening, you are sitting with a friend of yours who is a financial analyst. He tells you that increasing reported earnings in order to meet market expectations meets with widespread societal disapproval."

III. CONTROL group

(No further information was provided beyond the common information.)

After this interlude, all the participants were again provided with the same set of five options as in Phase N, requiring them to choose to announce earnings of either 31 or 35

cents per share.

2) We then conduct beliefs and manipulation checks, discussed in the results further below. We also measured participants' levels of pro-social concern (PSC).

3) In the effort task, participants engaged in a simple calculation task. This task created a time lag between the earnings-management task and the measurement of protected values for truthfulness.

4) In the stage of measurement of various controls and proxies for ethical types, we assessed participants' levels of protected values for truthfulness and their tendencies towards impression management and self-deception.

The experiment lasted about 20 minutes on average (of which the calculation task naturally took the biggest portion). The average total payment, received anonymously (see above) by each participant, was slightly less than CHF 30.5.

3.2 Variables of interest

TRUTHFUL CHOICE. This represents the dependent variable in the earnings-management task, coded as a binary variable that takes on the value of 1 if a participant chose to announce earnings per share of 31 cents (the honest option), while it takes on the value of 0 if a participant announced 35 cents (the dishonest option). **TRUTHFUL CHOICE** thus measures the extent of truth-telling, that is, the lack of earnings management.

SOCIAL NORM. This is a between-participants variation of *SNCOST*. The experiment did not offer continuous levels of *SNCOST*, but instead used three discrete levels. We define three dummies, making Phase N the omitted category in the regressions. **CONTROL** is equal to 1 for all observations from Phase SN with no additional information, and to 0 otherwise. **APPROVING** is equal to 1 for all observations from Phase SN with the social norm of approval of earnings management, and to 0 otherwise. **DISAPPROVING** is equal

to 1 for all observations from Phase SN with the social norm of disapproval of earnings management, and to 0 otherwise.

ECOST. This is a within-participants variation. The economic costs of truthfulness variable represents the amount of money a participant forfeited by announcing 31 cents. The ECOST variable takes on values from CHF 0 to CHF 1.20 ($= 1.50 - 0.30$), in increments of 30 cents.

AGENT TYPE. We use three proxies for ethical agent types: the primary focus is on two proxies for a concern for and a commitment to truthfulness as a protected value. Moreover, in Section 4.4.1, we use a proxy for pro-social concern (PSC). Two distinct subscales of protected values for truthfulness were developed in Tanner, Ryf, and Hanselmann (2009). PRV (reactions to violations of honesty) primarily captures affective reactions to (real or anticipated) violations of honesty (Tetlock, Kristel, Elson, Green, and Lerner, 2000). PNT (no trade-off) captures the more cognitive notion of an individual’s unwillingness to consider trade-offs based on an economic cost-benefit analysis of choosing between truthfulness and lying (Baron and Spranca, 1997). The details are in Supplementary Appendix A.2.1. Both scales have high Cronbach’s alpha (0.9 and 0.75, respectively). Both scales take on values between 0 (for an individual with no protected values) and 6 (for an individual with maximum protected values). We standardize the scales to have means of zero and standard deviations of unity (interdecile ranges: -1.25 to +1.24 and -1.25 to +1.47, respectively). The correlation between the two scales is 0.5, indicating that they are related but not identical. As the primary analysis here is concerned with responses to social norms without monetary consequences, we begin by considering the role of PRV. By contrast, PNT is expected to be primarily relevant in explaining resistance to economic, material incentives.

There exist other, broader and less specific measures of moral motivations. In a separate experiment with strong economic incentives to lie (but without social norms interventions)

we have evaluated the ability of such broader measures to predict truth-telling behavior: the HEXACO Honesty-Humility Scale (Ashton and Lee, 2009), moral identity (Aquino and Reed, 2002), social value orientation (Van Lange, Otten, De Bruin, and Joireman, 1997), and stated religiousness. As reported in Supplementary Appendix A.3.1, although the measures are correlated, protected values emerge, controlling for all the other variables, as the most significant determinant of resistance against economic costs of stating the truth. It is conceivable that, despite this evidence regarding resistance against economic incentives, the other measures do predict resistance against social norms.

DEMOGRAPHIC AND OTHER VARIABLES. Sex is equal to 1 for female participants and to 0 for male participants. Age is equal to each participant’s age in completed years (interdecile range: 20 to 29 years). Economics is equal to 1 for economics and finance students and to 0 otherwise. Other is equal to 1 for students of other fields and to 0 otherwise. Psychology students are the omitted category in the regressions. We also collect data on whether participants had recently read newspaper articles regarding CEOs, whether they worked part-time, and whether they had investment experience.

4 Experimental results

4.1 Descriptive evidence

We first verified that our participants perceived the announcement of 31 cents as the honest action that led to a personal loss, while the opposite was true of the announcement of 35 cents (both $p < 0.001$). As expected, the assessment of 31 cents as the honest action was neither different in the APPROVING norm condition than in the CONTROL condition ($p = 0.79$) nor was it different in the DISAPPROVING condition than in the CONTROL condition ($p = 0.42$). The same holds for the perception of which action led to a personal

loss. This is consistent with the notion that the assessment of what is the honest option is different from the knowledge of what is approved of or disapproved of by society.

Table 1 allows a first look at the choices the participants made in the experiment. This table reveals substantial variation in responses throughout the within-participants and between-participants conditions that were established in the experiment. Of particular relevance to our purpose are the choices made under the different social norms.

At the median cost level, around 31%-34% of participants told the truth in the CONTROL condition (and Phase N); by contrast, only 16% reported the truth under the earnings-management APPROVING norm, while 55% stated earnings truthfully under the DISAPPROVING norm. In aggregate, leaving aside the ECOST = 0 case, in 33% of cases, participants opted to suffer monetary losses relative to what they could have earned. Even when there was no economic cost of truthfulness, 23% of the participants chose the earnings-management solution. These participants may have experienced negative costs of lying. Social norms also had a significant impact at zero ECOST. In line with previous research, we thus find powerful direct effects of both types of social norms.

Table 1
Behavior across phases and costs of truthtelling

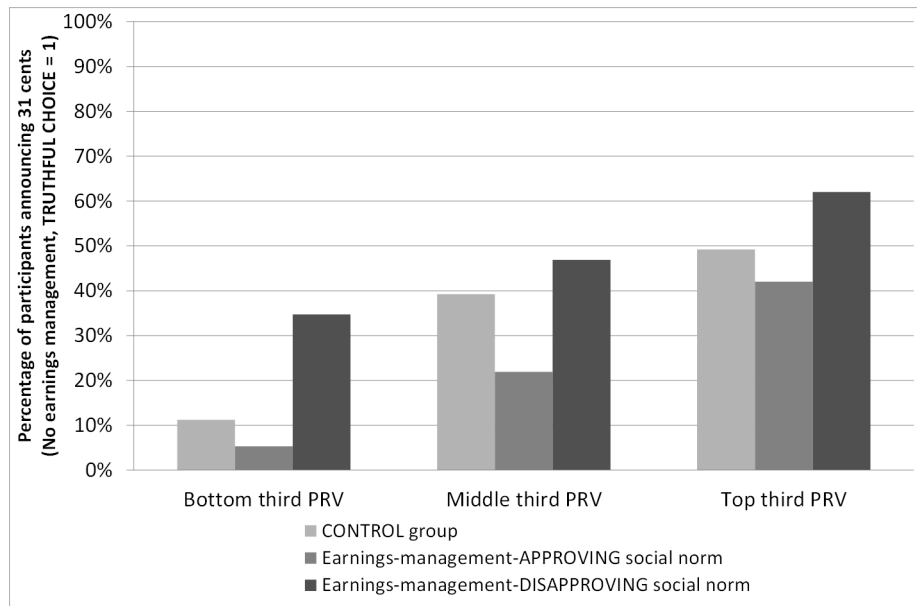
This table presents the percentages of participants announcing 31 cents of earnings per share (TRUTHFUL CHOICE = 1) across the various ECOST conditions and social norms (Phase SN) conditions of the experiment. Phase Neutral (N) contains data from Gibson, Tanner, and Wagner (2013) (GTW).

ECOST	Phase SN Data from social norm treatments			Phase N	All
	APPROVING	DISAPPROVING	CONTROL	GTW data	
	Percent of participants announcing 31 cents				
CHF 0	63.1%	80.7%	71.9%	82.0%	77.0%
CHF 0.3	33.3%	65.9%	49.4%	52.1%	51.0%
CHF 0.6	15.5%	54.5%	33.7%	31.4%	33.1%
CHF 0.9	15.5%	35.2%	27.0%	23.0%	24.5%
CHF 1.2	14.3%	31.8%	23.6%	21.1%	22.2%
Total	28.3%	53.6%	41.1%	41.9%	41.6%
Total except ECOST = 0	19.6%	46.9%	33.4%	31.9%	32.7%

Next, Figure 1 plots the percentages of truth-tellers across the three treatments of Phase SN (over all ECOST situations), using the CONTROL group as the reference point. To provide some first insights into heterogeneous responses, the figure considers separately, for each of the three PRV terciles, the behavior of participants.

Figure 1
Behavior across social norm conditions

This figure shows, for each of the three PRV (protected values for truthfulness, reactions to violations of honesty) terciles, the percentage of participants who announced 31 cents (TRUTHFUL CHOICE = 1) in Phase SN (social norms) of the experiment. The figure uses all ECOST situations.



While Table 1 shows strong effects of the APPROVING norm in the sample on average, Figure 1 demonstrates that, among the top-third PRV group, behavior was quite stable in the face of the APPROVING norm. This is striking, because it might have been expected that the effect of the APPROVING norm would be stronger for those who were more likely to tell the truth in phase 1.

And while the DISAPPROVING norm more than doubled the number of truth-tellers among the bottom-third PRV group, this norm had hardly any effect on behavior among the top-third PRV group. These results are not simply due to the fact that the percentages

of truth-tellers cannot exceed 100%. In the top PRV tercile, the percentage of truth-tellers in the CONTROL condition was in the 50% range (due to economic costs), so that significant behavioral changes were, in principle, possible under the disapproving norm. This descriptive evidence provides preliminary support for Hypothesis SUBST.

4.2 Empirical model

We estimate a discrete-choice / random-utility model (King, 1998; Wooldridge, 2006). Starting from equation (4), assuming a stochastic error with a logistic distribution (independent of the explanatory variables), and positing that agents choose the action that provides them with the higher utility, one obtains the logit model. We suitably adapt equation (4) by including dummy variables to identify the situational norm and control groups (see Section 3.2), making Phase N the omitted category. After relabeling coefficients,

$$\Pr(T_{ise} = 1|\mathbf{X}) = \Lambda \left[\begin{array}{c} \beta_0 + \beta_P PRV_i + \beta_{ASN} APPROVING_s + \beta_{DSN} DISAPPROVING_s + \\ \quad + \beta_{CSN} CONTROL_s + \beta_E ECOST_e + \\ + \beta_{PASN} PRV_i APPROVING_s + \beta_{PDSN} PRV_i DISAPPROVING_s + \\ \quad + \beta_{PCSN} PRV_i CONTROL_s + \beta_{PE} PRV_i ECOST_e \end{array} \right], \quad (5)$$

where $\Lambda(\bullet)$ is the logistic cumulative distribution function. The coefficient vector in Equation (5) is estimated by maximum likelihood. We cluster standard errors on the individual level.

The estimated coefficients are the implied estimates for the model parameters. We consider coefficients, rather than marginal effects, from the logit regressions in the subsequent analysis. Analyzing coefficients allows us to consider the counterfactual case in which participants would display identical initial probabilities of reporting the truth. (The highest marginal effects are expected to be found in the range of those participants with

medium initial probabilities of truthfulness.)

4.3 Regression results

4.3.1 Direct effects of situational norms and economic incentives

The regressions in Table 2 show that participants responded strongly either to society’s approval or to its disapproval of earnings management. These results support Hypothesis CONFORM, consistent with many prior studies, and they reject the idea of a uniform reactance against situational norms. These effects hold after controlling for PRV. Expressing as marginal effects the coefficients on the APPROVING and DISAPPROVING dummies implies that the approving norm made earnings management 15% more likely and that the disapproving norm made it 15% less likely.

