UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Corruption

A cross-country comparison of contagion and conformism

Schram, A.; Zheng, J.D.; Zhuravleva, T.

DOI
10.1016/j.jebo.2021.11.017

Publication date
2022

Document Version
Final published version

Published in
Journal of Economic Behavior and Organization

License
CC BY

Link to publication

Citation for published version (APA):
Schram, A., Zheng, J. D., & Zhuravleva, T. (2022). Corruption: A cross-country comparison of contagion and conformism. *Journal of Economic Behavior and Organization*, *193*, 497-518. https://doi.org/10.1016/j.jebo.2021.11.017

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Oct 2025

Journal of Economic Behavior and Organization 193 (2022) 497–518

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Corruption: A cross-country comparison of contagion and conformism[☆]

Arthur Schram[a,*], Jin Di Zheng[b,*], Tatyana Zhuravleva[c,*]

[a] CREED, University of Amsterdam, Department of Economics, European University Institute, Amsterdam School of Economics, Roetersstraat 11, 1018 WB Amsterdam, the Netherlands
[b] The Economics Experimental Lab (NEXL), Nanjing Audit University, No. 86 Yushan West Road, Pukou District, 211815 Nanjing, China.
[c] HSE University, International Laboratory for Experimental and Behavioural Economics, 11 Pokrovsky Boulevard, 119049 Moscow, Russia

ARTICLE INFO

Article history:
Received 26 February 2021
Revised 6 October 2021
Accepted 10 November 2021
Available online 15 December 2021

JEL classification:
C91
D73

Keywords:
Corruption
Bribery
Laboratory experiments
Contagion effect
Conformism effect

ABSTRACT

For successful anti-corruption policies, it is crucial to understand the basic social contract governing the interaction between people. Social norms are a key element of the social contract, but may vary across cultures. We investigate how descriptive social norms affect the development of corruption over time. In a laboratory experiment implemented in the Netherlands, Russia, Italy, and China we study a corruption game that is based on a real-effort task. To induce natural variation in descriptive norms, we vary the type of information about others' choices. Such information may lead to 'contagion' -where corruption increases in response to observing high corruption by others- or 'conformism' -where it decreases when low corruption by others is observed. Our results show evidence of contagion.

© 2021 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY license
(http://creativecommons.org/licenses/by/4.0/)

1. Introduction

Corruption is one of the most significant problems the world economy faces (Köbis et al., 2019). It is a widespread phenomenon, affecting every country across the globe. According to IMF estimates (IMF, 2016), the costs of bribery amount to 2% of the global GDP. While some researchers argue that corruption might increase efficiency (the "greasing the wheels" effect, Lui, 1985), most conclude that corruption has a negative effect on economic growth and development because it increases inequality (Gyimah-Brempong, 2002, Policardo and Carrera, 2018) and poverty (Gupta et al., 1998, Negin et al., 2010),

☆ We thank Katharina Brütt for her help in running the experimental sessions in Amsterdam, Yanwen Xing for her research assistance in sessions in Wuhan, and Alexey Belianin for his help in running the experimental sessions in Moscow. We are grateful for useful comments by editor Michele Gelfand, three anonymous reviewers, Klarita Gërxhani, Nils Köbis, and Shaul Salvi. We also thank the experimental laboratories BLESS (Bologna), CBER (Wuhan), CREED (Amsterdam), and HSE (Moscow) for allowing us to use their facilities. We gratefully acknowledge financial support from the Research Priority Area Behavioral Economics of the University of Amsterdam, and the Max Weber Program of the European University Institute. Finally, this article was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE), Russian Federation.

* Corresponding author.
E-mail addresses: schram@uva.nl (A. Schram), jindizheng@gmail.com (J.D. Zheng), tzhuravleva@hse.ru (T. Zhuravleva).

https://doi.org/10.1016/j.jebo.2021.11.017
0167-2681/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

reduces investments (Wei, 2000) and has a deleterious effect on the allocation of government expenditure (Mauro, 1997) and government debt (Cooray et al., 2017).

Attention for corruption and anticorruption policies has increased dramatically in recent decades, but in practice the results of these policies appear unimpressive. Some researchers now argue that “decision-makers should focus on policies that change the basic social contract, instead of relying solely on measures that are intended to change incentives for corrupt actors” (Rothstein, 2018: 35). To achieve this, it is very important to understand how citizens perceive corruption and whether it is seen as a violation of social norms, because social norms may be at the core of any social contract. This paper is an attempt to contribute to this understanding.[1]

We hope to contribute to this literature by studying the role of norms in corruptive behavior. A better understanding of this role might suggest alternative channels to govern corruption (Hauk and Saez-Marti, 2002). In particular, we investigate what we call a ‘contagion effect’ and a ‘conformism effect’. If an individual sees that corruption is common amongst her peers, then she might consider acting corruptly herself to be morally acceptable. In this way, corruption is ‘contagious’. In contrast, a corrupt-acting individual who sees that corruption is rare amongst her peers may conclude that it is deemed unacceptable and ‘conform’ to the more acceptable non-corrupt behavior.[2] We thus investigate whether corrupt behavior by an individual is sensitive to the disclosure of information about such choices by others. Such information constitutes a ‘descriptive norm’. Indeed, Köbis et al. (2015) provide evidence that information about others’ corruptive behavior affects one’s own decisions. We will discuss below the important differences between the Köbis et al. (2015) study and ours.

Contagion and conformism affect the development of behavior over time, which may or may not converge. If the extent of corruption converges, the level to which it does so may depend crucially on the environment in which it takes place. In particular, such dynamics may vary across countries. Aside from many other differences, distinct ‘prescriptive’ social norms may exist in various cultures, prescribing the extent to which corruptive choices are deemed (in)appropriate. Such prescriptive norms may interact with the descriptive norms, i.e, with the information individuals have about others’ corrupt choices.[3] The interaction between prescriptive and descriptive norms may strongly affect the dynamics of corrupt behavior (Bicchieri and Fukui, 1999). In fact, we would argue that this interaction is at the core of the ‘basic social contract’ and may therefore strongly impact corrupt behavior and the success of anti-corruption policies. In short, because prescriptive norms may differ across countries, the basic social contract may vary. As a consequence, contagion and corruption may depend on the country under observation. We therefore study these phenomena in four different countries.

We use laboratory experiments in this study. When comparing behavior across countries one needs to account for a plethora of differences in culture and institutions that might affect this behavior. The laboratory offers a unique environment to hold constant the many institutions that might differ and measure specific aspects of culture such as the relevant prescriptive norms. This allows us to correct for many of the confounds and focus primarily on what we are interested in: corrupt behavior by individuals and its relation to social norms.[4] All in all, we conduct a series of experiments in the Netherlands, Russia, Italy and China. As explained below, this allows us to collect data about the norms and behavior of people coming from societies with different perceptions of corruption.

We study corrupt behavior by applying the real-effort corruption game developed by Zheng et al. (2021). This builds on the paradigm introduced by Gneezy et al. (2019). We use a real-effort task as opposed to chosen effort because real effort allows for an objective measure of performance. Two performers of this task are grouped with a judge. We change the composition of groups between periods, in order to avoid long-run reciprocal relationships. The judge is informed about the performers scores on the task and must allocate a prize to one of the two. Before the judge decides, performers may transfer money to the judge. This is interpreted as a bribe. If the judge allocates the prize to a performer who performed worse but bribed more, we interpret this as corruption.[5] To allow for descriptive norms to develop, the type of information about others’ choices is varied. Either performers learn only the bribe choices of the other performers with whom they were grouped (a low-information treatment) or they learn the bribe choices of a much larger set of performers (high information). In the data pooled across all countries, we observe strong evidence of contagion, but little conformism. Performers strongly and significantly adjust their bribes upward if they observe that others are bribing more. Finally, we use separate sessions in each country and apply the elicitation method in Krupka and Weber (2013) to obtain information about existing prescriptive

[1] For successful policy, it is also important to understand whether and how people respond to the damage that their corruptive choices may cause to third parties and how such responses are affected by norms. Many find that negative externalities have no apparent effect (Abbink et al., 2002, Schulze and Franck, 2003, Cameron et al., 2009, Serra, 2011), while others find the opposite (Barr and Serra, 2009 Senci et al., 2019). With social preferences punitive policies may not be the most efficient path to follow (Armantier and Boly, 2011). As a case in point, consider China, which remains one of the most corrupt countries in the world, despite the existence of the death penalty for some types of corruption (Zhu, 2015). Similarly, it is important to understand the effects of norms on positive externalities. For example, some might act dishonestly if others benefit from this behavior (Weisel and Shalvi, 2015). Both types of externalities, and the role of norms therein, are beyond the scope of this paper.

[2] One way to think about the relationship between these terms is that contagion is conformism towards a negative act.

[3] The distinction between these two types of norm is important (Bicchieri, 2005; Brennan et al., 2013). Prescriptive (a.k.a. ‘injunctive’, or ‘moral’) norms prescribe what one ‘ought to do’, while descriptive norms, or ‘norms as social practice’, describe how things are usually done in a group, organization or society.

[4] We will discuss possible disadvantages to the use of experiments and related next steps in the concluding discussion.

[5] The judge receives both bribes in any case and in the subsequent period she plays with other bribers, thus there is no material gain for her to behave dishonestly.

norms with respect to corrupt behavior. The results show a stark difference in the prescriptive norms across the countries we selected.

The remainder of this paper is structured as follows. Section 2 provides a brief discussion of the state of the art in the related literature. Section 3 presents the experimental design and procedures, Section 4 describes and discusses the results, while Section 5 offers a concluding discussion.

2. State of the art

Social norms are unwritten rules that guide human activities within a society; they are “the standards of behavior that are based on widely shared beliefs how individual group members ought to behave in a given situation” (Fehr and Fischbacher, 2004). While prescriptive norms or ‘normative expectations’ (Bicchieri, 2005) indicate what one “ought to do” in a given social context, descriptive norms describe what a majority of group members are likely to do. Studies in social psychology have shed light on the “spreading of disorder” in rule violations (Keizer et al., 2008); this is an example of the effects that descriptive norms may have.[6] More recent research takes this one step further and studies the psychological mechanisms that are involved in the influence of descriptive and prescriptive social norms on corruptive behavior (Köbis et al., 2015; Shalvi et al., 2015; Guerra and Zhuravleva, 2021; Guerra and Zhuravleva, 2020).

In terms of the research questions asked, Köbis et al. (2015) is perhaps the study closest to ours. They present three experiments. In the first two, they show that perceived descriptive norms about corruptive behavior correlate with corrupt decisions. In the third, they show that there is a causal link from such norms to corrupt choices. An important difference with our study is that we collect data in four different countries, across which existing prescriptive norms may substantially differ. In addition, there are important methodological differences between the Köbis et al. (2015) study and ours. Their study creates a hypothetical situation (‘inviting a minister on a holiday’) and provides no choice-based incentives. In addition, they use deception. In contrast, we apply the standard methods of experimental economics by creating a real economic environment with incentivized decisions.

Recent studies from the laboratory show that descriptive norms may have direct effects on bribery behavior. Abbink et al. (2018) use a sequential bribery game to isolate the effects of descriptive social norms on bribe offers. In the first stage, a public official decides whether or not to accept a bribe from a firm; based on corrupt decisions of the officials, different official-firm pairs are formed. The effect of descriptive norms on bribe offers is clear; once a firm is paired with a corrupt official the bribes are more than doubled compared to when the firm is paired with an honest official. In addition, this effect persists independently of the sanctioning environment. Note that a different kind of descriptive norm is used by Abbink et al. (2018) than by us. Whereas they consider the effects of norms regarding bribees’ behavior on bribers’ choices, we consider norms relating to the bribers themselves. We will see, however, that our results do include a ‘mirror image’ of this finding; when our judges (bribees) are informed about the extent of bribing, they tend to make more corrupt choices.