The control group behaved about the same as in Phase N. This suggests that the information that their behavior would be closely observed by shareholders did not by itself change participants’ choices; it was the situational norm stating society’s approval (or disapproval) of earnings management that triggered a behavioral change.

As expected, the higher the economic incentives for earnings management were, the more likely participants were to manage earnings (see the negative coefficient on ECOST).¹⁵

4.3.2 Intrinsic preferences and heterogeneous responses to situational norms and economic incentives

We now consider the relevance of differences in preferences for truthfulness among individuals. As is evident from column (1) of Table 2, PRV is strongly significantly positively

¹⁵In the real world, managers are indeed faced with substantial cross-sectional and time-series variation in the economic cost of truth-telling. Our results are consistent with findings by Bergstresser and Philippon (2006), who showed that the use of discretionary accruals to manipulate reported earnings was more pronounced at firms where CEO compensation depended more on the stock price.

associated with the perceived attractiveness of truthfulness. Demographic variables are not systematically related to truthfulness.

Our main results, regarding who responds the most to social norms, referring to Hypotheses SUBST and COMPL, are reported in columns (2) and (3) of Table 2. Column (2) uses all ECOST situations. Column (3) only considers situations with positive ECOST.

We find a negative coefficient on the interaction PRV * DISAPPROVING. In other words, those with stronger PRV responded less to the dishonesty-disapproving social norm than those with weaker PRV. This occurred even though, as seen in Figure 1, the high-PRV individuals could also have increased their truthfulness. In the presence of the dishonesty-disapproving social norm, these intrinsic preferences were less important in guiding individuals towards truthfulness than in the absence of that norm.

Interestingly, we obtain a positive coefficient on the interaction term PRV * APPROVING. Thus, those with strong protected values were also steadfast in the face of a dishonesty-approving social norm. In other words, the effect of intrinsic preferences for truthfulness due to emotional reactions to violations of honesty was particularly pronounced when truthfulness was socially devalued. These findings starkly contrast with the notion that individuals who tend to be honest would be more affected by the dishonesty-approving norm simply because they have more “room” to turn towards dishonesty.

Recall that we have standardized the ethical type proxies. Because of the standardization, the coefficients shown for APPROVING represent the effects for a person of average PRV, in which case the standardized PRV score is zero, so that the interaction term with PRV cancels out. A person with PRV one standard deviation above the mean reacted only about half as strongly to an approving norm ($-0.68 + 0.39 = -0.29$ instead of -0.68) as the mean participant; in the range of positive ECOST, the substitution effect was stronger ($-0.82 + 0.67 = -0.15$).¹⁶ In the typical range of PRV, the total effect of norms still works

¹⁶As discussed by Bowles and Polanía-Reyes (2012), non-separability can be categorical or marginal;

Table 2

Main results: Testing hypotheses CONFORM, SUBST, and COMPL

This table presents coefficients of logit regressions. The dependent variable is TRUTHFUL CHOICE, which is equal to 1 when a participant chose to announce 31 cents and equal to 0 otherwise. The construction of Protected valued for truthfulness - reactions to violations of honesty (PRV) is described in the text. PRV is standardized to have a mean of zero and a standard deviation of unity. APPROVING is equal to 1 for all observations from Phase SN (social norm) with the social norm of approval of earnings management, and to 0 otherwise. DISAPPROVING is equal to 1 for all observations from Phase SN with the social norm of disapproval of earnings management, and to 0 otherwise. CONTROL is equal to 1 for all observations from Phase SN with no additional information, and to 0 otherwise. Phase Neutral (N) is the omitted category. Columns (1) and (2) use data from all ECOST situations. Column (3) considers situations where ECOST was strictly positive. Demographic controls (sex, age, fields of studies) are included; none have significant coefficients. T-statistics, obtained from robust standard errors clustered at the individual level, appear in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3) ECOST>0
APPROVING (social norm, Phase SN)	-0.64*** (-3.29)	-0.68*** (-3.54)	-0.82*** (-3.12)
DISAPPROVING (social norm, Phase SN)	0.59*** (3.37)	0.64*** (3.66)	0.77*** (4.46)
CONTROL group (Phase SN)	-0.16 (-0.88)	-0.22 (-1.21)	-0.11 (-0.53)
PRV (protected values for truthfulness)	0.69*** (5.48)	0.35*** (2.82)	0.37** (2.21)
PRV * APPROVING		0.39* (1.95)	0.67** (2.29)
PRV * DISAPPROVING		-0.37* (-1.93)	-0.43** (-2.23)
PRV * CONTROL		0.20 (1.05)	0.15 (0.72)
ECOST (cost of no earnings management)	-2.36*** (-13.97)	-2.51*** (-12.87)	-1.89*** (-8.62)
PRV * ECOST		0.64*** (3.36)	0.55*** (2.66)
Demographic controls (all insignificant)	Yes	Yes	Yes
Constant	1.59** (1.97)	1.46* (1.87)	0.45 (0.56)
Observations	2,610	2,610	2,088
Pseudo R^2	0.208	0.219	0.161
Pseudo Log Likelihood	-1403	-1384	-1107
Likelihood-ratio test statistic (χ^2 , p-value)	738 (<0.01)	776 (<0.01)	426 (<0.01)
Wald test statistic (χ^2 , p-value)	230 (<0.01)	237 (<0.01)	121 (<0.01)

in the same direction as the direct effect, that is, substitution is not complete. For the highest PRV-values, the direct effect of norms can even be overturned. PRV also induces significant resistance against economic costs.¹⁷

In sum, we find evidence in favor of hypothesis SUBST.

4.4 Further results

Before discussing a plausible explanation of our findings (see Section 5), we investigate some additional factors that could influence behavior in our experiment. In this section, we document that pro-social concern (PSC) is modestly related to resistance against norms (Section 4.4.1). The results are robust to controlling for marginal utility (Section 4.4.2). We also show that the experiment itself is unlikely to have had an effect on individuals' protected values that would distort our results (Section 4.4.3). Section 4.4.4 considers interactions between social norms and economic incentives. Section 4.4.5 reports on non-parametric and other robustness checks.

4.4.1 Pro-social concern

An alternative proxy for ethical preferences is pro-social concern. Significant evidence exists on the role of this motivation (Fehr and Fischbacher, 2002, 2003). To capture this facet, in an attempt to mirror reality, we are vague about the precise consequences for others and use participants' answers regarding the extent to which they believed that announcing 35 cents had negative consequences for some stakeholders or was manipulative. The details are in Supplementary Appendix A.2.2. We caution that this variable measures

that is, there may be a distinct effect of introducing any positive amount of ECOST. The findings here also suggest that the mere presence of monetary stakes has categorical effects on the role of intrinsic preferences and on their interplay with situational norms.

¹⁷Gibson, Tanner, and Wagner (2013) also find, in Phase N, that agents with strong protected values react less to economic costs of stating the truth. They do not discuss why this substitution effect arises.

how strongly participants assess the externality, but not explicitly how much participants would give up to avoid it. Participants who exhibited stronger pro-social concern score high on the resulting variable PSC. PSC is standardized to have a mean of zero and a standard deviation of unity (interdecile range: -1.57 to +1.16). PRV and PSC have a correlation of 0.39. PNT and PSC have a correlation of 0.28. We orthogonalize PRV and PSC, but similar results also hold if we include the main measures. Regression (1) of Supplementary Appendix Table A.5 shows that PSC's results parallel those of PRV. Thus, substitution, rather than complementarity of pro-social concern and social norms for honesty is documented in our experiment. Once we include PRV into the regression, the significance of PSC in explaining heterogeneity in responses is diminished. This is consistent with the design of the experiment, in which strategic and pro-social motivations are less likely to play a role.

4.4.2 Marginal utility of money

It is possible that the experiment or the situational norms manipulation differentially activated payoff-maximizing modes of thought. In an attempt to address this possibility, we control for a self-reported measure of marginal utility, MU (see Supplementary Appendix A.2.4 for details). In results available on request, we find that (1) MU is not correlated with PRV (correlation of -0.04, p-value of 0.48), (2) MU does not differ across the situational norms manipulations (Kolmogorov-Smirnov test p-value of 0.9 when comparing the MU of those who went through the approving-norm treatment and those who went through the disapproving-norm treatment), and (3) MU is insignificant (t-values of between -0.6 and -0.4) in the regressions (and all other results remain unchanged).

4.4.3 Experiment-participation effects

A potential concern as regards our findings is that the experiment may have affected participants' answers on the protected values survey. Either the experiment itself or the situational-norms manipulation may conceivably have played a role.

Collectively, three empirical observations mitigate this concern. First, we conducted a separate survey with 123 economics students who did not participate in any part of this experiment (the non-participants). We only measured the protected values of the students in this sample, and we did not involve them in any of the choice tasks. The means of the two PRV distributions of economics-student participants and non-participants are not statistically different (p-value of 0.24). Before rescaling, the average PRV of economics-student participants (non-participants) was 3.87 (3.68) while the 10th, 25th, 75th, and 90th percentiles were 2.6 (2.2), 3 (2.8), 4.8 (4.6), and 5.2 (5.6), respectively. A Kolmogorov-Smirnov test does not reject the hypothesis that the two PRV distributions are the same (p-value of 0.4).

Second, PRV was measured with a time lag after an interim (effort) task. This mitigates the concern that participants tried to answer the protected values survey consistently with their choices on the earnings-management task.

Third, a Kolmogorov-Smirnov test does not reject the hypothesis that the PRV distributions of those who went through the approving-norm treatment and those who went through the disapproving-norm treatment are the same (p-value of 0.26). And given that we control for situational norms in the regressions, any effect of the situational norms would be accounted for.

4.4.4 Interactions between situational norms and economic incentives

The role of economic costs may vary as the norms vary. In particular, Fischer and Huddart (2008) derive a model in which social norms augment the effects of incentives. Supplementary Appendix A.3.4 states the corresponding hypothesis for our context and presents results in line with the hypothesis that social norms and economic costs of truthfulness are complements. Hypothesis SUBST continues to be supported when the interaction between situational norms and economic incentives is accounted for.

4.4.5 Non-parametric tests and other robustness tests

We have considered many variations of the analysis, including non-parametric tests. These tests are briefly reported in Supplementary Appendix A.3.3.

5 Interpretation

We have found that individuals with strong intrinsic preferences for truthfulness, as proxied for by PRV, lean against both kinds of social norms. But what may be behind this two-edged resistance? Answering this question is important for the corporate governance implications one can draw from the findings of this intervention experiment.

One useful approach for understanding behavior in our experiment can be found in a self-signaling framework. For example, Bénabou and Tirole (2006) formalize the idea that, in the presence of extrinsic incentives, one’s action may be a less valuable signal of one’s own true preferences. As noted in Section 2.2.3, this model provides an account of both over-justification and control aversion.¹⁸

¹⁸Self-signaling can help build an identity, which can be an “asset” (Bénabou and Tirole, 2011). See Bodner and Prelec (2002) and Bénabou and Tirole (2004) for the first self-signaling models. By studying self-signaling, we add to the small experimental literature on the subject (Grossman and van der Weele, 2013; van der Weele and von Siemens, 2014; Grossman, 2012). Mazar, Amir, and Ariely (2008) argue that

In Supplementary Appendix A.1 we propose a model in this spirit with the novel feature that we allow for resistance to norms to be related to intrinsic preferences. Agents, when reflecting on their prior choices in order to infer their own true (ethical) types, understand that their actions depend both directly on their intrinsic preferences and on their responses to situational norms (and economic incentives). We prove that highly ethical types can, by resisting situational norms (and economic incentives) more than others, credibly self-signal their identities as truthful or pro-socially oriented individuals, respectively. Formally, a positive correlation of types and resistance ($r > 0$ in the notation of Section 2.1.2) supports a self-signaling equilibrium.

Thus, the results obtained so far are consistent with the self-signaling hypothesis. Additional tests arise from the observation that self-signaling can only occur regarding agent characteristics for which the action in question actually is an informative signal.

On the one hand, if agents wish to self-signal their identity as non-consequentialists, reporting the truth is an informative signal when money is at stake, but it is not informative when no money is at stake. Recall that PNT measures the extent to which individuals regard truthfulness as priceless and beyond the scope of an economic cost-benefit analysis (Baron and Spranca, 1997). Thus, we expect that $b_{AE} > 0$ for PNT. By contrast, in our experiment social norms do not bring about instrumental benefits. Thus, economic cost-benefit considerations are not directly applicable when it comes to situational norms. Hence, after observing their own actions, the agents' posterior type estimates are equal to their priors, that is, they cannot draw any inferences regarding their identity as non-consequentialists from their responses to such situational norms. Therefore, if it is self-signaling that drives people towards truthfulness, we would expect $b_{AASN} = b_{ADSN} = 0$ for

people may, at least to some extent, behave truthfully because they have a desire to maintain their self-concepts as honest persons (see also Fischbacher and Heusi (2013)). While self-deception is at the core of the Mazar, Amir, and Ariely (2008) analysis, the economic framework of self-signaling builds on Bayesian signaling. In results available on request, we find that the previous results are robust when tendencies for self-deception (see Supplementary Appendix A.2.3 for the measurement) are accounted for.

PNT. We find evidence in line with these predictions: The coefficient on PNT * ECOST is 0.79 (t-value: 4.48). The coefficients on PNT * APPROVING, PNT * DISAPPROVING, and PNT * CONTROL are all insignificant, with coefficients (t-values) of 0.23 (1.19), 0.11 (0.53), and 0.11 (0.65), respectively. (The full tables are available on request.)

On the other hand, no self-signaling can take place with respect to observable individual characteristics, in particular, demographics such as gender. Thus, these characteristics should not interact with situational norms. As is documented in the Supplementary Appendix, Table A.7, men and women responded similarly to situational norms,¹⁹ as did younger and older participants. Also, the extent to which participants had previously read newspaper articles about CEOs did not interact with situational norms in determining their truthful choices. Similarly, economics students did not respond to situational norms differently from psychology students, nor did participants with investment experience respond differently from those without investment experience.

While the results are, therefore, consistent with self-signaling, recall that in phase SN participants also received information regarding the fact that shareholders would be closely following their earnings announcements. This opens the possibility that to some extent participants were signaling to the market or to society more broadly. Such signaling is also supported by a positive correlation between the ethical type and the resistance to social norms. Two additional tests, however, provide further support for the self-signaling hypothesis.

First, column (1) in Table 3 shows that already in phase N, hypothesis SUBST was supported. Indeed, if anywhere, it is likely that self-signaling was strongly at play in that first phase. In phase SN, participants generally responded somewhat less to ECOST

¹⁹Our finding that preferences for truthfulness of both women and men are stable across situational norms is of interest, as other work suggests that women's social preferences are more malleable by context than men's (Croson and Gneezy, 2009). Some of our regressions suggest that, while women and non-investors told the truth more on average, they responded more strongly to economic incentives but this evidence is not robust across specifications.

overall, but this change is not systematically found in the high or low PRV individuals, as seen in the triple interaction term $PRV * ECOST * Phase\ SN$. These results implicitly suggest that the self-signaling motive was present throughout the experiment.

Second, in column (2) of Table 3, we include as additional explanatory variables the beliefs that participants stated regarding the determinants of project approval by shareholders (see the instructions for details).²⁰ We find that participants who believed that project approval after Phase SN would depend on whether they had always announced high earnings were more likely to report 35 cents of earnings, whereas those who believed that project approval depended on transparency were more likely to report the true earnings. Perceived competence played no role in the decisions. Importantly, our findings regarding Hypotheses CONFORM and SUBST remain robust. Moreover, columns (3) and (4) show that even allowing beliefs to differentially affect the responses to social norms does not change the results.²¹

Although the presented results yield a consistent picture in favor of the self-signaling hypothesis, we note that it is not possible to definitively rule out other channels. (1) Participants might have had an interest in pleasing the experimenter by appearing honest and non-greedy (Fischbacher and Heusi, 2013). The anonymity, relative to the experimenters, makes it unlikely that this was a major factor in our experiment. However, some researchers have argued that social esteem may, in fact, play a role even in anonymous settings (Ellingsen and Johannesson, 2008). We can, at least to some extent, control for these factors by including a measure of individuals' tendencies towards impression manage-

²⁰In untabulated results, we also included the information whether participants had recently read a newspaper article regarding CEOs. This variable controls, to some extent, for the possibility that behavior is driven by what participants thought that CEOs would do in such situations. The results remain robust.

²¹Because these beliefs refer directly to information given in phase SN, this analysis is restricted to phase SN data. Indeed, in untabulated results, we confirm that beliefs explain phase SN behavior much more significantly than phase N behavior. Beliefs and PRV are largely uncorrelated, except that there is a (statistically still insignificant) tendency for those who believe that earnings mattered for project approval to have lower PRV. Orthogonalizing PRV with respect to the beliefs does not change the results.

Table 3

Additional results: Phase SN, beliefs and impression management

This table presents coefficients of logit regressions. The dependent variable is TRUTHFUL CHOICE. For the main explanatory variables, see the notes to Table 2. Participants could indicate, with yes/no answers, that they believed project approval would depend on whether they had presented only high earnings (Belief-earnings), on how high their compensation was (Belief-compensation), on whether they were seen as competent (Belief-competence), and on whether they had reported transparently in the past quarters (Belief-transparency). The regressions consider ECOST > 0 situations. T-statistics, obtained from robust standard errors clustered at the individual level, appear in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Phases:	(1)	(2)	(3)	(4)	(5)
BELIEF interaction variable:	N+SN	SN	SN	SN	N+SN
			Earnings	Transparency	
APPROVING (social norm, Phase SN)		-0.79** (-2.01)	-0.91* (-1.76)	-0.05 (-0.05)	-0.96*** (-3.27)
DISAPPROVING (social norm, Phase SN)		0.76** (2.44)	0.71* (1.79)	-0.17 (-0.20)	0.79*** (4.55)
PRV * APPROVING		0.86* (1.65)	0.78 (1.55)	0.83* (1.69)	0.50* (1.65)
PRV * DISAPPROVING		-0.75** (-2.13)	-0.76** (-2.10)	-0.76** (-2.15)	-0.40** (-2.04)
Belief-earnings (1: yes, 0: no)		-1.60*** (-5.87)	-1.09** (-2.13)	-1.60*** (-5.81)	
Belief-compensation (1: yes, 0: no)		-0.25 (-0.69)	-0.26 (-0.70)	-0.22 (-0.62)	
Belief-competence (1: yes, 0: no)		-0.42 (-0.61)	-0.41 (-0.58)	-0.44 (-0.66)	
Belief-transparency (1: yes, 0: no)		1.56*** (3.11)	1.55*** (3.06)	0.76 (0.81)	
BELIEF * APPROVING			0.31 (0.42)	-0.84 (-0.83)	
BELIEF * DISAPPROVING			0.08 (0.14)	0.99 (1.08)	
BELIEF * ECOST			-0.87* (-1.81)	1.37 (1.12)	
PV (Protected Values)	0.38** (2.22)	0.48 (1.54)	0.45 (1.45)	0.44 (1.43)	0.39** (2.29)
ECOST (cost of no earnings management)	-1.99*** (-8.13)	-2.15*** (-7.70)	-1.76*** (-5.16)	-3.46*** (-2.85)	-1.95*** (-8.76)
ECOST * PRV	0.56** (2.35)	0.69*** (2.59)	0.77*** (2.80)	0.74*** (2.70)	0.51** (2.44)
Phase SN	-0.15 (-1.14)				
PRV * Phase SN	0.08 (0.69)				
ECOST * Phase SN	0.36** (2.08)				
PRV * ECOST * Phase SN	-0.15 (-0.84)				
CONTROL group (Phase SN)					-0.15 (-0.72)
PRV * CONTROL					0.18 (0.81)
IMPRESS					-0.00 (-0.00)
IMPRESS * APPROVING					0.78*** (2.97)
IMPRESS * DISAPPROVING					-0.19 (-0.99)
IMPRESS * CONTROL					0.07 (0.37)
IMPRESS * ECOST					0.31* (1.73)
Demographic control variables (insignificant)	Yes	Yes	Yes	Yes	Yes
Constant	0.59 (0.75)	1.12 (0.86)	0.93 (0.69)	1.80 (1.06)	0.61 (0.76)
Observations	2,088	1,044	1,044	1,044	2,088
Pseudo R-squared	0.132	0.295	0.297	0.301	0.179

ment, IMPRESS (see Supplementary Appendix A.2.3). While there is some evidence that IMPRESS interacted with situational norms, column (5) in Table 3 shows that our results regarding Hypotheses CONFORM and SUBST are robust even when controlling for this factor.²² In unreported results, triple interactions between PRV, IMPRESS, and the social norms are insignificant, further strengthening the conclusion that impression management does not explain the resistance of high-PRV types against the social norms. Also, (2) given that the situational norms manipulation did not change the payoff structure of the situation, and given the findings regarding the insignificant impact of the norms manipulation on self-reported marginal utility (see Section 4.4.2), we regard it as rather unlikely that moral disengagement (that is, an activation of people’s own payoff-maximizing modes of thought) was a primary driving force. Finally, (3), it is conceivable that the norms treatments induced stronger normative certainty for those with low PRV. However, the fact that PRV is a scale capturing predispositions or attitudes (not different levels of uncertainty) and the evidence reported in Section 4.4.3 that the PRV distributions of those who went through the approving-norm treatment and those who went through the disapproving-norm treatment are the same make this interpretation unlikely.

Overall, the rich set of evidence supporting hypothesis SUBST obtained in this experiment can be convincingly explained with the self-signaling model, though we cannot rule out that some degree of social or market signaling was also partly driving the observed resistance against social norm of high PRV individuals.

²²If anything, those with a stronger tendency to impress others by adhering to the social norm responded less to the earnings-management-disapproving norm; this is against what one would expect if impression management was driving responses to social norms. The interaction term is insignificant, though. There does appear, however, to be significantly stronger resistance to the approving norm among participants who are more prone to impression management.

6 Implications and conclusion

Cast in the broadest terms, this study presents an example of how, starting from a misbehavior, the effectiveness of (social norm) interventions can be evaluated. We conduct a simple, anonymous, and non-strategic earnings-management experiment using actual monetary incentives to lie. We present the consequences of a “clinical” intervention, varying social norms, representing either the approval or the disapproval of earnings management by society. This controlled experiment allows us to cleanly study the linking of reactions to these social norms with individuals’ intrinsic preferences. We find substantial evidence that those with strong intrinsic preferences for truthfulness due to emotional reactions to violations of honesty (high PRV individuals) react less to both kinds of social norms: On the one hand, they respond less to the “good” norm intervention that disapproves of dishonesty; on the other hand, they are also less susceptible to the “bad” social norm that approves of dishonesty.

Experimental simulation of corporate decision-making raises the question of external validity. Addressing this issue, several studies have shown that the levels of payments received by participants do not have a dramatic effect on their behavior if the subjects are paid proportionately to the opportunity cost of their time.²³ In addition, the available scientific evidence and the conclusions of evolutionary psychology suggest that the findings from our experiment have not only local validity but also relevance to similar situations in a wide range of settings.²⁴ Finally, we do not find that students of economics and

²³See, e.g., Davis and Holt (1992); Hoffman, McCabe, and Smith (1996); Forsythe, Horowitz, Savin, and Sefton (1991).

²⁴In particular, this research in evolutionary psychology shows that similar ecologically relevant problems are solved in a roughly similar way by most humans (controlling for their individual characteristics, such as the strengths of their values). Numerous studies find that the behavior of professional decision makers does not qualitatively differ from that exhibited by student subject groups, among them (DeJong, Forsythe, and Uecker, 1988; Dyer, Kagel, and Levin, 1989; Sade, Schnitzlein, and Zender, 2006; Smith, Suchanek, and Williams, 1988). See also Gillette, Noe, and Rebello (2008). If anything, real CEOs have been found to be more trusting (as well as trustworthy) and, thus, less conforming to the economic model of self-interest than students (Fehr and List, 2004).

students of psychology behave significantly differently in our experiment, even though their educational backgrounds differ quite substantially. Despite these pieces of evidence, we cannot ascertain that student participants behave identically to corporate executives.