Prescriptive norms are generally assumed to be much more stable than descriptive norms. Some have argued that beliefs and values are transmitted in a slow-moving process from generation to generation onto economic outcomes as part of the impact of culture (Guiso et al., 2006), so that corrupt behavior can be predicted by the prevailing social norms in a country. Fisman and Miguel (2007) use empirical data of diplomats from 149 countries in New York and find that diplomats from high-corruption countries accumulate significantly more unpaid parking violations than those from other countries. Gächter and Schulz (2016) use controlled experiments from 23 countries to demonstrate a robust link between social norms and individual’s intrinsic honesty. The authors build an index of “prevalence of rule violations (PRV)” (including tax evasion, corruption and political fraud) for each country and relate it to the results in a die-rolling experiment. They find that subjects from low-PRV value countries are less likely to lie in the experiment than those from high-PRV countries.

In spite of this growing literature, evidence remains mixed as to the effects of social norms on corrupt behavior. To begin with, such effects appear to be sensitive to specific microcultures. By conducting controlled experiments among employees from a large, international bank, Cohn et al. (2014) show that making the professional identity as bank employees salient increases dishonesty (though Rahwan et al., 2019 cast doubt on the generalizability of these results). These effects are not observed in other industries or among students that are primed with bank-related items. This suggests that a specific business culture in banks may have a strong effect on corruptive choices, which may interact with prescriptive norms that are active in the (macro-)culture where a bank is located.

Finally, cross-country comparisons produce inconclusive results about the effects of social norms on corruption. Various studies apply bribery games in different countries (Alatas et al., 2009, Cameron et al., 2009, Banuri and Eckel, 2012a, Banuri and Eckel, 2015, and Frank et al., 2015). Others conduct the game with subjects from different countries (Barr and Serra, 2010). For more examples, see Banuri and Eckel (2012b). By and large, results show that distinct social norms alone cannot perfectly predict corrupt behavior. Other elements such as institutional changes, gender, or group composition may play a role. In our analysis, we will correct for such factors.

[6] For descriptive norms to have such effects, people must prefer to act in accordance with the norm conditional on their empirical expectations (Bicchieri, 2005). In this paper, we assume this to be the case.

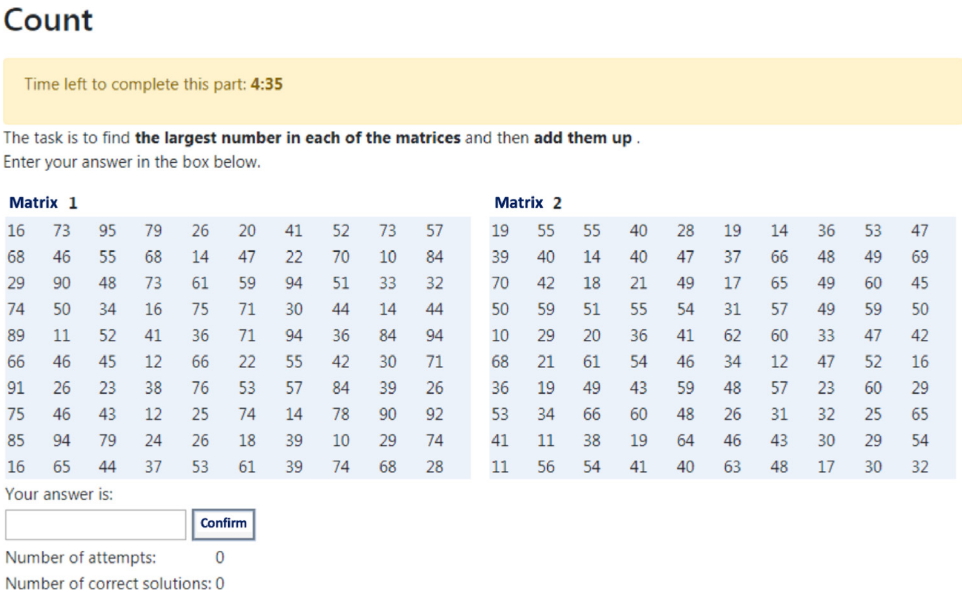


Fig. 1. Real effort task.

3. Experimental design

There are two types of sessions in each country. In one, participants play the 'corruption game' described below. In this game, we varied the possibility for descriptive norms to develop. In the second type, we measured existing prescriptive norms. We first describe the corruption game.

3.1. Corruption game

General setup

This corruption game was developed in Zheng et al. (2021), building on Gneezy et al. (2019) and used in other corruption experiments (Balafoutas et al., 2021).[7] There are two roles in the experiment: performers, and judges. Upon arrival in the laboratory, subjects are randomly assigned one of these, which remains fixed throughout the experiment. They subsequently play ten periods of the game described below. In each period they are (re)matched in groups of three, consisting of one judge and two performers.[8] After ten periods of the task, subjects are asked to fill out a questionnaire with personal characteristics and attitudes. At the end of the experiment, one period is randomly chosen for payment.

At the beginning of each period each subject receives an initial endowment of ten experimental tokens. Each period subsequently consists of three stages. At the first stage, the two performers in a group each and independently carry out the real-effort task explained below. At stage two, performers have an opportunity to transfer money to the judge (out of their ten-token endowment). We interpret any money transferred as a bribe. All bribe transfers are irrevocable; the judge keeps bribes from both performers irrespective of her subsequent decision. At stage three, the judge chooses a 'winner' from the two performers. This winner receives a ten-token prize, which is irrelevant for the judge's payoff.

Real-effort task

We use the real-effort task developed by Weber and Schram (2016). On their computer monitor, performers see two 10 by 10 matrices filled with two-digit numbers. Their task is to find the largest number in each of the two matrices and add them up (see Fig. 1 for an illustration). After entering a number, a new set of randomly chosen matrices appears on the next screen, irrespective of whether the number entered was correct or not. This is an individual task and each performer has five minutes to solve as many of these matrix summations as they can. Each correct answer adds one point to the performer's total score. On the monitor, performers can see the remaining time and also the number of attempts and correct trials. At this stage the judge waits.

[7] See Appendix A for the experimental instructions.

[8] Rematching takes place within a matching group of 12 subjects, with two matching groups per session. An exception is Russia, where the size of the laboratory only allows for one matching group per session.

Bribes and prize allocation

At the second stage, each performer is informed of her number of correct answers. Then she decides how many tokens (from 0 to 10) to transfer to the judge. Performers are informed that the transfer is nonrefundable and it does not guarantee them winning a prize. We denote performers i and j's performance by P_i and P_j, and their transfers to the judge by B_i and B_j, respectively. The judge remains inactive at this stage. Performers only know their own performance and transfer amounts.

At the third stage, the judge is informed about both performers' scores and transfers. That is, she knows P_i, P_j, B_i, and B_j. The judge then nominates and awards ten tokens to the winner. Zheng et al. (2021) show for an experiment run in Amsterdam, that in the absence of the possibility to bribe (and in the absence of social ties between the judge and performers) a very large majority of judges allocate the prize to the performer with the higher score. This suggests that the prize is seen as a reward for performance, even though it is never explicitly stated that this is so.

Treatments

We adopt a 4 × 2 full-factorial between-subject design. The treatment dimensions vary in which country subjects are living in and how much information (about other performers' bribes) they are provided with during the experiment. Our selection of countries is based on the wish to have a broad spectrum of actual corruption levels, and therefore presumably have a large variety in the prescriptive norms regarding corruption. This allows is to study contagion and conformism under divers settings with respect to experience with and attitudes towards corruption. In the end, we decided to collect data in The Netherlands (*NL*), Russia (*RS*), Italy (*IT*) and China (*CH*). We will discuss below how these countries vary in corruption and prescriptive norms.

The information level can be either low (*LI*) or high (*HI*). In *LI*, only information about decisions in one's own three-person group is provided. After the judge's decision is made, participants are told P_i, P_j, B_i, and B_j and who is the winner. A performer is therefore only informed about the other performer's score and transfer to the judge, as well as the judge's decision in that period. Because of the strangers' environment that we use, information about others' behavior gradually accumulates across periods. Participants in the high-information treatment (*HI*) are also informed about P_i, P_j, B_i, and B_j, and the judge's decision. In addition, they are informed of the average bribe in their matching group (recall that there are 12 participants in each matching group)[9]. This latter information is provided by way of a pop-up message that appears on each performer's and on the judge's monitor: *"This period, the average amount transferred per person to the judge is ... points"*. Each subject has to confirm this message before she can proceed to the next period.

Players' payoffs

Let i be the performer to whom the judge awards the prize and j the other performer. The monetary payoffs in any given period are then:

$$\begin{aligned} \pi_i &= 20 - B_i \\ \pi_j &= 10 - B_j \\ \pi_{judge} &= 10 + B_i + B_j \end{aligned} \tag{1}$$

All players are paid for one randomly selected period. In addition, each participant receives a participation fee. This was 7 euros in *NL*, 150 RuR (∼2.1 euros) in *RS*, 5 euros in *IT*, and 15 yuan in *CH* (∼2 euros). These participation fees were chosen in accordance with the standards of the laboratories concerned.

3.2. Social norm sessions

To measure prescriptive norms in each country we apply the incentivized elicitation method introduced by Krupka and Weber (2013).[10] We do so in a separate session using different participants from the same subject pool in each country.[11] In these sessions, we provide participants with a set of scenarios for the corruption game (varying performance, bribes, and the judge's decision). For each scenario, we ask respondents to rate the behavior as being "very socially inappropriate", "somewhat socially inappropriate", "somewhat socially appropriate", or "very socially appropriate". A participant receives a payment if she chooses the modal response for that scenario; otherwise, she receives only the participation fee. In this

[9] Due to an organizational error, one of the matching groups in Russia consisted of 15 participants.

[10] When studying corruption, there is always a danger of biased responses due to the 'sensitivity' of the topic or Hawthorne effects. This could conceivably play a role both in the corruption game and in the social norm sessions. Various design details contribute to our attempts to circumvent this problem. First, decisions were completely anonymous, which was stressed in the instructions. Moreover, all experiments were run in established laboratories of experimental economics where the subject pools know that anonymity is ensured when they register for an experiment. Second, we avoid terms like 'corruption' and 'bribes' throughout the instructions and experiment. Third, though one can never completely exclude the possibility of Hawthorne effects, we deem it unlikely that they would play a role here. In the social norm sessions, the task is to match the response of other participants, and it seems unlikely that the experimental setting would lead this coordination game to any specific outcome. For the corruption game, any Hawthorne effect would likely be the same in both treatments, so treatment effects would not be affected. Moreover, it is hard to see how contagion and conformism could fall prey to a Hawthorne effect.

[11] Instructions for these sessions are presented in Appendix B.

way, participants are incentivized to match the modal response provided by the majority of others rating the same scenario. Varying the scenarios allows us to identify in each country prescriptive social norms that govern bribe-giving and bribe-taking behavior. Note that this Krupka and Weber (2013) method uses a coordination game to reveal the collective perception regarding the appropriateness of different types of behavior.

We distinguish between seventeen hypothetical scenarios (see Appendix B), which are presented to participants in three sets of four scenarios and one set of five. These sets are presented in random order. To elicit norms concerning bribing behavior, we first posit an average score for the task. We then consider an environment where a hypothetical performer has a score below, equal to, or above this average. For each of these scores, we consider bribes of 0, 3, or 7 tokens. This results in nine different situations for which our subjects need to rate the appropriateness of the bribe. Next, we ask subjects to evaluate a hypothetical judge's decision for distinct environments. We posit one performer with an average score and the other with an above-average score. We consider two sets of bribe levels (7 - 3, and 3 - 0). Then we consider the judge allocating the prize to either player. This allows us to distinguish between the judge nominating (1) a player with a higher score and a higher bribe than the other, (2) a higher score and a lower bribe; (3) a lower score and a higher bribe; or a (4) a lower score and a lower bribe. We do so for each of the two sets of bribes, giving eight scenarios.

After subjects have indicated social appropriateness ratings for all 17 scenarios, one is randomly selected for payment. The experimenter computes the modal response for this scenario and pays every respondent with this answer 10 euros in Italy and the Netherlands, 500 RuR in Russia (~6.5 euros) and 30 yuan in China (~4 euros).

3.3. Procedures

The experiment was run at the BLESS laboratory at the University of Bologna in Italy in June 2018, the CREED laboratory of the University of Amsterdam in the Netherlands in September 2018, the CBER laboratory at Wuhan University in China in September 2018, and at the Higher School of Economics in Moscow, Russia, in April 2018 and September 2019. In all countries, we apply the same experimental instruction and protocol.[12] The corruption game sessions were computerized in oTree (Chen et al., 2016), the social norm sessions were run run in pencil and paper. Tokens are exchanged for the local currency of the country concerned.The exchange rate was 1 token = 1 euro in Amsterdam and Bologna, 1 token = 50 RuR (~0.65 euro) in Moscow, and 1 token = 3 yuan (~0.4 euro) in Wuhan.

In *NL, IT*, and *CH*, there were three sessions of *LI*, three sessions of *HI* and one social norm session. Each session of *LI* and *HI* consisted of 24 participants and was divided into two matching groups. The number of participants in the social norm sessions was 39, 40, and 40 in *NL, IT*, and *CH*, respectively. As mentioned earlier, the laboratory size in Moscow does not allow for 24 participants, so we had six sessions each for *LI* and *HI*, one of which had 15 participants instead of 12 (see footnote [9]). Here, the social norm session consisted of 32 participants. In total we thus had 579 participants for the corruption game and 151 participants for the social norm sessions.

Our post-experimental survey shows that subjects are mainly college students in the 18–24 age range.[13] The fraction of females is close to 50% in all countries except *CH*, where it is approximately 65%. The fraction of participants with a major in economics or business varies more across subject pools. It is highest (60%) in *CH* and lowest (30%) in *IT*. We also have information about participants' self-reported risk attitudes, income, life satisfaction, and trust. There is limited variation in these variables across subject pools, but we will correct for such differences in our regression analyses below. Finally, the cultural background of our participants varies across countries. All participants in *RS* are Russian, those in *IT* are Italian, and all *CH* are Chinese. In *NL*, participants' cultural background is mixed; 33% are Dutch, 14% come from the other parts of Western Europe, 17% from Italy, and 10% from China.

Descriptive norm sessions lasted approximately 1.5 hours and the average earnings including show-up fee were 20.3 euro in Amsterdam, 950 RuR (~13.3 euro) in Moscow, 18.3 euro in Bologna, and 55 yuan (~7 euro) in Wuhan. Prescriptive norm sessions lasted approximately 1 h. In *NL* 18 people out of 39 correctly guessed the modal response (yielding average earnings of 11.6 euro); this was 16 out of 32 in *RS* (400 RuR), 12 out of 40 in *IT* (8 euro), and 20 out of 40 in *CH* (42.5 yuan).

3.4. Hypotheses

Because the number of countries where we measure prescriptive norms is limited to four, we refrain from deriving formal hypotheses with respect to these norms. For the effects of descriptive norms, we focus on bribery, because the information we vary across treatments is about others' bribes.[14] As mentioned above, we distinguish between *Contagion* and *Conformism*. Contagion refers to a situation where observed high levels of bribes make bribing 'acceptable'. As a consequence, a decision maker observing that others bribe more than she does will subsequently increase her bribes. Conformism refers to the

[12] Languages in the instructions are English in Amsterdam, Russian in Moscow, Italian in Bologna, and Chinese in Wuhan. Some of the sessions in Moscow had to be postponed until September 2019 due to logistical reasons. These sessions were run precisely as originally planned and with the same subject pool.

[13] See Appendix C for the questions asked in the survey and Appendix D for the descriptive statistics.

[14] We could have, of course, provided information about judges' decisions at the matching-group level, but this would require much more detail than a simple average.

Table 1
Prescriptive Norms.

	NL	*RS*	*IT*	*CH*
Corruption Index	82	28	52	39
Bribes	−0.30	−0.09	0.03	−0.34
Corruption	−0.18	−0.28	0.14	−0.46

Notes: 'Corruption Index' shows the 2018 result measured by Transparency International. 'Bribes' gives the average appropriateness index for performers' bribing. 'Corruption' gives the average appropriateness index for judges' corruption. A positive (negative) number indicates (in)appropriateness.

reverse; someone who observes that others are bribing less than she is will subsequently diminish her bribes. Of course, contagion and conformism may well exist simultaneously. In fact, we hypothesize that both exist. Intuitively, we do predict that information based on the larger group in *HI* has a stronger effect than the pairwise information in *LI*.

Hypothesis 1: Contagion
The disclosure of information about average bribes leads performers to bribe more if this average is above their own previous bribe level. This effect is stronger in the treatment with information at the level of the matching group.

Hypothesis 2: Conformism
The disclosure of information about average bribes leads performers to bribe less if this average is below their own previous bribe level. This effect is stronger in the treatment with information at the level of the matching group.

Finally, we consider whether and how contagion and conformism might vary across the four countries. The corruption we observe in a country is likely to be affected by the prevailing prescriptive norms and by participants' experiences with their countries' institutions (Gërxhani and Wintrobe, 2021). Such influences will be strongest in early periods and diminish as information about others' choices accumulates and descriptive norms become more prominent. This will reduce differences in bribes and in judges' corruption. This yields our third hypothesis.

Hypothesis 3: Increasing influence of descriptive norms
Cross-country differences in bribes and judges' corruption decrease across periods.

4. Results

We ask three groups of questions. First, we consider the prescriptive norms that we measure in the four countries and relate them to corruption indices that are based on field data. Second, we analyze how bribes are affected by information about others' choices. That is, we investigate contagion and conformism and test hypotheses 1 and 2. Finally, a further analysis of the results per country and a closer look at judges' behavior allows us to test Hypothesis 3. Unless indicated otherwise, test results are based on permutation t-tests (henceforth, PtT), with 10,000 permutations.[15]

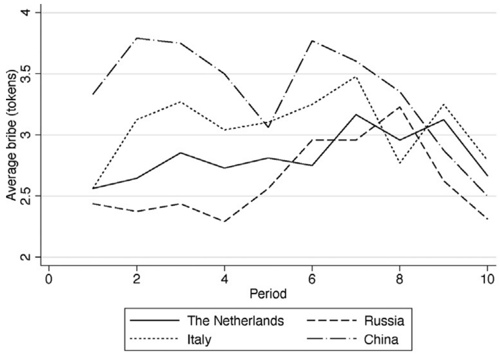
4.1. Corruption indices and prescriptive norms

The Transparency International corruption index 2018 measures perceptions by experts and businesspeople about the corruption in 180 countries around the world.[16] Corruption is indexed between 0 (highly corrupt) and 100 (not corrupt). The average value across countries in 2018 was 43. The first row of Table 1 shows the index for the four countries that we study. While Russia is perceived to experience much more corruption than average, the Netherlands is seen as particularly non-corrupt. The other two countries are closer to the average, with China slightly below and Italy above the mean of 43. Taken together, the four countries in which we collect data represent a broad spectrum of corruption levels, as was the intention.

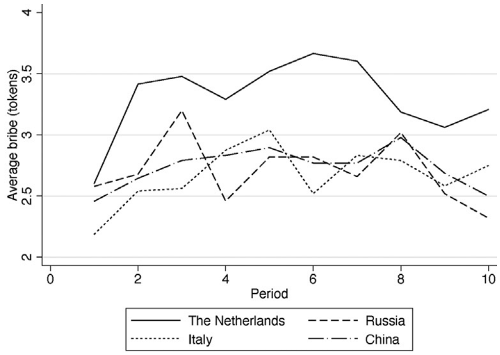
Using participants' decisions in the sessions where we measure prescriptive norms, we construct indices that reflect the extent to which these norms deem performers' bribes and judges' corruption appropriate.[17] The method we use is explained in Appendix E, where we also provide a detailed analysis of these norms. The indices take on a value between −1 (a strong prescriptive norm against bribes or corruption) and +1 (a strong prescriptive norm pro bribes or corruption). A positive (negative) index reflect a view that the behavior concerned is (in)appropriate. The mean values per country are shown in Table 1.

This measure of the appropriateness of bribes shows that bribing is deemed most inappropriate in *CH*, followed by *NL*, *RS*, and *IT*. Testing pairwise differences shows that prescriptive norms in *IT* are significantly more lenient towards bribes than in

[15] See Moir (1998) or Schram et al. (2019) for discussions of the PtT and its advantages.
[16] See https://www.transparency.org/cpi2018.
[17] We consider a judge to be corrupt if she awards the prize to B when (i) A has the better performance, and (ii) B bribes more than A.



(a) LI



(b) HI

Fig. 2. Bribes over 10 periods.

any of the other three countries (PtT, all $p < 0.01$). Pairwise differences between the other three countries are all statistically insignificant (PtT, all $p > 0.32$). A corrupted judge decision is also deemed most inappropriate in *CH*, followed by *RS*, *NL*, and *IT*, where it is considered appropriate. *CH* and *RS* thus have strong anti-corruption norms. Both have a significantly lower index than *IT* (PtT, both $p < 0.04$). All other pairwise comparisons are statistically insignificant (PtT, all $p > 0.11$).

What stands out in Table 1 is that the Chinese have the strongest prescriptive norms against bribes and corruption, while the Italians are most tolerant. The Dutch and the Russians are in between these two extremes, with the Dutch having stronger norms against bribes and the Russians against corruption. The two norms in China and the Russian anti-corruption norms are perhaps remarkable given the high corruption levels in these countries. This might be explained by the observation that people in countries with high corruption are confronted most often with the consequences thereof (Gatti et al., 2003; Smith, 2008). Alternatively, prescriptive norms may not be linearly related to the level of corruption. For example, Harrington et al. (2015) show that the extent of individual freedom in a society exhibits a non-linear relationship with societal well-being. In a similar vein, the relationship between prescriptive norms and corruption may be curvilinear.

What matters for our purposes, however, is that the four countries we selected show both a broad spectrum of corruption and of prescriptive norms, while the two are not aligned. This creates a divers playing field for our study of bribe contagion and conformism.

4.2. Bribes

Overview

Next, we consider bribes, and how they are affected by descriptive norms. Figure 2 shows the average bribes per period separately per country and information treatment.

We observe that the mean bribe lies between 2 and 4 (out of 10) tokens in both treatments and all countries. Moreover, no general trend across periods is discernible. In the left panel of Fig. 2 (the *LI* treatment) *CH* stands out with higher bribes, though bribes do decrease in the final four periods. The other countries appear to be comparable. In *HI*, it is *NL* that stands out with higher average bribes than in other countries. Permutation t-tests show that none of the pairwise differences in either panel, however, are significant (PtT, all $p > 0.14$)[18] This in itself is not surprising, because the averages across all periods are influence by prescriptive and descriptive norms.

Decisions in the first period cannot be influenced by descriptive norms since no information about others' choices has yet been disseminated. For this reason, we can pool in each country the observations for *LI* and *HI*. Any differences in period 1 bribes can then be attributed to cross-country differences, including distinct prescriptive norms. In addition, there has not been any interaction when performers decide on bribes in period 1, so their decisions may be considered as statistically independent. Average bribes in period 1 are 2.6 in *NL*; 2.5 in *RS*, 2.4 in *IT*, and 2.9 in *CH*. All pairwise differences are statistically insignificant except that period 1 bribes in *CH* are significantly higher than in *RS* (PtT, $p = 0.031$) and marginally so than in *IT* (PtT, $p = 0.065$); all other pairwise differences are statistically insignificant (PtT, all $p > 0.32$). We conclude that prescriptive norms and other differences across countries play only a minor role in period 1 bribes.[19]

Regression model

Our experiment is designed to identify how information about prevailing bribes affects performers' bribes in subsequent periods through the spread of descriptive norms. Specifically, we will analyze whether a player increases her bribe when she

[18] To account for interactions across periods, we use the matching group as the unit of observation for these tests.

[19] Recall from Table 1 that prescriptive norms against bribes are much stronger in *CH* than in *IT* and *RS*. Here we observe that first-period bribes are higher in *CH*.

Table 2
Measuring Contagion and Conformism.

Effect	Treatment	Coefficient	Sign
Group Contagion	*HI*	β_1	Positive
Group Conformism	*HI*	β_2	Negative
Competitor Contagion	*LI*	β_3	Positive
Competitor Conformism	*LI*	β_4	Negative
Competitor Contagion	*HI*	$\beta_3 + \beta_5$	
Competitor Conformism	*HI*	$\beta_4 + \beta_6$	

Notes: The final column shows the predicted sign of the coefficient(s) for the effects denoted in the first column. Recall that group contagion and conformism cannot occur in *LI* because no group-level information is provided there.

sees that her transfer is below the average bribe that she is informed of (the 'contagion' effect of information) and whether she reduces her bribe when she is informed that she has bribed more than others (the 'conformism' effect of information).