Keeping this caveat in mind, we believe the following theoretical, empirical, and practical implications of our findings deserve mention.

In terms of theory, our findings hold implications for modeling the impact of social norms,²⁵ and they suggest that self-signaling can be a useful framework for understanding managerial responses to social norms (and explicit incentives).

For empirical work, the basic flavor of our analysis, in line with standard economic theory, is that actions can provide an important source of inference regarding the type of the observed agent. Subject to the caveat that earnings management may be a noisy measure of ethical behavior, CEOs whose earnings management behavior remains steady even as economic incentives and social norms in favor of or against earnings management change should be inferred to be more committed to honesty. Empirical work should investigate whether the market attributes stronger credibility to them and, for example, reacts more strongly when they make announcements regarding the outlooks of their companies.

From a practical point of view,²⁶ the results confirm the potential importance of social norms for corporate governance, but highlight important caveats. Specifically, first, it is individuals with relatively low intrinsic ethical motivations for whom injunctive social norms prove to be a powerful guiding beacon. Second, corporations have to bear in mind that social norms (and/or incentives) aimed at fostering ethical behavior may undermine or

²⁵Ostrom (2000) and Andreoni and Bernheim (2009) argue that individuals care heterogeneously about norm conformity. Formally, the sensitivity of individuals to social norms in models such Fischer and Huddart (2008), Burks and Krupka (2012), Huck, Kübler, and Weibull (2012), and Sliwka (2007) would depend on personal norms and preferences, whether the setting is one that emphasizes honesty or fairness choices.

²⁶We thank Anke D'Angelo (Chief Compliance Officer of Swiss Re), Robert Kuipers (Lead Partner Reward PwC Switzerland), and Maurice Zufferey (CEO Spencer Stuart Switzerland) for discussions on these issues.

even eliminate intrinsic motivations. However, the bright side of the substitution effect is that when the situational norm does approve of dishonesty, those with a strong commitment to honesty will resist that “bad” norm.

Third, the findings suggest a caveat regarding uniform “best practice” recommendations in the use of corporate social norms as a tool of corporate governance. A “one size fits all” approach to the use of social norms and ethical exhortations to support the implementation of codes of conduct can backfire. In practice, cultural differences (e.g., Asia vs. Europe) are often recognized as important. Our results instead highlight that even within a fairly homogenous culture, there is great, systematic heterogeneity in how individuals react to situational norms.

Fourth, navigating the interplay between social norms, incentives, and intrinsic preferences is tricky. In their theoretical analysis, Fischer and Huddart (2008) highlight the interaction effects between social norms and incentives. Our work instead emphasizes that to get the full return on the investment into finding the “right” people, into incentives, and into social norms, corporations need to be aware of the triangle of interactions between the three behavior drivers.

Fifth, the results raise the question of how to design and implement social norm interventions appropriately. In light of control aversion and over-justification effects for the highly intrinsically motivated, it is important that norms are designed such that they convey information and induce those who comply with the norms to conclude that they have high moral competence, rather than framing the norms as instruments of control. Exploring how and why social norms interventions (and more direct extrinsic incentives) are perceived to have different meanings in different corporations (and even business units) is, therefore, an important area for further research.

References

- Adam, Avshalom M., and Dalia Rachman-Moore, 2004, The methods used to implement an ethical code of conduct and employee attitudes, *Journal of Business Ethics* 54, 225–244.
- Ajzen, Icek, and Martin Fishbein, 1980, *Understanding Attitudes and Predicting Social Behavior* (Prentice-Hall: Englewood Cliffs).
- Allcott, Hunt, and Todd Rodgers, 2014, The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservations, *American Economic Review* 104, 3003–3037.
- Andreoni, James, and B. Douglas Bernheim, 2009, Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects, *Econometrica* 77, 1607–1636.
- Aquino, Karl, and Americus Reed, 2002, The self-importance of moral identity, *Journal of Personality and Social Psychology* 83, 1423–1440.
- Asch, Solomon E., 1956, Studies of independence and conformity: a minority of one against a unanimous majority, *Psychological Monographs* 70, 1–70.
- Ashton, Michael C., and Kibeom Lee, 2009, The HEXACO-60: A short measure of the major dimensions of personality, *Journal of Personality Assessment* 91, 340–345.
- Bandura, Albert, 1986, *Social foundations of thought and action* (Prentice-Hall: Englewood Cliffs).
- Baron, Jonathan, and Mark Spranca, 1997, Protected values, *Organizational Behavior and Human Decision Processes* 70, 1–16.
- Bem, Daryl J., 1972, *Self-Perception Theory* vol. 6 . pp. 1–62 (McGraw-Hill).
- Bénabou, Roland, and Jean Tirole, 2004, Willpower and personal rules, *Journal of Political Economy* 112, 848–886.
- , 2006, Incentives and prosocial behavior, *American Economic Review* 96, 1652–1678.
- , 2011, Identity, morals, and taboos: Beliefs as assets, *Quarterly Journal of Economics* 126, 805–855.

- Bodner, Ronit, and Drazen Prelec, 2002, Self-signaling and diagnostic utility in everyday decision making, in Isabelle Brocas, and Juan D. Carrillo, ed.: *Collected essays in psychology and economics* . pp. 105–126 (Oxford University Press: Oxford).
- Bowles, Samuel, and Sandra Polanía-Reyes, 2012, Economic incentives and social preferences: Substitutes or complements?, *Journal of Economic Literature* 50, 368–425.
- Brehm, Jack, 1966, *A theory of psychological reactance* (Academic Press: New York).
- Burks, Stephen V., and Erin L. Krupka, 2012, A multimethod approach to identifying norms and normative expectations within a corporate hierarchy: Evidence from the financial services industry, *Management Science* 58, 203–217.
- Cai, Hongbin, and Joseph Tao-Yi Wang, 2006, Overcommunication in strategic information transmission games, *Games and Economic Behavior* 56, 7–36.
- Capco, 2003, Understanding and mitigating operational risk in hedge fund investments, *White paper*.
- Cialdini, Robert B., and Noah J. Goldstein, 2004, Social influence: Compliance and conformity, *American Review of Psychology* 55, 591–621.
- Cialdini, Robert B., Carl A. Kallgren, and Raymond R. Reno, 1991, A focus theory of normative conduct, *Advances in Experimental Social Psychology* 24, 201–234.
- Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren, 1990, A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places, *Journal of Personality and Social Psychology* 58, 1015–1026.
- Croson, Rachel, and Uri Gneezy, 2009, Gender differences in preference, *Journal of Economic Literature* 47, 1–27.
- Davis, Douglas D., and Charles A. Holt, 1992, *Experimental Economics* (Princeton University Press: Princeton).
- Deci, Edward L., 1975, *Intrinsic Motivation* (Plenum: New York).
- , Richard Koestner, and Richard M. Ryan, 1999, A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation, *Psychological Bulletin* 125, 627–668.

- DeJong, Douglas V., Robert Forsythe, and Wilfred C. Uecker, 1988, A note on the use of businessmen as subjects in sealed offer markets, *Journal of Economic Behavior and Organization* 9, 87–100.
- Deutsch, Morton, and Harold B. Gerhard, 1955, A study of normative and informational influence on individual judgment, *Journal of Abnormal and Social Psychology* 51, 629–636.
- Dyck, Alexander, Adair Morse, and Luigi Zingales, 2010, Who blows the whistle on corporate fraud?, *Journal of Finance* 65, 2213–2253.
- , 2014, How pervasive is corporate fraud?, *Working paper*.
- Dyer, Douglas, John H. Kagel, and Dan Levin, 1989, A comparison of naive and experienced bidders in common value offer auctions: A laboratory analysis, *Economic Journal* 99, 108–115.
- Ellingsen, Tore, and Magnus Johannesson, 2008, Pride and prejudice: The human side of incentive theory, *American Economic Review* 98, 990–1008.
- Elster, Jon, 1989, Social norms and economic theory, *Journal of Economic Perspectives* 3, 99–117.
- Erat, Sanjiv, and Uri Gneezy, 2012, White lies, *Management Science* 58, 723–733.
- Evans, John H., R. Lynn Hannan, Ranjani Krishnan, and Donald V. Moser, 2001, Honesty in managerial reporting, *Accounting Review* 76, 537–559.
- EY, 2009, European Fraud Survey 2009, *Available at www.ey.com*.
- Fehr, Ernst, and Urs Fischbacher, 2002, Why social preferences matter, *Economic Journal* 112, C1–C33.
- , 2003, The nature of human altruism, *Nature* 425, 785–791.
- Fehr, Ernst, and John A. List, 2004, The hidden costs and returns of incentives - trust and trustworthiness among CEOs, *Journal of the European Economic Association* 2, 743–771.

- Fischbacher, Urs, and Franziska Heusi, 2013, Lies in disguise: An experimental study on cheating, *Journal of the European Economic Association* 11, 525–547.
- Fischer, Paul, and Steven Huddart, 2008, Optimal contracting with endogenous social norms, *American Economic Review* 98, 1459–1475.
- Fiske, Alan Page, and Philip E. Tetlock, 1997, Taboo tradeoffs: Reactions to transactions that transgress spheres of exchange, *Political Psychology* 17, 255–294.
- Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton, 1991, Fairness in simple bargaining experiments, *Games and Economic Behavior* 6, 347–369.
- Gibson, Rajna, Carmen Tanner, and Alexander F. Wagner, 2013, Preferences for truthfulness: Heterogeneity among and within individuals, *American Economic Review* 103, 532–548.
- Gillette, Ann B., Thomas H. Noe, and Michael J. Rebbello, 2008, Board structures around the world: An experimental investigation, *Review of Finance* 12, 93–140.
- Gino, Francesca, Shahar Ayal, and Dan Ariely, 2009, Contagion and differentiation in unethical behavior, *Psychological Science* 20, 393–398.
- Gneezy, Uri, 2005, Deception: The role of consequences, *American Economic Review* 95, 384–394.
- , Stephan Meier, and Pedro Rey-Biel, 2011, When and why incentives (don't) work to modify behavior, *Journal of Economic Perspectives* 25, 191–210.
- Grossman, Zachary, 2012, Self-signaling versus social-signaling in giving, *Working paper*.
- , and Joël van der Weele, 2013, Self-image and strategic ignorance in moral dilemmas, *Working paper*.
- Grover, Steven L., 2005, The truth, the whole truth and nothing but the truth: The causes and management of workplace lying, *Academy of Management Executive* 19, 148–157.
- Guiso, Luigi, Paolo Sapienza, and Luigi Zingales, 2008, Trusting the stock market, *Journal of Finance* 63, 2557–2600.

- Gurun, Umit, Noah Stoffman, and Scott Yonker, 2015, Trust busting: The effect of fraud on investor behavior, *Working paper*.
- Healy, Paul M., and James M. Wahlen, 1999, A review of the earnings management literature and its implications for standard setting, *Accounting Horizons* 13, 365–383.
- Hoffman, Elizabeth, Kevin A. McCabe, and Vernon Smith, 1996, On expectations and the monetary stakes in ultimatum games, *International Journal of Game Theory* 25, 289–301.
- Huck, Steffen, Dorothea Kübler, and Jürgen Weibull, 2012, Social norms and economic incentives in firms, *Journal of Economic Behavior and Organization* 83, 173–185.
- Innes, Robert, and Arnab Mitra, 2013, Is dishonesty contagious?, *Economic Inquiry* 51, 722–734.
- Kaptein, Muel, and Mark S. Schwartz, 2007, The effectiveness of business codes: A critical examination of existing studies and the development of an integrated research model, *Journal of Business Ethics* 77, 111–127.
- Karpoff, Jonathan M., D. Scott Lee, and Gerald S. Martin, 2008, The cost to firms of cooking the books, *Journal of Financial and Quantitative Analysis* 43, 581–611.
- King, Gary, 1998, *Unifying Political Methodology* (The University of Michigan Press: Ann Arbor).
- Krupka, Erin L., and Roberto A. Weber, 2013, Identifying social norms using coordination games: Why does dictator-game sharing vary?, *Journal of the European Economic Association* 11, 495–524.
- Lepper, Mark, David Greene, and Richard Nisbett, 1973, Undermining children’s interest with extrinsic rewards: A test of the overjustification hypothesis, *Journal of Personality and Social Psychology* 28, 129–137.
- López-Pérez, Raúl, and Eli Spiegelmann, 2013, Why do people tell the truth? Experimental evidence for pure lie aversion, *Experimental Economics* 16, 233–247.
- Lundquist, Tobias, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson, 2009, The aversion to lying, *Journal of Economic Behavior and Organization* 70, 81–92.