We adopt an OLS model for this analysis. The dependent variable is performer *i*'s change in bribes between two periods, $\Delta B_t^i = B_t^i - B_{t-1}^i$ (thus defined to be positive for an increase). The independent variables are the following.

(i) The difference between *i*'s own bribe and the average bribe in the matching-group in the previous period. This is denoted by $D_{A,t-1}^i$. We distinguish between whether *i*'s bribe is less than or more than the matching group, denoted by $D_{<A,t-1}^i$ and $D_{>A,t-1}^i$, respectively. Denoting the average previous bribe in the matching group by A_{t-1},

$$D_{<A,t-1}^i = \begin{cases} A_{t-1} - B_{t-1}^i, & \text{if } A_{t-1} \geq B_{t-1}^i \\ 0, & \text{otherwise} \end{cases}$$

$$D_{>A,t-1}^i = \begin{cases} B_{t-1}^i - A_{t-1}, & \text{if } A_{t-1} < B_{t-1}^i \\ 0, & \text{otherwise} \end{cases}.$$

Because -by design- participants in *LI* have no information about other bribes in the matching group, $D_{<A,t-1}^i$ and $D_{>A,t-1}^i$ are both set to 0 in *LI*. $D_{<A,t-1}^i$ then allows us to investigate how a performer's change in bribe level responds to previously having bribed less than the average bribe in her matching group. We call this 'group contagion'. $D_{>A,t-1}^i$ does the same in case the performer previously bribed more than group average, which we call 'group conformism'.

(ii) The difference between *i*'s own and her competitor's bribe in the previous period. This difference is denoted by $D_{j,t-1}^i$. We distinguish between $D_{<j,t-1}^i$ for cases when *i*'s bribe is less than her competitor's and $D_{>j,t-1}^i$ if she bribes more. Formally,

$$D_{<j,t-1}^i = \begin{cases} B_{t-1}^j - B_{t-1}^i, & \text{if } B_{t-1}^j \geq B_{t-1}^i \\ 0, & \text{otherwise} \end{cases}$$

$$D_{>j,t-1}^i = \begin{cases} B_{t-1}^i - B_{t-1}^j, & \text{if } B_{t-1}^j < B_{t-1}^i \\ 0, & \text{otherwise} \end{cases}.$$

$D_{<j,t-1}^i$ allows us to investigate how a performer's change in bribe level responds to previously having bribed less than her direct competitor. We call this 'competitor contagion'. $D_{>j,t-1}^i$ does the same in case the performer previously bribed more than the competitor. We denote this effect by 'competitor conformism'.

(iii) A dummy variable T, equal to 1 for the *HI* treatment, and 0, otherwise.

(iv) A dummy variable $Unfair_{t-1}^i$, equal to 1 if the judge in the previous period allocated the prize to the performer with a lower performance and a higher bribe, and 0 otherwise. This allows us to capture possible influences of having experienced unfair judgement in the previous period.

(v) A control variable P_{t-1}^i, capturing *i*'s previous performance as well as a series of personal characteristics C_k^i, including age, gender, educational background, nationality, field of study, financial situation, and risk attitude.

Since the competitor's bribe information is available in both *LI* and *HI*, we add the interaction terms $D_{<j,t-1}^i * T$ and $D_{>j,t-1}^i * T$ to allow the effect of knowing the competitor's bribe to vary between the two treatments.

Together, this gives regression Eq. (2).

$$\Delta B_t^i = \beta_o + \beta_1 * D_{<A,t-1}^i + \beta_2 * D_{>A,t-1}^i + \beta_3 * D_{<j,t-1}^i + \beta_4 * D_{>j,t-1}^i + \beta_5 * D_{<j,t-1}^i * T + \beta_6 * D_{>j,t-1}^i * T + \beta_7 * T + \beta_8 * Unfair_{t-1}^i + \beta_9 * P_{t-1}^i + \sum_k \gamma_k C_k^i + \varepsilon_{ti} \quad (2)$$

We use the coefficients in (2) to test for contagion and conformism. Table 2 lists how the estimated coefficients inform us. For example, a significantly negative estimate of β_2 provides evidence of group conformism.

Table 3
Changes in Bribe: Contagion and Conformism.

β_1 (group contagion): Group average bribe - Own bribe (previous period)	1.153*** (0.084)
p-value permutation test	0.000
β_2 (group conformism): Own bribe - Group average bribe (previous period)	−1.098 (0.041)
p-value permutation test	0.457
β_3 (competitor contagion): Competitor bribe - Own bribe (previous period)	0.387*** (0.060)
p-value permutation test	0.006
β_4 (competitor conformism): Own bribe - Competitor bribe (previous period)	−0.731 (0.049)
p-value permutation test	0.265
β_5 (competitor contagion): [Comp. bribe - Own bribe (prev. per.)]*treatment	−0.472** (0.069)
p-value permutation test	0.023
β_6 (competitor conformism): [Own bribe - Comp. bribe (prev. per.)]*treatment	0.672 (0.071)
p-value permutation test	0.249
β_7: Treatment dummy for *HI*	−0.218 (0.167)
β_8: Unfair judge previous period	−0.226* (0.115)
β_9: Own performance in previous period	0.131*** (0.039)
Observations	3,465

Notes: The dependent variable is the change in bribe (Eq. (2)). Standard errors (clustered by matching group) are in parentheses. For reasons discussed in the main text, we use permutation tests to establish significance of the contagion and conformism variables. Controls include 23 personal characteristics; these are derived from survey questions listed in Appendix C. The regression also includes three dummy variables for Russia, Italy, and China. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Hypotheses 1 predicts contagion, that is, for *HI* $\beta_1 > 0$ and for *LI* $\beta_3 > 0$. It also predicts that contagion is larger in *HI* than in *LI*: $\beta_1 > \beta_3$. Note that we have no hypotheses on the effect of information about the competitor's bribe in *HI* ($\beta_3 + \beta_5$). This is because having information about the matching group may make the competitor's bribe less informative. Similarly, Hypotheses 2 predicts conformism, that is, for *HI* $\beta_2 < 0$ and for *LI* $\beta_4 < 0$. It also predicts that conformism is stronger in *HI* than in *LI*: $\beta_2 < \beta_4$. Again, we have no hypotheses on the effect of information about the competitor's bribe in *HI* ($\beta_4 + \beta_6$).

Note that in Eq. (2) the lagged bribe appears on both sides of the equality (more precisely, in ΔB^i_t, $D^i_{<A,t-1}$, $D^i_{>A,t-1}$, $D^i_{<j,t-1}$, and$D^i_{>j,t-1}$). For this reason, one would obtain positive coefficients for β_1 and β_3 and negative coefficients for β_2 and β_4 even if bribes were determined randomly (and these coefficients would likely be significantly different than zero).[20] To circumvent this problem, we permute the observed bribes 10,000 times and calculate the fraction of observed β values that lie above the estimated coefficient for β_1 and β_3 and below it for β_2 and β_4. The resulting fraction is denoted as the *p*-value in Table 3, below.

Regression results

We report here the results for the pooled data, the regression results per country are presented in Appendix F. In each matching group, there are 8 performers. With ten periods there are 80 bribe observations per matching group and $80 * 12 = 960$ observations.[21] Dropping the first period data for the regression and one participant in *CH* who did not provide personal data leaves us with 864 observations in *NL* and *IT*, 855 in *CH*, and 882 in Russia, for a total of 3465 observations.

presents the results of estimating the coefficients of Eq. (2) for the pooled data.[22] As explained above (Table 2), $\beta_1, \ldots, \beta_6$ are the coefficients related to contagion and conformism. We use the results to test Hypotheses 1 and 2. Recall that the former (the contagion effect) predicts $\beta_1 > 0$ and $\beta_3 > 0$. We strongly reject the null hypothesis of no effect in favor of these contagion effects. Contagion thus appears to be a strong force in the development of descriptive norms.[23] Hypothesis 2 (conformism) predicts $\beta_2 < 0$ and $\beta_4 < 0$. We find no evidence of a conformism effect in response to average bribes in

[20] A Stata code is available upon request to assert this claim.

[21] Recall that in one matching group in *RS* there were 10 performers; see footnote 9. This gives 20 additional observations in *RS*.

[22] As a robustness check, we also estimated these coefficients using a Tobit regression with the lower and upper bounds to the change in bribe determined by the previous bribe and the maximum of ten tokens that can be sent. This leads to no qualitative changes. Details are available upon request.

[23] The results in Appendix F, Table F1 show that contagion is particularly strong in *RS* and *IT*.

the matching group (β_2 is not significantly different than what random bribes would yield). Similarly, there is no evidence of conformism in response to the competitor's bribe (β_4) in the pooled data.[24]

Hypothesis 1 also predicts that group contagion is stronger than competitor contagion ($\beta_1 > \beta_3 > 0$), while Hypothesis 2 predicts stronger group conformism than competitor conformism ($\beta_2 < \beta_4 < 0$). For the reasons mentioned above when discussing how to test the significance of the coefficients, we cannot use F-tests to investigate these predictions. Instead, we consider the estimates per country (cf. Appendix F). The eight comparisons between coefficients that this yields all have the predicted sign. A binomial test rejects a hypothesis of equal likelihoods of contagion being stronger or weaker than conformism ($prob[\beta_1 > \beta_3] = prob[\beta_3 > \beta_1]$ and $prob[\beta_2 > \beta_4] = prob[\beta_4 > \beta_2]$) in favor of the hypothesis that that information at the group level has a stronger affect than information about one's competitor ($p = 0.004$).

Recall that we have no prediction about the effect of information about the competitor's previous bribe in *HI*, that is, we do not formally predict the signs of β_5 and β_6. Intuitively, one may expect the impact of this information in *HI* to be lower than in *LI* because performers also have aggregate information in the former case. This would mean that $\beta_5 < 0$ and $\beta_6 > 0$. The results support this intuition, though the effect is statistically insignificant for conformism. The remaining question, then, is whether in *HI* information about the competitor affects bribes, that is, whether we can reject $\beta_3 + \beta_5 = 0$ and $\beta_4 + \beta_6 = 0$. We again apply permutations to assess these hypotheses. The result shows that both sums are statistically insignificant ($p = 0.64$ and $p = 0.69$, respectively). We conclude that when there is information about the group average, then information about the bribe of the other performer one was previously grouped with plays no role.

Finally, we consider the effects of two of our controls, having previously encountered an unfair judge (β_8) and previous performance (β_9). An unfair judge has a marginally significant negative effect on subsequent bribes. As for previous performance, the effect is positive and statistically significant. A higher previous performance then leads to an increase in current bribes. This may be related to a sense of entitlement.

Differences across countries

We saw above (Table 1) that the strongest prescriptive norms against bribing are observed in *CH*, followed by *NL*, *RS*, and *IT* (cf. Table 1). Because these norms in *IT* and *RS* appear to be similar, as do those in *CH* and *NL*, we simplify the analysis by pairing the countries in this way. This allows us to consider Hypothesis 3 (that cross-country differences in bribes and judges' corruption decrease across periods). We first compare the mean bribe in the two countries with strong prescriptive bribe norms (*NL* and *CH*) to that in the countries with low prescriptive norms (*RS* and *IT*). We do so for the first and second half of the experiment to see whether the difference between the two pairs of countries diminishes as predicted. In the first five periods, mean bribes are 3.05 and 2.71, respectively, in the last five periods they are 3.06 and 2.82; the difference thus reduces from 0.34 to 0.24. To test whether this difference is significant, we regress bribes on dummies representing (i) the second half of the experiment, and (ii) a dummy variable representing *RS* or *IT*, and their interaction (clustering at the level of matching groups). The results show that the coefficient of the interaction term is statistically insignificant ($p = 0.440$). We therefore find no support for the hypothesis that the effect of prescriptive norms on bribes weakens across periods.