- Mazar, Nina, On Amir, and Dan Ariely, 2008, The dishonesty of honest people: A theory of self-concept maintenance, *Journal of Marketing Research* 45, 633–644.
- Milgram, Stanley, 1974, *Obedience to Authority* (Harper and Row).
- Miller, Nolan, Alexander F. Wagner, and Richard J. Zeckhauser, 2013, Solomonic separation: Risk decisions as productivity indicators, *Journal of Risk and Uncertainty* 46, 265–297.
- Musch, Jochen, Robbi Brockhaus, and Arndt Bröder, 2002, Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit, *Diagnostica* 48, 121–129.
- Nolan, Jessica M., P. Wesley Schultz, Robert B. Cialdini, Noah J. Goldstein, and Vldas Griskevicius, 2008, Normative social influence is underdetected, *Personality and Social Psychology Bulletin* 34, 913–923.
- Ostrom, Elinor, 2000, Collective action and the evolution of social norms, *Journal of Economic Perspectives* 14, 137–158.
- Paulhus, Delroy, 1984, Two-component models of socially desirable responding, *Journal of Personality and Social Psychology* 46, 598–609.
- Sade, Orly, Charles Schnitzlein, and Jaime F. Zender, 2006, Competition and cooperation in divisible good auctions: An experimental investigation, *Review of Financial Studies* 19, 195–235.
- Sánchez-Pagés, Santiago, and Marc Vorsatz, 2007, An experimental study of truth-telling in a sender-receiver game, *Games and Economic Behavior* 61, 86–112.
- Sherif, Muzafer, 1936, *The Psychology of Social Norms* (Harper and Row: New York).
- Shleifer, Andrei, 2004, Does competition destroy ethical behavior?, *American Economic Review* 94, 414–418.
- Sliwka, Dirk, 2007, Trust as a signal of a social norm and the hidden costs of incentive schemes, *American Economic Review* 97, 999–1017.
- Smith, Vernon, Gerry L. Suchanek, and Arlington W. Williams, 1988, Bubbles, crashes, and endogenous expectations in experimental spot asset markets, *Econometrica* 56, 1119–1151.

- Tanner, Carmen, Bettina Ryf, and Martin Hanselmann, 2009, Geschützte Werte Skala: Konstruktion und erste Validierung eines Messinstrumentes (Protected Values Measure: Construction and first validation of an instrument to assess protected values), *Diagnostica* 55, 174–183.
- Tetlock, Philip E., Ori V. Kristel, S. Beth Elson, Melanie C. Green, and Jennifer S. Lerner, 2000, The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals, *Journal of Personality and Social Psychology* 78, 853–870.
- van der Weele, Joël, and Ferdinand von Siemens, 2014, Bracelets of pride and guilt? An experimental test of self-signaling in charitable giving, *CESifo working paper no. 4674*.
- Van Lange, Paul A. M., Wilma Otten, Ellen M. N. De Bruin, and Jeffrey A. Joireman, 1997, Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence, *Journal of Personality and Social Psychology* 73, 733–746.
- Wang, Joseph Tao-Yi, Michael Spezio, and Colin F. Camerer, 2010, Pinocchio’s pupil: Using eyetracking and pupil dilation to understand truth-telling and deception in game, *American Economic Review* 100, 984–1007.
- Wooldridge, Jeffrey M., 2006, *Introductory Econometrics: A Modern Approach* (Thomson Higher Education: Mason).
- Zingales, Luigi, 2015, The ‘cultural revolution’ in finance, *Journal of Financial Economics* 117, 1–4.

Please note: The following supplementary appendices are not meant for publication in print. They are included for the benefit of referees.

A.1 Self-signaling model

This Supplementary Appendix provides a model for why one might expect a systematic positive relationship between ρ_i and the agent's type, i.e., $r > 0$.

As before, all agents experience morally driven costs of lying, though the amount depends on their type: $C_i = m\theta_i$, where θ_i is the ethical type and m is the common marginal value of truthfulness (which we normalized to unity in the main text). There are extrinsic consequences of truthfulness, denoted by $EXCO$. As in the main text, these might entail the non-monetary situational norm-driven costs of stating the truth, $SNCOST_s$, where s denotes the type of situational norm. They might also entail economic costs of stating the truth, $ECOST_e$, where the index e denotes the $ECOST$ situation. For the purposes of this theoretical framework we subsume the impact of both $SNCOST$ and $ECOST$ under the term extrinsic consequences, $EXCO_{se}$.

Additionally, as a key ingredient of the self-signaling model, the agent is unsure about (and has imperfect memory of) his type, but can interpret his actions as self-signals of his intrinsic preferences.

Suppose that there is a continuum of ethical types, distributed continuously with $F(\theta)$ between upper and lower bounds of $\underline{\theta}$ and $\bar{\theta}$, respectively.

Self-signaling is incorporated into the utility function by positing

$$V_{ise}(T) = \begin{cases} -EXCO_{se}(1 - r\theta_i) + \eta\zeta_1 & \text{if } T = 1 \\ -m\theta_i + \eta\zeta_0 & \text{if } T = 0. \end{cases} \quad (6)$$

Here, ζ_1 is the posterior estimate the agent has about his own type if he tells the truth, ζ_0 is the posterior estimate the agent has about his own type if he lies, and $\eta > 0$ is a parameter which indicates how much the agent cares about his (moral) self-image.

The difference between the utilities of truthtelling and of lying is given by

$$Y_{ise}^* = m\theta_i - EXCO_{se}(1 - r\theta_i) + \eta(\zeta_1 - \zeta_0). \quad (7)$$

An individual exhibits truthfulness when $Y_{ise}^* > 0$.

Consider a self-signaling separating equilibrium defined by $\hat{\theta}$ such that for agents with $\theta \geq \hat{\theta}$, $T = 1$ and all other agent types lie. In additional materials available on request, we show that if θ is uniformly distributed over the interval $[\underline{\theta}, \bar{\theta}]$, the cutoff is given by

$$\hat{\theta} = \frac{EXCO_{se} - \frac{\eta(\bar{\theta} - \underline{\theta})}{2}}{m + rEXCO_{se}}. \quad (8)$$

We also show that, if $\underline{\theta}=0$, a necessary condition for a unique separating equilibrium of the form postulated to exist is that

$$r > -\frac{m}{EXCO_{se}}. \quad (9)$$

Showing this result formally requires the use of fixed point theorems; the details are available on request. (We also derive similar results for the case when θ is truncated normally distributed.) Intuitively, two agent characteristics support an interpretation of truthfulness as an act of self-signaling. First, when $m > 0$, higher types have higher marginal utility of truthtelling, giving rise to a single-crossing condition. Note, though, that with $m > 0$ high types would be more likely to tell the truth whether or not they engage in self-signaling.

Second, and of primary interest for our paper, the resistance parameter is bounded from below. Sufficiently strong resistance of high types against extrinsic rewards for dishonesty allows self-signaling to work even if being a high type per se does not mean that one values truth as such more. If the direct marginal costs of lying, m , tend to zero, $r > 0$ is necessary to ensure $\hat{\theta} \geq 0$. The more costly it is to tell the truth, the less the range of possible values for r extends below zero.

A separating equilibrium of the form postulated fails if $EXCO_{se} < 0$; if truthfulness brings benefits, truthtelling is no effective self-signal. Even if $EXCO_{se} > 0$, no equilibrium with $\hat{\theta} \geq 0$ exists if the agent cares too much about self-image, that is, if η is too large. Finally, if r is too small or negative, $\hat{\theta} \leq \bar{\theta}$ may not exist.¹ Of course, in all these cases, people may tell the truth for non-self-signaling reasons.

Overall, $r > 0$ supports a self-signaling equilibrium (though it is not a necessary condition in general).

¹If η is large and r or $EXCO_{se}$ are negative, a (pathological) equilibrium where lower types tell the truth can exist.

A.2 Experimental instructions

The instructions of the experiment are attached at the end of this document. This section briefly discusses the construction of the proxies for individual characteristics.

A.2.1 Protected values for truthfulness survey

According to the correspondence (or compatibility) principle established by Ajzen and Fishbein (1980), values and behavior need to be assessed at a similar level of specificity in order to be able to uncover a link between the two. This principle underlies the protected values for truthfulness measure. The questionnaire contains two subscales designed to approach protected values for truthfulness from different angles.

(1) PRV (reactions to violations of honesty): Five items assessed the participants' reactions to violations of honesty by a hypothetical CEO who was reporting company information. This scale focuses on the affective dimension of individuals' commitment to honesty.

[*PRV*] Because CEOs' compensation levels depend on the earnings they report to their shareholders, CEOs have an incentive to modify reports to shareholders. What is your opinion on CEOs modifying company information in reports?

Please choose the appropriate category. This is:

Very immoral	1	2	3	4	5	6	7	Very moral
Not at all praiseworthy	1	2	3	4	5	6	7	Very praiseworthy
Not at all blameworthy	1	2	3	4	5	6	7	Very blameworthy
Not at all outrageous	1	2	3	4	5	6	7	Very outrageous
Not at all acceptable	1	2	3	4	5	6	7	Very acceptable

(2) PNT (no trade-off): Four additional items assessed the participants' own protected values by examining how much importance they attributed to trade-off reluctance, unwillingness to sacrifice a value against material or other benefits, or incommensurability, again referring to the specific context of a hypothetical CEO's decisions regarding the reporting of information.

[*PNT*] CEOs have an opportunity to modify information in the reports they provide to their shareholders. Some view such modification as a violation of truthfulness; others regard it as acceptable protection of personal interests. What do you think about the value of truthfulness in such a situation?

Truthfulness is something

... that one should not sacrifice, no matter what the (material or other) benefits									
strongly disagree	1	2	3	4	5	6	7		strongly agree
... for which I think it is right to make a cost-benefit analysis									
strongly disagree	1	2	3	4	5	6	7		strongly agree
... that cannot be measured in monetary terms									
strongly disagree	1	2	3	4	5	6	7		strongly agree
... about which I can be flexible if the situation demands it									
strongly disagree	1	2	3	4	5	6	7		strongly agree

After appropriate recoding of some items, indices of the degrees of protected values for truthfulness were constructed, based on the means across the first five items (for PRV), or the second four items (for PNT). The combined PVT is the mean of all nine items. The original protected values survey was conducted in German. In the paper, for ease of interpretation of the empirical results, we changed the scale to range from 0 to 6. The survey parts were not labeled for participants.

A.2.2 Pro-social concern

We use participants' answers regarding the extent to which they believed that announcing 35 cents had negative consequences for some stakeholders (-2 = hurting some stakeholders to +2 = not hurting some stakeholders) or was manipulative (-2 = manipulative to +2 = not manipulative). Answers to these questions are reordered so that participants who exhibited stronger pro-social concern score high on these scales. We then calculate the mean of the two items. The resulting variable PSC is standardized to have a mean of zero and a standard deviation of unity

A.2.3 Impression management and self-deception

We used the standard Deception Scales (PDS) of Paulhus (1984); Musch, Brockhaus, and Bröder (2002) for the German version. This is a self-reporting questionnaire designed to measure individuals' tendencies to give socially desirable responses (SDR). See the full instructions for details. It measures two distinct forms of SDR: self-deception and impression management. Accordingly, we coded two variables SELFDECEIT and IMPRESS. Participants who exhibited more socially acceptable responses scored higher on both scales.

A.2.4 Marginal utility

We asked the following question (drawn from Miller, Wagner, and Zeckhauser (2013)):

Please imagine that you find a CHF 50 bill on the street. It is impossible to identify the owner, and it is, therefore, completely acceptable and morally unobjectionable that you keep the CHF 50. Think about your average peer who earns about the same amount of money as you do, and is approximately equally wealthy. Would you say that, relative to this average peer, you benefit

a lot more

more

equally

less

a lot less

from this additional amount of money?