4.3. Judges' decisions

To address the second part of Hypothesis 3 (decreased country differences in corruption), we need to consider judges' corruption. First note that our information treatments also affect the judges. In *LI*, a judge only observes the bribes by the two performers in her group. Judges in *HI* also observe the average bribe in their matching groups.

We first define corruption. As introduced above, we say that a judge makes a corrupt decision if she nominates a less successful performer who offers a larger bribe, i.e., the judge allocates the prize to player i while $(P_i < P_j \wedge B_i > B_j)$. This means that there are various situations where a corrupt choice is not possible. In particular, no corruption can occur if there are ties in bribes or performance or if the better performer bribes more.[25] Within each country in each matching group there are four judges, each making ten decisions. We have 12 matching groups per country, which results in 480 observations per country (490 in *RS* because of the one case with a larger matching group). We summarize all possible cases with respect to corruption in Appendix G. Corruption was possible in a total of 582 cases (varying from 121 in *NL*, to 171 in *IT*. Of these opportunities, corruption was observed in 61% of the cases.

To obtain an impression of the determinants of corruption we use a Probit regression. We consider only the cases where corruption was possible and the data pooled across countries. Our focus is first on the effects of bribes and performance on corruption and then on Hypotheses 3. The complete regression results for the pooled data and an overview per country are given in Appendix G. The dependent variable in the probit regressions is a dummy variable indicating whether or not the judge made the corrupt choice. The marginal effects of interest are reported in Table 4. This shows that having the chances of a judge acting corruptly increases with the difference in bribes.[26] A larger performance difference reduces the chances that the poorer performer will win the prize. The information treatment has no significant effect on the likelihood of making a corrupt decision.

[24] We do observe marginally significant conformism in *NL* and strongly significant conformism in *CH*; see Appendix F.

[25] Of course, in the world outside the laboratory, taking a bribe is considered to be a corrupt act irrespective of the subsequent decision. In this respect, our definition may be seen as providing a lower bound on corrupt choices.

[26] Recall that in all these cases, the performer with the higher bribe has the lower performance.

Table 4
Judges' Corruption.

	Marginal effect	Standard error
Treatment dummy for *HI*	0.126	(0.082)
Total bribes	−0.005	(0.009)
Bribe difference	0.052***	(0.013)
Performance difference	−0.054***	(0.014)
Number of observations	582	

Notes: The table reports marginal effects for the probit regression described in the main text. Regression coefficients are given in Table G3 of Appendix G. Standard errors (clustered at the level of individuals and matching groups) are in parentheses. Period (separately per treatment), country dummies, and a set of personal characteristics are included as regressors but not reported in the table; more details are available upon request. *** $p < 0.01$.

The second part of Hypothesis 3 predicts that country differences will diminish over time. For our Probit regression this would mean that the likelihood of corruption becomes more similar across countries as the periods progress. To test this, we use the pooled regression and add interactions between country dummies and the period. The results are presented in Table G4 of Appendix G. To start, we note that baseline corruption is lowest in *CH* and highest in *NL* (setting the period equal to 0, all countries have lower corruption than *NL* and *CH* has the lowest). Over periods, corruption increases in *CH*, while it decreases in *NL*. The other two countries remain in between these two. None of the trends in countries are statistically significant (all $p > 0.24$). These trends do, however, narrow the gap between countries. For example, the (estimated) baseline gap between *NL* and *CH* is 0.23. Over ten rounds, the probability that a judge in *CH* makes a corrupt decision increases by an estimated 0.11 while that for a judge in *NL* decreases by 0.03. Together, this means that the original gap is reduced over time by more than 56%. This remains the largest gap between any two countries. These results suggest that differences between countries indeed diminish over time in spite of the trends per country being statistically insignificant.

5. Discussion and conclusions

To develop anti-corruption policies, it is important to understand the determinants of corrupt behavior. In this paper, we have focused on the role of prescriptive and descriptive norms. We observe that prescriptive norms differ per country. For example, we find evidence of norms against bribery and corruption in China and the Netherlands while both acts are deemed more appropriate in Italy. Chinese and Russian participants are more averse to corruption than the Dutch. This is interesting, because China and Russia score higher on standard indices of corruption than Italy and the Netherlands. This suggests that corrupt practices and anti-corruption norms could coexist (as argued in different contexts by Gërxhani and Schram (2006), Hoffmann and Patel, 2017, and Jackson and Köbis, 2018). We believe that the decision to bribe or act corruptly is affected by a cost-benefit analysis of the decision involved. The costs include the disutility of acting against a social norm. Even if these costs are the same in two distinct cultures, the benefits may differ. In some countries, for example, it might simply be far more difficult than in others to realize economic gains from trade if one does not participate in bribery and corruption.[27]

We show that the disclosure of information about a common level of corruption induces players to adjust their behavior and to converge towards this common level. Similar results have been received in other contexts (Bicchieri et al., 2019; Köbis et al., 2019). Our results show that the convergence process is asymmetric. We observe contagion (increasing bribes when one observes that others are bribing more) but not conformism (reducing bribes when others are bribing less). The relative strength of contagion compared to conformism that we observe suggests that a policy that reveals average bribe levels will reduce bribes by those who are above this average less than it will increase the bribes of those below average. Together, such a policy would then increase corruption.

Italy appears to be the odd one out when it comes to creating a situation where corruption is possible.[28] In contrast to the other countries, our Italian participants do not adjust their bribes to their performance in an attempt to avoid putting the judge in a position where corruption is possible. As a consequence there are more situations with a potential for corruption in Italy. Yet, we do not observe more corruption in our Italian experiments. Furthermore, in contrast to judges in the Netherlands and China (but similar to those in Russia), Italian judges do not increase corrupt choices when information about average bribes is disclosed. It appears that corruption is simply accepted from the start as a 'reasonable' possibility. This is supported by the prescriptive norms that we elicited in Italy; Italian participants in our experiment tend to find bribes and corruption (mildly) appropriate.

[27] China is an interesting case, because President Xi Jinping's has implemented a number of anti-corruption reforms since taking office in 2013. This anti-corruption campaign may affect norm measurement while not (yet) affecting behavior in the experiment (which subjects may not relate to the reforms) or in the field. We consider further corruption experiments in China -both in the laboratory and in the field- to be an interesting avenue for future research.

[28] It should be noted that our Italian sessions were run in Bologna and our conclusions about our participants in 'Italy' therefore reflect behavior in Northern Italy. Stark cultural differences exist between the North and the South of Italy (e.g., Bigoni et al., 2016).

We believe that the laboratory provides a useful and necessary environment to isolate the effects of prescriptive and descriptive norms in different cultures. The relevance of our conclusions are, of course, dependent on the external validity of our experiments. In particular, one might ask whether the subject pools of university students are representative of the general populations in the countries. We would argue that cultural differences between students are, if anything, smaller than between general populations. In this respect, any differences we find across countries might provide a lower bound for more general cross-country differences. The fact that we do find differences speaks to the external validity of our results. This is not very surprising; social norms evolve slowly and students may be expected to be strongly influenced by the norms of their parents and educators.[29] Moreover, our main interest lies in studying the occurrence of contagion and conformism. In doing so, we are not so much interested in measuring the levels of these phenomena, as in their comparative statics (e.g., how contagion is affected by increased information). Comparative statics with a student subject pool in the laboratory tend to mirror those observed in the field (e.g., Camerer, 2015).

Nevertheless, an interesting avenue for future research would be to use other subject pools and conduct similar experiments in the field. Another interesting direction might be to incorporate punishment into this game. Punishment adds to the costs in the cost-benefit analysis referred to above. Since punishment continues to be one of the most common methods used to combat corruption, it would be highly policy-relevant to allow performers to punish judges if the latter are corrupt. One could then analyze how this affects, first, the propensity of judges to be corrupt and, second, the behavior of bribers.

Appendix A. Instructions for corruption game

Please find below the instructions used in Amsterdam. The same instructions were used in China, Italy and Russia translated in Chinese,Italian and Russian respectively.

Treatment with no information.

Welcome to our experiment!

This is an experiment on decision-making where you will earn money. The amount of money you earn will depend upon the decisions you make and on the decisions other people make. You will be paid privately at the end of the experiment, there is no obligation to tell others how much you earn. In the experiment you will remain anonymous. The experiment will take approximately one-and-a-half hour.

You have already received 7 Euros for showing up. Your total earnings will be the sum of this 7 euros and your payoffs in the experiment. In this experiment we use experimental tokens. At the end of the experiment these will be exchanged to euros at a rate of 1 token = 1 euro and you will be paid in cash.

Please read the instructions carefully and do not communicate with each other during the experiment. If you have a question, please raise your hand and an experimenter will come to help you. There are 24 participants in this experiment. All the participants are randomly divided into two types. Sixteen of you are players, eight are judges. There will be 10 independent periods in this experiment.

At the beginning of every period each participant of each type receives 10 tokens. Groups of three are formed, each consisting of two players and one judge. The two players will perform a task during 5 minutes. The task is explained below. The task gives each of these players a score. The better a player does at the task, the higher will be her or his score. After they have finished, the judge will decide on the winner. The winner will receive 10 points. The judge will not perform the task. He or she can give the prize to either of the two players in his or her group. Please note that the judge must allocate this prize to one of the two players. The amount of the prize will be added to the final payoff of the players. This prize will NOT be deduced from the earnings of the judge. Before the judge decides, each of the two players decides whether to transfer tokens to the judge in their group. They may transfer any amount between 0 and 10. As soon as both players have made their choices, the judge will see on her or his computer screen the information about the score of both players and their transfers. Then a new period starts and all will be randomly rematched into new groups of three (two players and one judge). You will not change your type, players remain players and judges remain judges. The rules for all 10 periods are identical. All periods are independent.

At the end of the experiment every participant of each type will receive his or her earnings from one randomly picked period plus the show-up fee of 7 euros.

The Task. You will see two matrices on the computer screen. Each matrix has 10 rows and 10 columns and is filled with randomly generated numbers. Your task is to find the largest number in each of the matrices and add them up. You are not allowed to use calculators but you can use the paper and pencil that you have found on your desk. After entering the number, the computer will tell you whether it is correct or incorrect (the time will continue to run while you see the result). Then, irrespective of whether your answer is correct or incorrect a new pair of matrices will appear. New matrices will appear as long as you are within the 5 minutes limit with the max of 10 matrices. When the 5 minutes limit ends, participants A and B will see the total number of correct solutions that they have achieved.

[29] More generally, there is high correlation between behavior of student participants in the laboratory, Amazon Turk participants in experiments and choices by representative samples (Snowberg and Yariv, 2021).

Treatment with information

Welcome to our experiment!

This is an experiment on decision-making where you may earn money. The amount of money you earn will depend upon the decisions you make and on the decisions other people make. You will be paid privately at the end of the experiment, there is no obligation to tell others how much you earn. In the experiment you will remain anonymous. The experiment will take approximately one-and-a-half hour.

You have already received 7 Euros for showing up. Your total earnings will be the sum of this 7 euros and your payoffs in the experiment. In this experiment we use experimental tokens. At the end of the experiment these will be exchanged to euros at a rate of 1 token = 1 euro and you will be paid in cash.

Please read the instructions carefully and do not communicate with each other during the experiment. If you have a question, please raise your hand and an experimenter will come to help you. There are 24 participants in this experiment. All the participants are randomly divided into two types. Sixteen of you are players, eight are judges. There will be 10 independent periods in this experiment.

At the beginning of every period each participant of each type receives 10 tokens. Groups of three are formed, each consisting of two players (A and B) and one judge. The two players will perform a task during 5 minutes. The task is explained below. The task gives each of these players a score. The better a player does at the task, the higher will be her or his score. After they have finished, the judge will decide on the winner. The winner will receive 10 points. The judge will not perform the task. He or she can give the prize to either of the two players in his or her group. Please note that the judge must allocate this prize to one of the two players. The amount of the prize will be added to the final payoff of the players. This prize will NOT be deduced from the earnings of the judge.

Before the judge decides, each of the two players decides whether to transfer tokens to the judge in their group. They may transfer any amount between 0 and 10. As soon as both players have made their choices, the judge will see on her or his computer screen the information about the score of both players and their transfers. Then a new period starts and all will be randomly rematched into new groups of three (two players and one judge). You will not change your type, players remain players and judges remain judges. The rules for all 10 periods are identical. All periods are independent. At the end of the experiment every participant of each type will receive his or her earnings from one randomly picked period plus the show-up fee of 7 euros.