We assigned a value of 5 to “a lot more” answers, and a value of 1 to “a lot less” answers. This measure captures each participant’s self-reported marginal utility of income.

A.3 Additional evidence

A.3.1 Broader measures of ethical commitment

Conceptually, protected values have a plausible claim to be able to predict resistance against extrinsic incentives (Baron and Spranca, 1997; Tetlock, Kristel, Elson, Green, and Lerner, 2000). Nonetheless, it is useful to assess the ability of broader measures of ethical commitment in explaining such resistance. We have conducted another, hitherto unpublished experiment with the same truth-telling task (earnings management choices). Although the design was very similar, there were three important structural differences: First, the economic stakes were higher (ranging up to CHF 4 per choice). Second, we collected information on protected values as well as on some other personality traits at time 1, about 1-2 weeks before the experiment (time 2), thus introducing a time lag between personality trait collection and choices. Third, in the experiment itself, there were two groups: one exposed to a manipulation of interest (a depletion task) before the truth-telling task,

and one control group that was not exposed to this manipulation. In what follows, we only use data from the control group to ensure comparability with our experiment. However, the findings reported also hold when including the group for which we studied treatment effects. There were 51 participants in the control group. (A few participants were older than 40 years, and they are not included in the results that follow, though this does not materially affect the results.) 23 were economics students, 25 were psychology students, 3 studied in other fields. 29 were women. The median age was 21.

We caution that this sample is small. The ancillary findings from this additional experiment reported here should not be taken as general evidence on the relation between the various concepts of ethical commitment. Still, we believe they provide useful background information for the choice of measures in the main experiment.

At time 1, we collected information on the following traits, and participants also did some other tasks (including an effort task) in between answering these questions.

1. Protected Values (as in the current paper). For brevity, we present results for the combined protected values scale (combining PRV and PNT) below and comment briefly on the differences.
2. HEXACO Honesty-Humility Scale, the average of 10 items, in a German translation of Ashton and Lee (2009). In what follows, we refer to this as HEXACO.
3. Stated religiousness, the average of 2 items on a five-level scale (with a 6th item “no answer” allowed): “As how religious would you describe yourself?” and “How important is your religion / your religious belief to you?” In what follows, we refer to this as Religiousness.
4. Moral identity, the average of 5 items (5-level scale) in a German translation of Aquino and Reed (2002). Specifically, we computed the average of the five most relevant internalization items when evaluating the nine traits: caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind, see p. 1426 and Table 3 of Aquino and Reed (2002).
5. Social value orientation (SVO), 9 items in a German translation of the instrument in the Appendix of Van Lange, Otten, De Bruin, and Joireman (1997). Individuals were classified into Prosocials (SVO=1) if they made 6 or more prosocial choices and into Proselfs (SVO=0) otherwise. (Results do not depend on the exact classification method.)

The descriptive statistics are as follows. (1 participant did not answer the religiousness question, so in the regression analysis below the sample size is $N = 50$.) The measures are generally positively correlated.

Table A.1
Summary statistics of additional experiment

The data are from an experiment similar to the main experiment, but with ECOST levels CHF 0, CHF 0.5, CHF 1, CHF 2, CHF 4. For the explanation of the variables, see the text above.

Variable	Obs	Mean	Std. Dev.	Min	Max	Cronbach's alpha
PV	51	4.057	0.943	2.100	5.875	0.810
HEXACO	51	3.255	0.490	2.000	4.400	0.630
Religiousness	50	1.980	0.974	1.000	4.500	0.905
Moral identity	51	4.271	0.594	2.400	5.000	0.817
SVO	51	0.559	0.486	0.000	1.000	0.870

Table A.2
Correlation of individual characteristics in the additional experiment

The data are from an experiment similar to the main experiment, but with ECOST levels CHF 0, CHF 0.5, CHF 1, CHF 2, CHF 4. For the explanation of the variables, see the text above. In the table, * indicates significance on the 1% level, + indicates significance on the 5% level.

	PV	HEXACO	Religiousness	Moral identity	SVO
PV	1.000				
HEXACO	0.448*	1.000			
Religiousness	-0.050	0.079	1.000		
Moral identity	0.324+	0.333+	0.143	1.000	
SVO	0.395*	0.406*	0.122	0.442*	1.000

Table A.3
Choices in the additional experiment

This table shows the numbers of participants and the percentages of participants announcing 31 cents of earnings per share (TRUTHFUL CHOICE = 1) across various ECOST conditions. The data are from an experiment similar to the main experiment, but with ECOST levels CHF 0, CHF 0.5, CHF 1, CHF 2, CHF 4.

ECOST (CHF)	0	0.5	1	2	4	Total
Number of participants	44	30	27	22	11	134
Percent of participants	86%	59%	53%	43%	22%	53%
Total N	51	51	51	51	51	255

At time 2, one to two weeks after measuring these traits, participants returned to the lab. In this second round, they went through the earnings management task (as well as some other tasks). Table A.3 shows the descriptive statistics for the choices, presenting a similar pattern as in our current study.

Then, we conduct regression analysis as for the main experiment. We first find in untabulated results that each of the moral motivation measures is positively related to

truth-telling (with religiousness not significant on conventional levels). Because the individual measures are positively correlated, Table A.4 uses orthogonalized measures. Column (1) shows that PV and HEXACO remain as the by far most significant measures. Next, and of primary interest, we consider the interaction terms of the individual traits with ECOST to test for the ability of the individual traits to predict trade-off resistance. As a key result, PV most strongly explains resistance against economic costs. In results available on request, we confirm that PNT has somewhat stronger predictive power for the resistance against economic incentives than PRV, as in the main experiment.

Overall, these ancillary results support using the two PV scales to study resistance against social norms in our study.

Table A.4
Alternative measures of ethical motivations

This table presents coefficients of logit regressions. The dependent variable is TRUTHFUL CHOICE. The data are from an experiment similar to the main experiment, but with ECOST levels CHF 0, CHF 0.5, CHF 1, CHF 2, CHF 4. For the explanatory variables, see the text above. PV, HEXACO, Religiousness, Moral Identity, and SVO are orthogonalized. T-statistics, obtained from robust standard errors clustered at the individual level, appear in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)
ECOST	-0.813*** (-6.88)	-0.976*** (-6.03)
PV	0.962*** (3.51)	0.570* (1.90)
HEXACO	0.842*** (3.16)	0.515* (1.67)
Religiousness	0.513 (1.42)	0.577 (1.36)
Moral Identity	0.224 (1.13)	0.329 (1.43)
SVO	0.130 (0.48)	-0.110 (-0.37)
PV * ECOST		0.354*** (3.18)
HEXACO * ECOST		0.325* (1.68)
Religiousness * ECOST		-0.006 (-0.05)
Moral identity * ECOST		-0.027 (-0.20)
SVO * ECOST		0.182 (0.94)
Demographics (insignificant)	Yes	Yes
Constant	-0.681 (-0.21)	-1.389 (-0.43)
Observations	250	250

A.3.2 Pro-social concern

Table A.5 presents the results discussed in Section 4.4.1.

Table A.5
The role of pro-social concern

This table presents coefficients of logit regressions. The dependent variable is TRUTHFUL CHOICE. For the main explanatory variables, see the notes to Table 2. PRV and PSC are standardized to have a mean of zero and a standard deviation of unity, and are orthogonalized. The regressions consider situations where ECOST was strictly positive. T-statistics, obtained from robust standard errors clustered at the individual level, appear in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)
APPROVING (social norm, Phase SN)	-0.88*** (-3.45)	-0.86*** (-3.50)
DISAPPROVING (social norm, Phase SN)	0.76*** (4.43)	0.80*** (4.65)
CONTROL group (Phase 2)	-0.06 (-0.31)	-0.12 (-0.58)
PRV (protected values for truthfulness)		0.40** (2.36)
PRV * APPROVING		0.64** (2.17)
PRV * DISAPPROVING		-0.40** (-2.11)
PRV * CONTROL		0.10 (0.46)
PSC (pro-social concern)	0.18 (1.13)	0.09 (0.60)
PSC * APPROVING	0.62* (1.80)	0.09 (0.25)
PSC * DISAPPROVING	-0.48** (-2.41)	-0.29 (-1.59)
PSC * CONTROL	0.38* (1.86)	0.38** (2.02)
Demographic control variables	Yes	Yes
Constant	0.57 (0.76)	0.61 (0.78)
Observations	2,088	2,088
Pseudo R-squared	0.150	0.186
Pseudo Log Likelihood	-1122	-1074
Likelihood-ratio test statistic (?2, p-value)	396 (j0.01)	492 (j0.01)
Wald test statistic (?2, p-value)	130.8 (j0.01)	154.2 (j0.01)

A.3.3 Non-parametric tests and other robustness tests

As a complementary approach, to test the robustness of the findings and to investigate potential non-linearities, in Table A.6 we consider the impact of situational norms separately for different quantiles of PRV, that is, non-parametrically. The maximum PRV of participants in the first quartile is -0.71; the minimum PRV of participants in the fourth

quartile is +0.71. Columns (1) to (4) show that participants with higher PVT-RV values were generally more impervious to both norms as well as to economic incentives, in accordance with Hypothesis SUBST. In the case of the disapproving norm, the coefficients between the quantiles are not monotonic; they are, however, overall decreasing from the first to the fourth quartiles. This analysis highlights the finding that the participants with protected values significantly below the median were those who drove the strong response to situational norms.²

We have conducted many other variations of the analysis. Results available upon request show that models based on random effects logit regressions yield results similar to those from before. Additional results include the following: Comparing the behaviors of the participants in the cost and in the timing randomizations, we find no statistically significant differences. Using a logarithmic transformation of PRV slightly strengthens the results. Defining PSC to include also the extent to which participants regard announcing 35 cents as corresponding to a personal gain (thus presumably corresponding to a loss for somebody else) and as short-term behavior yields similar results.

A.3.4 Hypothesis regarding interactions between economic incentives and situational norms

The role of economic costs may vary as the norms vary. For example, Fischer and Huddart (2008) derive a model in which social norms augment the effects of incentives. In our context, this model implies that a dishonesty-approving norm would complement the effects of economic incentives against truthfulness; a dishonesty-disapproving norm would work against *ECOST*. Thus:

Hypothesis ECOST-COMPL: *Economic costs of stating the truth and social norms are complements: $b_{EASN} < 0$ and $b_{EDSN} > 0$.*

Alternatively, economic incentives may weaken the effects of situational norms, perhaps by the same mechanisms by which they crowd out effects of intrinsic preferences. This would imply that *ECOST* and social norms are substitutes, or $b_{EASN} > 0$ and $b_{EDSN} < 0$. Or, if preferences are separable in situational norms and economic incentives, we would

²Naturally, the regressions also imply that the marginal effect on the probability of truth-telling of approving and disapproving social norms was greatest for participants whose protected values approximated the median (not shown). Intuitively, for those who were strongly opportunistically inclined, social norms regarding earnings management did not have measurable behavioral effects, either because an approving norm encountered people who were already lying or because a disapproving norm failed to dislodge participants who were initially unmotivated to consider truthfulness as a viable option.

have $b_{EASN} = b_{EDSN} = 0$.

Consider regressions (5) to (8) in Panel B of Table A.6. Here, we add the interactions between *ECOST* and situational norms as an explanatory variable. We find that the coefficient on *ECOST*APPROVING* is strongly negative for participants with low PRV. In other words, for the typical person who is not motivated by moral preferences, Hypothesis ECOST-COMPL is a good description of reality. This corresponds to the prediction of the model in Fischer and Huddart (2008). By contrast, for those with strong PRV, the *ECOST*APPROVING* interaction term is in fact positive, though insignificant. The difference between Q1 and Q4 is highly significant, consistent with Hypothesis SUBST. The results regarding steadfastness with respect to the disapproving norm are somewhat weaker, though they trend in the same direction.

A.3.5 Demographics

Table A.7 documents that demographic characteristics do not interact with social norms (see the discussion in Section 5).