At the end of each period we will show you on your monitor information about what occurred in four randomly chosen groups (eight players and four judges). This information will involve the average transfers by players to the judges.

The Task. You will see two matrices on the computer screen. Each matrix has 10 rows and 10 columns and is filled with randomly generated numbers. Your task is to find the largest number in each of the matrices and add them up. You are not allowed to use calculators but you can use the paper and pencil that you have found on your desk. After entering the number, the computer will tell you whether it is correct or incorrect (the time will continue to run while you see the result). Then, irrespective of whether your answer is correct or incorrect a new pair of matrices will appear. New matrices will appear as long as you are within the 5 minutes limit with the max of 10 matrices. When the 5 minutes limit ends, players A and B will see the total number of correct solutions that they have achieved.

Appendix B. Instructions for social norm measurement

Welcome to our experiment!

This is an experiment on decision-making where you may earn money. The amount of money you earn will depend upon the decisions you make and on the decisions other people make. You will be paid privately at the end of the experiment, there is no obligation to tell others how much you earn. In the experiment you will remain anonymous. The experiment will take approximately half an hour. You have already received 5 Euros for showing up. Your total earnings will be the sum of this 5 euros and your payoffs in the experiment.

Please read the instructions carefully and do not communicate with each other during the experiment. If you have a question, please raise your hand and an experimenter will come to help you.

There are 39 participants in this experiment.

Assume that there is the following situation.

There is a group of three that consists of two players, A and B, and one judge. Initially each player has 10 points. The two players perform a task during 5 minutes. The task gives each of these players a score. The better a player does at the task, the higher will be her or his score. On average, players score 12 on this task. After the two players have finished, the judge will decide on the winner. The winner will receive 10 points. The judge will not perform the task. He or she can give the prize to either of the two players. Please note that the judge must allocate this prize to one of the two players. The amount of the prize will be added to the final payoff of the players. This prize will NOT be deduced from the earnings of the judge.

Before the judge decides, each of the two players decides whether to transfer tokens to the judge. They may transfer any amount between 0 and 10. As soon as both players have made their choices, the judge will see on her or his computer screen the information about the score of both players and their transfers.

You will not participate in this situation. Instead, your task today is to analyze the potential outcomes. We are interested in what you think about what most people feel is the appropriate choice in certain situations. So, we are not asking what you personally think that one should do, but what you think that most people feel is appropriate.

On the following pages, you will read descriptions of a series of situations. These descriptions correspond to situations in which one person, either the judge or one of the players, must make a decision. For each situation, you will be given a description of the decision faced by the individual. This description may include several possible choices available to the individual. After you read the description of the decision, you will be asked to evaluate the different possible choices and to decide, for each of the possible actions, whether taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior.” By socially appropriate, we mean behavior that most people agree is the “correct” or “ethical” thing to do. Another way to think about what we mean is that if the individual A were to select a socially inappropriate choice, then someone else might be angry at her or him for doing so.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

Your payoffs for this experiment will be formed in the following way. You will be asked to respond to 17 different situations (split across four answer sheets). At the end we will randomly select one of these 17 situations for payment. For this situation we will check which answer was given most often (this is called the “modal” answer). You will receive 10 euros if you guess the modal answer among all the 39 participants. If you do not guess the modal answer, you will receive only the participation fee. In other words, if you give the same response as that most frequently given by other people, then you will receive an additional 10 euro.

To summarize, your task for each situation is to predict the answer that is most often chosen by everyone in this room today. If you guess correctly, you will earn 10 euros.

Answer Sheet 1

Recall that the average player scores 12 on this task. Consider the situation where

- player A scores 20
- player B scores 12

In the table below we show different transfers by player B and different decisions by the judge. Indicate for each of the situations whether the judge’s decision is “very socially inappropriate”, “somewhat socially inappropriate”, “somewhat socially appropriate”, or “very socially appropriate”. For each situation, you may indicate your response by marking the corresponding cell with an “X”. Remember, if one of these situations is the one chosen for payment, then you will receive 10 euros if your choice is the response most often given by other people in today’s session.

The judge decision is:						
Player A transfers:	Player B transfers:	The judge gives the prize to:	very socially inappropriate	somewhat socially inappropriate	somewhat socially appropriate	very socially appropriate
7	3	A				
7	3	B				
3	7	A				
3	7	B				

Notes: Recall that player A has a score of 20 and player B has a score of 12. Note that answering for these situations requires that you place four X (one in each row).

Answer Sheet 2

Recall that the average player scores 12 on this task. Consider the situation where

- player A scores 20
- player B scores 12

In the table below we show different transfers by player B and different decisions by the judge. Indicate for each of the situations whether the judge’s decision is “very socially inappropriate”, “somewhat socially inappropriate”, “somewhat socially appropriate”, or “very socially appropriate”. For each situation, you may indicate your response by marking the corresponding cell with an “X”. Remember, if one of these situations is the one chosen for payment, then you will receive 10 euros if your choice is the response most often given by other people in today’s session.

The judge decision is:						
Player A transfers:	Player B transfers:	The judge gives the prize to:	very socially inappropriate	somewhat socially inappropriate	somewhat socially appropriate	very socially appropriate
3	0	A				
3	0	B				
0	3	A				
0	3	B				

Notes: Recall that player A has a score of 20 and player B has a score of 12. Note that answering for these situations requires that you place four X (one in each row).

Answer Sheet 3

Recall that the average player scores 12 on this task. In the table below we show different scores for one of the players (which we call player C) and different transfers by the same player. Indicate for each of the situations whether this transfer is “very socially inappropriate”, “somewhat socially inappropriate”, “somewhat socially appropriate”, or “very socially appropriate”. For each situation, you may indicate your response by marking the corresponding cell with an “X”. Remember, if one of these situations is the one chosen for payment, then you will receive 10 euros if your choice is the response most often given by other people in today’s session.

Player C’s transfer decision is:					
Player C’s score:	Player C transfers:	very socially inappropriate	somewhat socially inappropriate	somewhat socially appropriate	very socially appropriate
12	0				
20	3				
4	7				
4	0				

Notes: Note that answering for these situations requires that you place four X (one in each row).

Answer Sheet 4

Recall that the average player scores 12 on this task. In the table below we show different scores for one of the players (which we call player C) and different transfers by the same player. Indicate for each of the situations whether this transfer is “very socially inappropriate”, “somewhat socially inappropriate”, “somewhat socially appropriate”, or “very socially appropriate”. For each situation, you may indicate your response by marking the corresponding cell with an “X”. Remember, if one of these situations is the one chosen for payment, then you will receive 10 euros if your choice is the response most often given by other people in today’s session.

Player C’s transfer decision is:					
Player C’s score:	Player C transfers:	very socially inappropriate	somewhat socially inappropriate	somewhat socially appropriate	very socially appropriate
12	7				
12	3				
20	0				
20	7				
4	3				

Notes: Note that answering for these situations requires that you place four X (one in each row).

Appendix C. The survey

1. Your age.
2. Your nationality[30]
   - the Netherlands
   - Western Europe excluding the Netherlands
   - Russia
   - Eastern Europe excluding Russia
   - Italy
   - Southern Europe excluding Italy
   - China
   - Other
3. Your gender:
   - male;
   - female
4. Your height (in centimeters)
5. Your field of studies:
   - economics, finance, management;
   - social science, psycology, political science ;
   - law ;
   - international relation ;
   - mathematics, computer science ;
   - humanities ;
   - media, journalism, design ;
   - other
6. Do you like risk or avoid risk?
   - I like risk;
   - rather like risk;

[30] This question was included only in the survey in Amsterdam.

- neutral to risk;
- rather avoid risk;
- avoid risk

7. Which statement most accurately describes the financial situation of your family?
   - money is not enough to survive;
   - enough money only for urgent needs;
   - There is enough money for daily expenses, but already buying clothes requires savings;
   - There is enough money, even some savings, but large purchases need to be planned in advance;
   - We can afford large expenses at the first necessity.
8. Given all the circumstances, how satisfied are you with your life in general? (from 1 "completely dissatisfied" to 10 "completely satisfied")
9. In your opinion, in general, most people can be trusted, or when communicating with other people caution never hurts? Please mark the position on the scale, where 1 means "You have to be very careful with other people" and 10 means "Most people can be completely trusted"
10. Some people feel that they have complete freedom of choice and control their lives, while other people feel that what they are doing does not have a real impact on what is happening to them. To what extent are these characteristics applicable to you and your life? Please mark the position on the scale, where 1 means "I do not have freedom of choice" and 10 means "I have total freedom of choice"

Appendix D. Descriptive statistics

	NL (T)		NL (C)		IT (T)		IT (C)		CN (T)		CN (C)		RS (T)		RS (C)	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Age	21.01	2.45	22.38	5.14	22.89	2.68	22.07	2.56	21.73	2.27	21.61	2.11	21.77	5.60	20.69	3.49
Female	0.50	0.50	0.40	0.49	0.56	0.50	0.54	0.50	0.69	0.46	0.61	0.49	0.51	0.50	0.47	0.50
Risk attitude	3.07	1.03	2.97	1.11	2.85	1.14	2.79	1.11	2.92	1.03	2.57	0.91	3.12	1.02	2.81	0.91
Income	3.92	0.86	4.06	0.64	3.71	0.70	3.50	0.87	3.52	0.73	3.56	0.58	3.83	0.70	3.65	0.65
Life satisfaction	7.29	2.00	7.53	1.49	7.51	1.44	7.42	1.79	6.49	1.74	6.93	1.54	7.27	1.88	6.71	2.07
Trust	5.57	1.92	5.89	2.15	4.78	2.08	4.99	2.34	5.89	1.89	6.31	1.80	5.36	1.98	5.22	2.03
Freedom	7.40	1.82	7.31	1.78	7.24	1.81	7.32	1.94	6.41	1.69	6.96	1.75	7.30	1.79	7.32	1.75
Major in economics	0.67	0.47	0.47	0.50	0.28	0.45	0.31	0.46	0.58	0.49	0.62	0.48	0.63	0.48	0.62	0.48
Observations	720		720		720		720		720		720		810		720	

Appendix E. Social norm indices

In this appendix, we provide details about the prescriptive norms measures presented in the main text. To analyze subjects' choices in the Krupka-Weber coordination games, we first convert their responses to a numerical index. Following Krupka and Weber (2013), a rating of "very socially inappropriate" is given a score of -1, "somewhat socially inappropriate" a score of $-\frac{1}{3}$, "somewhat socially appropriate" a score of $\frac{1}{3}$, and "very socially appropriate" a score of 1. The precise conversion used is somewhat arbitrary, of course. Our conclusions do not change if we score the 'somewhat' categories as $-\frac{1}{2}$ and $\frac{1}{2}$. We arrive at an index that varies from -1 (very socially inappropriate) to 1 (very socially appropriate). Table E.1 shows the mean converted indices per country for each of the bribe scenarios we presented participants with.

For each possible score (4, 12, or 20) the scenarios involved a case without a bribe and two additional bribe levels. We are interested in how the participants in various countries view the appropriateness of bribes in comparison to the no-bribe case. For this purpose, we normalize the appropriateness index for the no-bribe scenario in each country and at each performance level to 0. The normalized index for a bribe thus represents how appropriate participants in that country consider that bribe by a person with a specific performance level compared to that person not giving a bribe. The

Table E1
Prescriptive Norms.

	NL	*RS*	*IT*	*CH*
	mean	mean	mean	mean
Score: 4, bribe: 0	0.50	0.02	0.38	0.25
Score: 4, bribe: 3	−0.06	−0.29	−0.13	−0.27
Score: 4, bribe: 7	−0.21	−0.15	−0.23	−0.15
Score: 12, bribe: 0	0.30	0.19	−0.05	0.35
Score: 12, bribe: 3	0.23	0.42	0.18	0.08
Score: 12, bribe: 7	0.03	0.33	0.03	0.10
Score: 20, bribe: 0	0.21	0.56	−0.42	0.58
Score: 20, bribe: 3	0.11	0.44	−0.17	0.30
Score: 20, bribe: 7	0.13	0.27	0.30	0.23
Observations	39	32	40	40

Table E2
Normalized Prescriptive Norms.

Score	Bribe	NL	RS	IT	CH
4	3	−0.56	−0.31	−0.52	−0.52
	7	−0.72	−0.17	−0.62	−0.40
12	3	−0.07	0.23	0.23	−0.27
	7	−0.27	0.15	0.08	−0.25
20	3	−0.10	−0.13	0.25	−0.28
	7	0.09	−0.29	0.72	−0.35

Notes: Numbers in the cells give the mean appropriateness of the bribe concerned, for the given score *minus* the mean appropriateness of giving no bribe at that score.

Table E3
Prescriptive norms: judges' behavior.

	NL	RS	IT	CH
Judge nominates A				
A bribes: 7, B bribes: 3	0.66	0.88	0.35	0.73
A bribes: 3, B bribes: 0	0.61	0.88	0.50	0.72
A bribes: 3, B bribes: 7	0.28	0.21	−0.13	0.63
A bribes: 0, B bribes: 3	0.32	0.33	−0.23	0.62
Judge nominates B				
A bribes: 7, B bribes: 3	−0.50	−0.69	−0.22	−0.73
A bribes: 3, B bribes: 0	−0.62	−0.79	−0.38	−0.67
A bribes: 3, B bribes: 7	−0.13	−0.23	0.15	−0.47
A bribes: 0, B bribes: 3	−0.23	−0.33	0.13	−0.45
Observations	39	32	40	40

Notes: The table reports the index of social appropriateness. 1=very socially appropriate. −1 = very socially inappropriate. In all cases, A scores 20 and B scores 12.

normalized indices are presented in Table E2. Finally, to obtain a measure across all performance levels we simply take the average across the six normalized indices per country. These averages are given in the row 'Bribes' of Table 1 in the main text.

Consider first a performer with a score of 4 (below the average performance of 12). Here, the pattern is similar across countries. Bribing is considered less appropriate than not bribing (the normalized index is negative in all countries for both bribe levels). The effect is about twice as strong in *NL* and *IT* than in *RS*, with *CH* in between. These effects are statistically significantly different from zero, except for a bribe of 7 in *RS* (PtT, $p < 0.05$ for all except *RS*, where $p = 0.39$). The normalized index is insignificantly different between bribes of 3 and 7 in all countries (PtT, all $p > 0.10$). Moreover, for both bribe levels all pairwise comparisons between countries are statistically indistinguishable (PtT, $p > 0.10$) except the comparison between *NL* and *RS* for a bribe of 7 (PtT, $p = 0.08$).

When performance is average (a score of 12) bribing becomes more acceptable in all countries. In fact, the only significant difference (PtT, $p = 0.01$) with a bribe of zero appears for a bribe of 3 in *CH*, where bribing is still considered to be less appropriate than not bribing. For all other countries, $p > 0.10$ so there the social norms concerning bribing 3 or bribing nothing are indistinguishable. Across countries, bribing 3 is considered significantly less appropriate in *CH* than in *RS* and *IT* (PtT, $p < 0.01$ in both cases). All other country differences are statistically insignificant (PtT, $p > 0.17$). For a bribe of 7, only the difference between *RS* and *CH* is marginally significant (PtT, $p = 0.10$); for all other comparisons $p > 0.13$.

Finally, when performance is high (20), bribing is again deemed more appropriate than not bribing in *IT* (PtT, $p = 0.07$ for a bribe of 3; $p < 0.01$ for a bribe of 7). The reverse still holds for *CH* (PtT, $p = 0.04$ for a bribe of 3; $p = 0.05$ for a bribe of 7). The Dutch (*NL*) are now indifferent about the appropriateness of bribing versus not bribing (PtT, $p > 0.1$ for both bribe levels). An interesting case is *RS*. While they consider bribing inappropriate compared to not bribing for low performers (a score of 4) and appear to be indifferent for average performers, they revert to finding it relatively inappropriate for high performers. This effect is, however, insignificant for a bribe of 3 (PtT, $p = 0.32$) and only marginally significant for a bribe of 7 (PtT, $p = 0.06$). The pattern does suggest that Russians may feel that bribery is inappropriate for those who deviate from average performance.This is reminiscent of the finding for Russia reported in Gächter and Herrmann (2007). In their public good games with punishment, people are punished for any deviation from the average contribution, not just for free riding. Across countries, *IT* clearly stands out. For a bribe of 3, the differences with *NL*, *RS*, and *CH* are all significant (PtT, $p = 0.09$, $p = 0.05$, $p < 0.01$, respectively). The differences are even stronger for a bribe of 7 (PtT, $p < 0.01$ for all comparisons). All other country comparisons are statistically insignificant (PtT, all $p > 0.34$).

Next, we turn to the appropriateness scores concerning judge's behavior. First note that there is no straightforward way in which to normalize these norms like we did for performers. To start, Table E3 shows the mean (converted) appropriate scores for all scenarios concerning judges' decisions.

All situations in Table 3 reflect cases where performer A outperforms performer B (20 *vs.* 12). The top panel shows appropriateness scores for situations where A receives the prize, and in the lower panel B is chosen as winner. For *NL*, *RS*, and *CH*, it is immediately clear that it is deemed appropriate to award the prize to A, the better performer (all numbers in the top panel are positive and all number in the lower panel are negative).Nevertheless, bribes do matter in *NL*, *RS*, and *CH*. The appropriateness index of giving the prize to A diminishes if B bribes more than A. Similarly, the *in*appropriateness of giving the prize to B diminishes if B bribes more than A. In *IT*, the norms appear to be different than in the other countries. Here, it is deemed most appropriate that the prize is awarded to the performer who bribes more (the index is negative if the judge gives the prize to the performer with the lower bribe and positive if the opposite is true, irrespective of A's better performance).

Our main interest lies in the case where the judge is corrupt. We consider this to be the case if she awards the prize to B when (i) A has the better performance, and (ii) B bribes more than A. This occurs in the last two cases of Table E3. As a measure of the prescriptive norm with respect to judges' corruption, we take the average appropriateness of these two cases. This measure of anti-corruption norms is given in the row 'Judges' Corruption' in Table 1 of the main text.

With respect to the cases where the judge is deemed corrupt, we first note that for all countries the difference between Bribes of 0 (A) and 3 (B) as opposed to 3 (A) and 7 (B) is statistically insignificant (PtT, all $p > 0.32$). For bribes of 3 and 7 corruption is deemed significantly more inappropriate in *NL*, *RU*, and *CN* than in *IT* (PtT, $p = 0.06$, $p = 0.01$, $p < 0.01$, respectively). All other country comparisons are statistically insignificant (PtT, all $p > 0.20$). For bribes of 0 and 3, corruption is deemed significantly more inappropriate in *CN* than in *NL* (PtT, $p = 0.04$) or *IT* (PtT, $p < 0.01$). The difference between *IT* and *RS* is also significant (PtT, $p = 0.04$). All other country comparisons are statistically insignificant (PtT, all $p > 0.12$).

Appendix F. Contagion and conformism at the country level (Table F1)

Table F1
Bribes: Contagion and Conformism.

	Netherlands	Russia	Italy	China	Pooled
β_1 (group contagion)	0.995	1.130***	1.297***	1.103	1.153***
Group average bribe - Own bribe (previous period)	(0.127)	(0.182)	(0.130)	(0.173)	(0.084)
p-value permutation test	0.169	0.006	0.001	0.202	0.000
β_2 (group conformism)	−1.060	−1.165	−1.081	−1.067	−1.098
Own bribe - Group average bribe (previous period)	(0.116)	(0.102)	(0.113)	(0.076)	(0.041)
p-value permutation test	0.703	0.447	0.477	0.272	0.457
β_3 (competitor contagion)	0.423	0.544***	0.441**	0.163	0.387***
Competitor bribe - Own bribe (previous period)	(0.063)	(0.113)	(0.051)	(0.109)	(0.060)
p-value permutation test	0.436	0.000	0.017	0.984	0.006
β_4 (competitor conformism)	−0.709*	−0.688	−0.710	−0.855***	−0.731
Own bribe - Competitor bribe (previous period)	(0.075)	(0.110)	(0.077)	(0.094)	(0.049)
p-value permutation test	0.052	0.949	0.598	0.006	0.265
β_5 (competitor contagion)	−0.366	−0.659***	−0.539**	−0.284	−0.472**
[Comp. bribe - Own bribe (prev. per.)]*treatment	(0.078)	(0.131)	(0.063)	(0.116)	(0.069)
p-value permutation test	0.786	0.000	0.021	0.811	0.023
β_6 (competitor conformism)	0.571	0.716	0.714	0.775*	0.672
[Own bribe - Comp. bribe (prev. per.)]*treatment	(0.084)	(0.182)	(0.116)	(0.109)	(0.071)
p-value permutation test	0.392	0.423	0.452	0.076	0.249
β_7	−0.212	0.029	−0.217	−0.562	−0.218
Treatment dummy	(0.427)	(0.370)	(0.223)	(0.356)	(0.167)
β_8	−0.204	−0.280	0.029	−0.356	−0.226*
Unfair judge previous period	(0.279)	(0.175)	(0.194)	(0.272)	(0.115)
β_9	0.221**	0.053	−0.013	0.133**	0.131***
Own performance in previous period	(0.072)	(0.052)	(0.073)	(0.052)	(0.039)
Observations	864	882	864	855	3,465

Notes: The dependent variable is the change in bribe (Eq. (2)). Standard errors (clustered by matching group) are in parentheses. For reasons discussed in the main text, we use permutation tests to establish significance of the contagion and conformism variables. Controls include 23 personal characteristics; these are derived from survey questions listed in Appendix C. The pooled regression also includes three dummy variables for *Russia*, *Italy*, and *China*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix G. Corruption opportunities

In this appendix, we provide an overview of the situations in which judges have the possibility to make a corrupt decision. Recall that this occurs when one of the performers has a better performance and the other bribes more. Table G1 summarizes all possible cases and reports the number of observations in each. A few interesting things can be noted. First, corruption is possible in 25.2%, 29.0%, 35.6%, and 30.8% of the cases in *NL*, *RS*, *IT*, and *CH*, respectively. Second, the cases where one performer has a higher score than the other (406, 403, 402, 406, respectively) is more or less the same across countries.

Table G1
Possibility of corruption: distribution.

Case		Country				Possibility of corruption
		NL	RS	IT	CH	
equal performance $P_i = P_j$	equal bribe $B_i = B_j$	14	28	10	12	corruption not possible
	unequal bribe $B_i \neq B_j$	60	59	68	62	corruption not possible
unequal performance $P_i \neq P_j$	equal bribe $B_i = B_j$	69	89	63	66	corruption not possible
	player with higher performance bribes more	216	172	168	192	corruption not possible
	player with higher performance bribes less	121	142	171	148	corruption possible
Total number of observations		480	490	480	480	

Notes: Cell entries give the number of observations for each situation and country.

Table G2
Determinants of corrupt judges per country.

	NL	*RS*	*IT*	*CH*
Treatment	0.350**	−0.086	0.125	0.272*
	(0.160)	(0.161)	(0.168)	(0.140)
Total bribes	0.003	0.006	−0.023	0.004
	(0.011)	(0.021)	(0.015)	(0.017)
Bribe difference	0.077***	0.025	0.049**	0.053**
	(0.023)	(0.026)	(0.022)	(0.024)
Performance difference	−0.020	−0.017	−0.085***	−0.110***
	(0.022)	(0.022)	(0.024)	(0.032)
Observations	121	135	168	147

Notes: The table reports marginal effects for the probit regression described in the main text. Standard errors (clustered at the level of individuals and matching groups) are in parentheses. Period (separately per treatment), and a set of personal characteristics are included as regressors but not reported in the table; more details are available upon request. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Third and interestingly, when one performer scores better than the other and the bribes differ, then the fraction of times that the better performer also bribes more differs across countries. This fraction is 0.64 in *NL*, 0.55 in *RS*, 0.50 in *IT*, and 0.57 in *CH*. We use two-sided binomial tests to investigate whether these fraction differ from 50%. The results show that the fraction is significantly different than 50% in *NL* ($p < 0.01$) and *CH* ($p = 0.02$). In *RS* ($p = 0.10$) and *IT* ($p = 0.91$), however, when one performer scores better than the other, she is equally likely to bribe more or less than this other performer. This suggests that there is a tendency in *NL* and *CH* to avoid situations where the judge might act corruptively, by 'outbribing' the other if one outperforms her. In *RS* and *IT* there is no relationship between having the better performance and the higher bribe. Recall that our prescriptive norm measurements in *RS* and *IT* show little evidence of norms that find bribing inappropriate (see Table 1 in the main text). This is consistent with our finding here, that 'winning the bribe' is unrelated to relative performance in *RS* and *IT*.