Table A.6

Differential resistance to social norms—non-parametric analysis

This Supplementary Appendix table presents coefficients of logit regressions. The dependent variable is TRUTHFUL CHOICE, which is equal to 1 when a participant chose to announce 31 cents and equal to 0 otherwise. The regressions are calculated separately for the participants in the quantiles of PRV described at the tops of the respective columns. The explanatory variables are defined in the text and in the notes to Table II. T-statistics, obtained from robust standard errors clustered at the individual level, appear in parentheses below coefficient estimates. All demographic controls are included; none have significant coefficients. The final column shows differences between coefficient estimates of interest for the top and bottom quantiles relevant to the respective regression sets are shown, for the variables listed in each row. The z-statistics for the significance of these differences are in parentheses.

As we have independent samples, these statistics are computed as $(\beta_i - \beta_j) / \left(\sqrt{se(\beta_i)^2 + se(\beta_j)^2} \right)$ for two quantiles i, j . ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A	(1) Bottom quarter PRV	(2) Q2 PRV	(3) Q3 PRV	(4) Top quarter PRV	Difference Q4 - Q1
APPROVING (social norm, Phase SN)	-1.23*** (-3.44)	-0.80** (-2.03)	-0.72** (-2.19)	0.08 (0.18)	1.31** (2.24)
DISAPPROVING (social norm, Phase SN)	1.30*** (3.42)	0.37 (1.05)	0.73** (2.19)	0.26 (0.83)	-1.04** (2.12)
CONTROL group (Phase SN)	-1.13*** (-2.99)	-0.06 (-0.13)	0.30 (0.94)	-0.35 (-1.26)	0.78* (1.65)
ECOST (cost of no earnings management)	-3.52*** (-6.41)	-2.52*** (-7.08)	-2.44*** (-7.33)	-1.63*** (-6.61)	1.89*** (3.13)
Demographic control variables	Yes	Yes	Yes	Yes	
Constant	1.36 (0.71)	0.21 (0.12)	1.42 (1.15)	1.53 (0.93)	
Observations	670	640	710	590	
Pseudo R^2	0.3	0.18	0.17	0.1	
Pseudo Log Likelihood	-361.5	-481.3	-481.3	-489.8	
Likelihood-ratio test statistic (χ^2 , p-value)	214 (j0.01)	150 (j0.01)	172 (j0.01)	80 (j0.01)	
Wald test statistic (χ^2 , p-value)	96 (j0.01)	60 (j0.01)	57 (j0.01)	51 (j0.01)	
Panel B	(5) Bottom quarter PRV	(6) Q2 PRV	(7) Q3 PRV	(8) Top quarter PRV	Difference Q4 - Q1
(ECOST*APPROVING)	-4.40** (-2.23)	-0.33 (-0.33)	-0.05 (-0.07)	1.07 (1.61)	5.48*** (2.64)
(ECOST*DISAPPROVING)	1.63** (2.22)	0.29 (0.53)	0.71 (0.94)	0.43 (1.22)	-1.20 (1.47)
(ECOST*CONTROL)	-3.74** (-2.05)	0.94* (1.70)	0.60 (1.37)	0.13 (0.33)	3.86** (2.07)
APPROVING (social norm, Phase SN)	-0.70 (-1.54)	-0.69 (-1.63)	-0.72* (-1.95)	-0.60 (-1.05)	0.04 (0.05)
DISAPPROVING (social norm, Phase SN)	0.45 (1.09)	0.21 (0.47)	0.27 (0.44)	-0.03 (-0.10)	-0.48 (0.94)
CONTROL group (Phase SN)	-0.62 (-1.38)	-0.54 (-1.28)	-0.06 (-0.16)	-0.44 (-1.44)	0.18 (0.32)
ECOST (cost of no earnings management)	-3.76*** (-5.46)	-2.72*** (-6.81)	-2.65*** (-7.40)	-1.86*** (-6.90)	1.90*** (2.56)
Demographic control variables	Yes	Yes	Yes	Yes	
Constant	1.48 (0.77)	0.33 (0.18)	1.54 (1.25)	1.69 (1.02)	
Observations	670	640	710	590	
Pseudo R^2	0.32	0.18	0.18	0.1	
Pseudo Log Likelihood	-353.2	-479.9	-479.9	-488.6	
Likelihood-ratio test statistic (χ^2 , p-value)	230 (j0.01)	153 (j0.01)	174 (j0.01)	82 (j0.01)	
Wald test statistic (χ^2 , p-value)	94 (j0.01)	61 (j0.01)	62 (j0.01)	60 (j0.01)	

Table A.7

Situational social norms, economic incentives, and agent-specific costs of lying: Demographics

This Supplementary Appendix table presents coefficients of logit regressions. The dependent variable is TRUTHFUL CHOICE, which is equal to 1 when a participant chose to announce 31 cents and equal to 0 otherwise. Data from the *ECOST*₀ situations are used. Newspaper is a binary indicator equal to 1 if a participant has recently read newspaper articles regarding CEOs (and 0 otherwise). Work is a binary indicator equal to 1 if a participant has a part-time job (and 0 otherwise). Investor is a binary indicator equal to 1 if a participant owns shares, bonds, or mutual funds (and 0 otherwise). All regressions control for all demographic characteristics. The column heading states which of the six characteristics is interacted with *ECOST* and the social norm indicators. T-statistics, obtained from robust standard errors clustered at the individual level, appear in parentheses below coefficient estimates. **, *, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	(1) Sex	(2) Age	(3) Economics	(4) Newspaper	(5) Work	(6) Investor
Demographic variable used in interactions:						
APPROVING (social norm, Phase SN)	-0.65** (-2.08)	-1.88 (-1.33)	-0.84*** (-2.89)	-0.69*** (-2.72)	-1.07** (-2.50)	-0.76*** (-2.98)
DISAPPROVING (social norm, Phase SN)	0.80*** (3.36)	1.89** (2.14)	0.68*** (2.91)	0.57*** (2.85)	0.83** (2.09)	0.52** (2.45)
CONTROL group (Phase SN)	-0.04 (-0.17)	0.31 (0.39)	0.20 (0.83)	0.07 (0.34)	0.67* (1.89)	-0.02 (-0.08)
ECOST (economic cost of no earnings management)	-1.27*** (-5.93)	-1.96** (-2.47)	-1.72*** (-7.28)	-1.71*** (-8.32)	-1.89*** (-4.98)	-1.78*** (-8.29)
Demographic variable * APPROVING	0.00 (0.00)	0.05 (0.90)	0.43 (0.97)	0.19 (0.36)	0.61 (1.18)	0.42 (0.78)
Demographic variable * DISAPPROVING	-0.30 (-0.90)	-0.05 (-1.43)	-0.05 (-0.15)	0.27 (0.73)	-0.22 (-0.50)	0.41 (1.16)
Demographic variable * CONTROL	0.23 (0.63)	-0.01 (-0.33)	-0.29 (-0.76)	-0.04 (-0.09)	-0.80* (-1.94)	0.26 (0.67)
Demographic variable * ECOST	-0.68** (-1.98)	0.02 (0.49)	0.30 (0.88)	0.49 (1.37)	0.39 (0.93)	0.64* (1.84)
Sex (1: Female, 0: Male)	0.83** (2.57)	0.36 (1.28)	0.34 (1.20)	0.35 (1.25)	0.37 (1.32)	0.35 (1.24)
Age (Years)	-0.00 (-0.17)	-0.01 (-0.22)	-0.01 (-0.20)	-0.00 (-0.18)	-0.00 (-0.17)	-0.01 (-0.19)
Economics (1: Economics, 0: Psychology)	-0.30 (-0.99)	-0.28 (-0.93)	-0.49 (-1.44)	-0.27 (-0.91)	-0.27 (-0.89)	-0.28 (-0.93)
Newspaper (1: Yes, 0: No)	-0.28 (-0.95)	-0.27 (-0.92)	-0.28 (-0.96)	-0.70** (-2.08)	-0.30 (-1.03)	-0.29 (-0.98)
Work (1: Yes, 0: No)	0.29 (1.10)	0.31 (1.17)	0.31 (1.18)	0.30 (1.16)	0.13 (0.40)	0.30 (1.16)
Investor (1: Yes, 0: No)	-0.03 (-0.12)	-0.02 (-0.07)	-0.03 (-0.11)	-0.04 (-0.14)	-0.02 (-0.06)	-0.66** (-2.19)
Constant	0.13 (0.17)	0.39 (0.43)	0.45 (0.59)	0.45 (0.59)	0.45 (0.59)	0.54 (0.71)
Observations	2,088	2,088	2,088	2,088	2,088	2,088
Pseudo R ²	0.09	0.09	0.09	0.09	0.09	0.09
Pseudo Log Likelihood	-1200	-1201	-1202	-1203	-1198	-1201
Likelihood-ratio test statistic (χ^2 , p-value)	240 (0.01)	238 (0.01)	236 (0.01)	234 (0.01)	244 (0.01)	238 (0.01)
Wald test statistic (χ^2 , p-value)	111 (0.01)	102 (0.01)	106 (0.01)	106 (0.01)	104 (0.01)	108 (0.01)

Experimental instructions

1

Welcome! This is a study on managerial decision-making.

The study will take around 20 minutes. You will be in the position of a "Chief Executive Officer" (CEO). You will have to take a number of decisions.

Your choices will be treated absolutely confidentially and anonymously. If you participate in the full study, you will receive CHF 8 plus a compensation depending on your choices. More information on what your compensation will depend on will follow later.

2

Participation code

To guarantee 100% anonymity, we kindly ask that you choose a participation code of the following form:

- The last three digits of your student ID +
- Three letters of your choice

Example: Student ID = 01-705-234 → 234
 3 letters of your choice: bsp

→ Your code: 234bsp

You will receive your compensation anonymously by stating your code. Remember your code as we have no way of identifying you otherwise.

Please enter your participation code into the field below and click "Continue."

Participation code:

3

General information

Please note:

- Please read the instructions for each task carefully.
- Please answer all questions.
- Please answer openly. Only your personal view counts, and there are no right or wrong answers (except for the clearly labeled questions checking understanding).

4

Personal information

- Gender:
 - ___ male
 - ___ female
- Age (e.g., 24):
 - ___
- Major
 - ___ Psychology
 - ___ Economics
 - ___ Other
- Specialization in psychology
 If you are majoring in psychology, please provide your field of specialization: (List)
- Specialization in economics
 If you are majoring in economics, please provide your field of specialization (List)
- Do you work part-time?
 - ___ yes
 - ___ no
- If yes, how much (in percent)?
 - ___
- Do you own individual stocks, mutual funds, or bonds?
 - ___ yes
 - ___ no

5

Information on compensation

In the following, you will have to put yourself in the position of a Chief Executive Officer (CEO). The financial payment that you receive after the experiment depends on the compensation of the CEO.

A higher compensation for the CEO means higher payment to you personally.

In addition, you will be asked some questions checking your understanding. For each incorrect answer, CHF 1 will be deducted from your final payment. By reading the instructions carefully, these questions are easy to answer.

In any case, you will receive CHF 8 at a minimum if you participate in the full study.

6

Introduction

Please read carefully the following description.

Imagine...

You are the Chief Executive Officer (CEO) of the company Castor AG. Castor AG is a publicly listed company. All shareholders are long-term investors. One of your tasks is to inform, each quarter, shareholders about the course of business and the earnings per share.

The end of the quarter, when you will have to report to shareholders, is around the corner.

7

Your compensation consists of a fixed and a variable salary component. The variable component is a bonus which depends on the announced earnings of the company. The higher the announced earnings, the higher will be your bonus.

The market currently anticipates the announcement of 35 cents as earnings per share. This is known to shareholders.

As the CEO, you know the true earnings of Castor AG, but the market does not. You know that earnings of 31 cents per share would more accurately reflect the actual value of the company.

It is your decision whether you make a legally permitted accounting modification to announce 35 cents instead of 31 cents for the earnings per share.

You know that the shareholders are closely following the earnings announcement of your company.

8

Before the actual decision, the following page asks some questions of understanding.

Please note that each incorrect answer will result in a deduction of CHF 1 from your total payment.

All questions need to be answered correctly before you can continue with the experiment.