Next, Table G2 presents the marginal effects of treatment, bribes and performance per country (compare to the pooled effects presented in Table 4 of the main text). Note that the effect of bribe differences that appears in the pooled data is not significant in *RS* while that of performance differences is only significant in *IT* and *CH*.

Table G3
Probit regression corruption.

	coefficient	s.e.
RS	−0.282	0.342
IT	−0.236	0.363
CH	−0.462	0.331
Period	0.014	0.027
Treatment	0.367	0.238
Total bribes	−0.014	0.025
Bribe difference	0.152	0.039***
Performance difference	−0.156	0.041***
Period*Treatment	−0.040	0.037
Individual characteristics		
N	582	

Notes: Cells give the estimated coefficients and standard errors (s.e.) for variables depicted in the first column. The independent variable is a judges choice to be corrupt or not. Country effects are in comparison to *NL*. No individual characteristics are statistically significant at the 5% level. Majoring in economics or business (positive), majoring in maths (positive) and being from Eastern Europe (negative) have a significant effect at the 10% level. *** $p < 0.01$.

Table G4
Probit regression corruption.

	coefficient	s.e.
RS	−0.595	0.458
IT	−0.165	0.460
CH	−0.666	0.439
*NL**Period	−0.007	0.047
*RS**Period	0.049	0.040
*IT**Period	−0.021	0.040
*CH**Period	0.033	0.039
Treatment	0.374	0.237
Total bribes	−0.016	0.025
Bribe difference	0.149	0.039***
Performance difference	−0.151	0.041***
Period*Treatment	−0.040	0.036
Individual characteristics		
N	582	

Notes: Cells give the estimated coefficients and standard errors (s.e.) for variables depicted in the first column. The independent variable is a judges choice to be corrupt or not. Country effects are in comparison to *NL*. No individual characteristics are statistically significant at the 5% level. Majoring in economics or business (positive), majoring in maths (positive) and being from Eastern Europe (negative) have a significant effect at the 10% level.

Finally, Table G3 presents the probit results for the regression underlying Table 4 of the main text and Table G4 shows the probit estimates for the regression used to investigate country differences in corruption over time (Hypotheses 3).

References

Abbink, K., Freidin, E., Gangadharan, L., Moro, R., 2018. The effect of social norms on bribe offers. J. Law Econ. Organ. 34 (3), 457–474.
Abbink, K., Irlenbusch, B., Renner, E., 2002. An experimental bribery game. J. Law Econ. Organ. 18 (2), 428–454.
Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., Gangadharan, L., 2009. Gender, culture, and corruption: insights from an experimental analysis. South. Econ. J. 663–680.
Armantier, O., Boly, A., 2011. A controlled field experiment on corruption. Eur. Econ. Rev. 55 (8), 1072–1082.
Balafoutas, L., Sandakov, F., Zhuravleva, T., 2021. No moral wiggle room in an experimental corruption game. Front. Psychol. 3509.
Banuri, S., Eckel, C., 2012. The Effects of Sanctions on Bribery: US versus Pakistan. Technical Report. CBEES Working Paper Series 09-01.
Banuri, S., Eckel, C., 2012b. Experiments in culture and corruption: a review (vol. 15, pp. 51–76).
Banuri, S., Eckel, C., 2015. Cracking down on bribery. Soc. Choice Welfare 45 (3), 579–600.
Barr, A., Serra, D., 2009. The effects of externalities and framing on bribery in a petty corruption experiment. Exp. Econ. 12 (4), 488–503.
Barr, A., Serra, D., 2010. Corruption and culture: an experimental analysis. J. Public Econ. 94 (11–12), 862–869.
Bicchieri, C., 2005. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press.
Bicchieri, C., Dimant, E., Gaechter, S., Nosenzo, D., 2019. Observability, Social Proximity, and the Erosion of Norm Compliance. Available at SSRN 3355028.
Bicchieri, C., Fukui, Y., 1999. The great illusion: ignorance, informational cascades, and the persistence of unpopular norms. Bus. Ethics Q. 9 (1), 127–155.
Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., Pancotto, F., 2016. Amoral familism, social capital, or trust? The behavioural foundations of the Italian North–South divide. Econ. J. 126 (594), 1318–1341.
Brennan, G., Eriksson, L., Goodin, R., Southwood, N., 2013. Explaining Norms. Oxford University Press.
Camerer, C., 2015. The Promise of Lab-field Generalizability in Experimental Economics: A Reply to Levitt and List (2007), pp. 249–295.
Cameron, L., Chaudhuri, A., Nisvan, E., Gangadharan, L., 2009. Propensities to engage in and punish corrupt behavior: experimental evidence from Australia, India, Indonesia and Singapore. J. Public Econ. 93 (8), 843–851.
Chen, D., Schonger, M., Wickens, C., 2016. oTree—an open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Finance 9, 88–97.
Cohn, A., Fehr, E., Maréchal, M.A., 2014. Business culture and dishonesty in the banking industry. Nature 516 (7529), 86.
Cooray, A., Dzhumashev, R., Schneider, F., 2017. How does corruption affect public debt? An empirical analysis. World Dev. 90, 115–127.
Fehr, E., Fischbacher, U., 2004. Social norms and human cooperation. Trends Cogn. Sci. 8 (4), 185–190.
Fisman, R., Miguel, E., 2007. Corruption, norms, and legal enforcement: evidence from diplomatic parking tickets. J. Polit. Economy 115 (6), 1020–1048.
Frank, B., Li, S., Bühren, C., Qin, H., 2015. Group decision making in a corruption experiment: China and Germany compared. Jahrbücher für Nationalökonomie und Statistik 235 (2).
Gächter, S., Herrmann, B., 2007. The limits of self-governance when cooperators get punished: experimental evidence from urban and rural Russia. Eur. Econ. Rev. 55, 193–210.
Gächter, S., Schulz, J.F., 2016. Intrinsic honesty and the prevalence of rule violations across societies. Nature 531 (7595), 496–499.
Gatti, R., Paternostro, S., Rigolini, J., 2003. Individual Attitudes toward Corruption: Do Social Effects Matter?. The World Bank.
Gërxhani, K., Schram, A., 2006. Tax evasion and income source: a comparative experimental study. J. Econ. Psychol. 27 (3), 402–422.
Gërxhani, K., Wintrobe, R., 2021. Understanding tax evasion: combining the public choice and new institutionalist perspectives. In: Douarin, E., Havrylyshyn, O. (Eds.), Handbook of Comparative Economics. Palgrave MacMillan, pp. 785–810.
Gneezy, U., Saccardo, S., van Veldhuizen, R., 2019. Bribery: behavioral drivers of distorted decisions. J. Eur. Econ. Assoc. 17 (3), 917–946.
Guerra, A., Zhuravleva, T., 2020. Do Women Always Behave as Corruption Cleaners? Available at SSRN 3601696.
Guerra, A., Zhuravleva, T., 2021. Do bystanders react to bribery? J. Econ. Behav. Organ. 185, 442–462.
Guiso, L., Sapienza, P., Zingales, L., 2006. Does culture affect economic outcomes? J. Econ. Perspect. 20 (2), 23–48.
Gupta, S., Davoodi, H., Alonso-Terme, R., 1998. Does Corruption Affect Inequality and Poverty? IMF Working Paper WP/98/76.
Gyimah-Brempong, K., 2002. Corruption, economic growth, and income inequality in Africa. Econ. Governance 3(3), 183–209.
Harrington, J.R., Boski, P., Gelfand, M.J., 2015. Culture and national well-being: should societies emphasize freedom or constraint? PLoS ONE 10 (6), e0127173.
Hauk, E., Saez-Marti, M., 2002. On the cultural transmission of corruption. J. Econ. Theory 107 (2), 311–335.

Hoffmann, L.K., Patel, R.N., 2017. Collective Action on Corruption in Nigeria: A Social Norms Approach to Connecting Society and Institutions. Chatham House London.
IMF, 2016. Costs and Mitigating Strategies. IMF Staff Discussion Note May 2016.
Jackson, D., Köbis, N., 2018. Anti-corruption through a social norms lens. U4 Issue 7.
Keizer, K., Lindenberg, S., Steg, L., 2008. The spreading of disorder. Science 322 (5908), 1681–1685.
Köbis, N., Troost, M., Brandt, C.O., Soraperra, I., 2019. Social norms of corruption in the field: social nudges on posters can help to reduce bribery. Behav. Public Policy 1–28.
Köbis, N.C., Van Prooijen, J.-W., Righetti, F., Van Lange, P.A.M., 2015. "Who doesn't?"—the impact of descriptive norms on corruption. PLoS ONE 10 (6), e0131830.
Krupka, E.L., Weber, R., 2013. Identifying social norms using coordination games: why does dictator game sharing vary? J. Eur. Econ. Assoc. 11(3), 495–524.
Lui, F.T., 1985. An equilibrium queuing model of bribery. J. Polit. Economy 93(4), 760–781.
Mauro, P., 1997. The effects of corruption on growth, investment, and government expenditure: a cross-country analysis. Corruption Global Economy 83–107.
Moir, R.D., 1998. A monte carlo analysis of the fisher randomization technique: reviving randomization for experimental economists. Exp. Econ. 1, 87–100.
Negin, V., Abd Rashid, Z., Nikopour, H., 2010. The causal relationship between corruption and poverty: a panel data analysis.
Policardo, L., Carrera, E.J.S., 2018. Corruption causes inequality, or is it the other way around? An empirical investigation for a panel of countries. Econ. Anal. Policy 59, 92–102.
Rahwan, Z., Yoeli, E., Fasolo, B., 2019. Heterogeneity in banker culture and its influence on dishonesty. Nature 575 (7782), 345–349.
Rothstein, B., 2018. Fighting systemic corruption: the indirect strategy. J. Am. Acad. ArtsSci..
Schram, A., Brandts, J., Gërxhani, K., 2019. Social-status ranking: a hidden channel to gender inequality under competition. Exp. Econ. 22 (2), 396–418.
Schulze, G., Franck, B., 2003. Deterrence versus intrinsic motivation: experimental evidence on the determinants of corruptility. Econ. Governance (4(2)) 143–160.
Senci, C.M., Hasrun, H., Moro, R., Freidin, E., 2019. The influence of prescriptive norms and negative externalities on bribery decisions in the lab. Rationality Soc. 31 (3), 287–312.
Serra, D., 2011. Combining top-down and bottom-up accountability: evidence from a bribery experiment. J. Law Econ. Organ. (28(3)) 569–587.
Shalvi, S., Gino, F., Barkan, R., Ayal, S., 2015. Self-serving justifications: doing wrong and feeling moral. Curr. Dir. Psychol. Sci. 24 (2), 125–130.
Smith, D.J., 2008. A Culture of Corruption: Everyday Deception and Popular Discontent in Nigeria. Princeton University Press.
Snowberg, E., Yariv, L., 2021. Testing the waters: behavior across participant pools. Am. Econ. Rev. 111 (2), 687–719.
Weber, M., Schram, A., 2016. The non-equivalence of labour market taxes: areal-effort experiment. Econ. J. 127 (604), 2187–2215.
Wei, S., 2000. How taxing is corruption on international investors. Rev. Econ. Stat. (82(1)) 1–11.
Weisel, O., Shalvi, S., 2015. The collaborative roots of corruption. Proc. Natl. Acad. Sci. 112 (34), 10651–10656.
Zheng, J., Schram, A., Dögan, G., 2021. Friend or foe? Social distance in bribery. Exp. Econ 24, 854–882.
Zhu, L., 2015. Punishing corrupt officials in China. China Q. (223) 595–617.