9

Questions checking understanding

- How high does the market estimate the earnings per share for Castor AG?
 - _ 30 cents
 - _ 31 cents
 - _ 35 cents
- How high are the actual earnings per share of Castor AG according to internal accounting?
 - _ 30 cents
 - _ 31 cents
 - _ 35 cents
- With which announced earnings per share would you as the CEO receive a lower bonus?
 - _ with 31 cents of earnings per share
 - _ with 35 cents of earnings per share
- Can you as CEO announce earnings that deviate from the actual earnings?
 - _ yes
 - _ no
- Your compensation as CEO...
 - _ depends on the announced earnings per share
 - _ does not depend on the announced earnings per share

10

A quarter has passed and today you have to publish an earnings report. The market estimates earnings to be 35 cents per share. You know that earnings of 31 cents per share would more accurately reflect the actual value of the company.

In the following, there will be five possible situations for each of which you have to make a choice.

11

Which earnings will you announce?

(Please choose one alternative in each of the five situations.)

- 31 cents per share - In this case, your compensation will be **CHF 60,000** (CHF 0.30)*
- 35 cents per share - In this case, your compensation will be **CHF 300,000** (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be **CHF 120,000** (CHF 0.60)*
- 35 cents per share - In this case, your compensation will be **CHF 300,000** (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be **CHF 180,000** (CHF 0.90)*
- 35 cents per share - In this case, your compensation will be **CHF 300,000** (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be **CHF 240,000** (CHF 1.20)*
- 35 cents per share - In this case, your compensation will be **CHF 300,000** (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be **CHF 300,000** (CHF 1.50)*
- 35 cents per share - In this case, your compensation will be **CHF 300,000** (CHF 1.50)*

* This corresponds to the bonus payment that you receive after the experiment.

12

You expect that next year there will be a good opportunity for the acquisition of another company. However, you will need the shareholders' approval for that project. At the shareholder meeting you have an opportunity to convince the long-term investors of this idea. The investors are following closely your earnings announcements as well as those of the competition.

[Then, the participants were randomly assigned to one of the following three groups, which were not labeled for the participants]

[I. CONTROL GROUP]: --- [no additional information on this page]

[II. Earnings-management-APPROVING situational social norm group]: One evening, you are sitting with a friend of yours who is a financial analyst. He tells you that increasing reported earnings in order to meet market expectations meets with widespread societal approval.

[III. Earnings-management-DISAPPROVING situational social norm group]: One evening, you are sitting with a friend of yours who is a financial analyst. He tells you that increasing reported earnings in order to meet market expectations meets with widespread societal disapproval.

13

Again, a quarter has passed and today you again have to publish an earnings report. The market again estimates earnings to be 35 cents per share. You know that earnings of 31 cents per share would more accurately reflect the actual value of the company.

In the following, there will again be five possible situations for each of which you have to make a choice.

14

Which earnings will you announce?
(Please choose one alternative in each of the five situations.)

- 31 cents per share - In this case, your compensation will be CHF 60,000 (CHF 0.30)*
- 35 cents per share - In this case, your compensation will be CHF 300,000 (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be CHF 120,000 (CHF 0.60)*
- 35 cents per share - In this case, your compensation will be CHF 300,000 (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be CHF 180,000 (CHF 0.90)*
- 35 cents per share - In this case, your compensation will be CHF 300,000 (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be CHF 240,000 (CHF 1.20)*
- 35 cents per share - In this case, your compensation will be CHF 300,000 (CHF 1.50)*
- 31 cents per share - In this case, your compensation will be CHF 300,000 (CHF 1.50)*
- 35 cents per share - In this case, your compensation will be CHF 300,000 (CHF 1.50)*

* This corresponds to the bonus payment that you receive after the experiment.

15

A year has passed and you present the shareholders your proposed acquisition project. You will need the shareholders' approval for that project. What does this approval depend on in your view?

- That I have only presented high earnings (Yes, No)
- How high my compensation was (Yes, No)
- Whether I am seen as competent (Yes, No)
- Whether I have reported transparently in the past quarters (Yes, No)
- Do you think the project will be approved? (Yes, No)

16

From your own point of view, how do the two choices 31 cents vs. 35 cents earnings per share differ from each other?

To which extent do you judge the decision to announce 31 cents earnings per share as

dishonest	-2	-1	0	+1	+2	honest
manipulative	-2	-1	0	+1	+2	not manipulative
associated with personal costs	-2	-1	0	+1	+2	associated with personal gains
short-term oriented	-2	-1	0	+1	+2	long-term oriented
long-term hurting other stakeholders (e.g., shareholders, employees, suppliers)	-2	-1	0	+1	+2	long-term not hurting other stakeholders (e.g., shareholders, employees, suppliers)

To which extent do you judge the decision to announce 35 cents earnings per share as

dishonest	-2	-1	0	+1	+2	honest
manipulative	-2	-1	0	+1	+2	not manipulative
associated with personal costs	-2	-1	0	+1	+2	associated with personal gains
short-term oriented	-2	-1	0	+1	+2	long-term oriented
long-term hurting other stakeholders (e.g., shareholders, employees, suppliers)	-2	-1	0	+1	+2	long-term not hurting other stakeholders (e.g., shareholders, employees, suppliers)

17

Introduction

Please read carefully the following description.

Imagine...

You are the **Chief Executive Officer (CEO)** of a company. Your compensation consists of a fixed and a variable salary component. The variable component is a bonus which depends on the realized earnings of the company.

You have an opportunity to affect the firm's earnings by your effort. The more you work, the more time you need to invest, and the more money you will receive at the end of the experiment.

18

Calculations

In this task, you can increase earnings per share and, therefore, your compensation, by working. You will work on five sets of calculations. In each set, you can decide whether to do 1 or 5 simple calculations. Doing 5 calculations takes approximately five times as long as doing 1 calculation, and you will be paid more for this. The compensation you receive for 1 and for 5 calculations will vary over the five sets of calculations.

Moreover, you will receive CHF 0.2 for each correct calculation.

19

Before the actual decision, the following page asks some questions of understanding.

Please note that each incorrect answer will result in a deduction of CHF 1 from your total payment.

All questions need to be answered correctly before you can continue with the experiment.

20

Questions of understanding

- Can you as the CEO influence the earnings?
_ yes
_ no
- Your compensation as CEO...
_ depends on the realized earnings per share
_ does not depend on the realized earnings per share
- Your compensation as CEO is higher, if you solve...
_ 1 calculation
_ 5 calculations

21

Overview of the sets

Here you see an overview of the five sets of calculations and your compensation. You will make choices on these after two example calculations. In each set, you decide whether to do 1 or 5 calculations. You do not have to remember the payment levels now, as they will be shown in each choice situation later.

First set:

- 1 calculation - In this case, your compensation will be CHF 60,000 (CHF 0.30)*
- 5 calculations - In this case, your compensation will be CHF 300,000 (CHF 1.50)*

Second set:

- 1 calculation - In this case, your compensation will be CHF 120,000 (CHF 0.60)*
- 5 calculations - In this case, your compensation will be CHF 300,000 (CHF 1.50)*

Third set:

- 1 calculation - In this case, your compensation will be CHF 180,000 (CHF 0.90)*
- 5 calculations - In this case, your compensation will be CHF 300,000 (CHF 1.50)*

Fourth set:

- 1 calculation - In this case, your compensation will be CHF 240,000 (CHF 1.20)*
- 5 calculations - In this case, your compensation will be CHF 300,000 (CHF 1.50)*

Fifth set:

- 1 calculation - In this case, your compensation will be CHF 300,000 (CHF 1.50)*
- 5 calculations - In this case, your compensation will be CHF 300,000 (CHF 1.50)*

* This corresponds to the bonus payment that you receive after the experiment

22

Here are two examples of calculations. You do not have to solve them. The following calculations will have the same level of difficulty.

$$3 + 4 - 5 + 8 + 3 - 9 = \quad \text{Result: } \boxed{4}$$

$$15 + 3 - 4 + 18 + 6 - 12 = \quad \text{Result: } \boxed{26}$$

23

You can now decide whether to do 1 or 5 calculations. The payment for each option is given in parentheses below.

1. I decide to do...

1 calculation
(CHF 0.30)

5 calculations
(CHF 1.50)

[If participant clicks "1 calculation"]:

$$2 + 27 - 7 + 3 - 9 - 3 = \quad \text{Result: } \boxed{[13]}$$

[If participant clicks "5 calculations"]:

$$2 + 27 - 7 + 3 - 9 - 3 = \quad \text{Result: } \boxed{[13]}$$

$$34 - 5 + 16 - 7 - 12 - 2 = \quad \text{Result: } \boxed{[24]}$$

$$1 + 35 + 2 - 5 + 9 - 6 = \quad \text{Result: } \boxed{[36]}$$

$$11 - 7 + 26 - 2 + 4 - 5 = \quad \text{Result: } \boxed{[27]}$$

$$93 - 31 + 5 + 7 - 9 + 2 = \quad \text{Result: } \boxed{[67]}$$

24

You again can decide whether to do 1 or 5 calculations. The payment for each option is given in parentheses below.

2. I decide to do...

1 calculation (CHF 0.60)	5 calculations (CHF 1.50)
-----------------------------	------------------------------

[If participant clicks "1 calculation"] :
 $19 - 4 - 3 + 33 - 11 - 6 =$ Result:

[If participant clicks "5 calculations"] :
 $19 - 4 - 3 + 33 - 11 - 6 =$ Result:

$15 + 4 - 11 + 18 + 3 + 9 =$ Result:

$24 - 7 + 28 - 2 - 16 - 4 =$ Result:

$40 + 8 - 2 + 6 - 5 - 4 =$ Result:

$12 + 11 - 16 + 9 + 7 - 2 =$ Result:

25

You again can decide whether to do 1 or 5 calculations. The payment for each option is given in parentheses below.

3. I decide to do...

1 calculation (CHF 0.90)	5 calculations (CHF 1.50)
-----------------------------	------------------------------

[If participant clicks "1 calculation"] :
 $16 - 4 + 2 + 9 - 15 - 3 =$ Result:

[If participant clicks "5 calculations"] :
 $16 - 4 + 2 + 9 - 15 - 3 =$ Result:

$17 - 13 + 42 - 7 - 2 + 4 =$ Result:

$19 + 3 - 17 + 12 - 16 + 8 =$ Result:

$57 - 19 - 2 + 11 - 3 - 5 =$ Result:

$8 + 14 - 16 + 9 + 5 - 12 =$ Result:

26

You again can decide whether to do 1 or 5 calculations. The payment for each option is given in parentheses below.

4. I decide to do...

1 calculation (CHF 1.20)	5 calculations (CHF 1.50)
-----------------------------	------------------------------

[If participant clicks "1 calculation"] :
 $67 - 11 - 5 - 11 + 3 - 3 =$ Result:

[If participant clicks "5 calculations"] :
 $67 - 11 - 5 - 11 + 3 - 3 =$ Result:

$65 - 5 - 7 + 20 - 3 - 8 =$ Result:

$3 + 14 - 7 + 2 - 4 + 15 =$ Result:

$34 - 19 + 2 + 6 - 21 + 1 =$ Result:

$43 + 22 - 53 - 9 + 18 - 3 =$ Result:

27

You again can decide whether to do 1 or 5 calculations. The payment for each option is given in parentheses below.

5. I decide to do...

1 calculation (CHF 1.50)	5 calculations (CHF 1.50)
-----------------------------	------------------------------

[If participant clicks "1 calculation"] :
 $74 - 24 - 4 + 12 - 3 + 2 =$ Result:

[If participant clicks "5 calculations"] :
 $74 - 24 - 4 + 12 - 3 + 2 =$ Result:

$2 + 9 + 4 + 5 + 63 + 11 =$ Result:

$45 - 28 + 8 - 1 - 14 + 9 =$ Result:

$7 - 2 + 3 + 5 + 6 - 18 =$ Result:

$23 + 8 - 17 + 3 + 8 - 14 =$ Result:

28

In what follows, please answer a few questions on your personal attitudes.

29

Please imagine that you find a CHF 50 bill on the street. It is impossible to identify the owner, and it is, therefore, completely acceptable and morally unobjectionable that you keep the CHF 50. Think about your average peer who earns about the same amount of money as you do, and is approximately equally wealthy. Would you say that, relative to this average peer, you benefit

a lot more
 more
 equally
 less
 a lot less

from this additional amount of money?

30

