

CULTURE AND PREVALENCE OF SANCTIONING INSTITUTIONS*

Mehmet Y. Gürdal[†] Özgür Güreker[‡] Mustafa Yahşi[§]

July 26, 2019

Abstract

The interplay of culture and institutions is of high relevance for the economic well-being of societies. Identifying causality between culture and institutions, however, is difficult using field data. Trajectories based on country-specific histories can influence the adoption and the success of institutions in different societies. To investigate the causal impact of culture on the acceptance and prevalence of sanctioning institutions, we report the results of controlled lab experiments in two countries with different cultural trajectories: Germany and Turkey. We find, if the same sanctioning institution is one of two alternatives that people can freely choose in an endogenous choice setting, then the sanctioning institution is the clear winner against a non-sanctioning institution in Germany, as well as in Turkey. Though there are some differences in people's initial institutional preferences and contribution behavior, the dynamics of institution choice, the evolution of contributions and sanctioning behavior are remarkably similar in both countries. This result is in contrast with previous studies which show that exogenous sanctioning institutions perform well in establishing cooperation in Germany, but substantially less so in Turkey.

Keywords: cooperation, culture, endogenous choice, experiments, institutions, cultural economics, trajectory.

JEL Classification Numbers: C72, C92, H41.

*We thank the participants of the 2018 Workshop on Endogenous Institutions in Copenhagen and the 2019 IMEBESS Conference in Utrecht for helpful comments. We acknowledge the financial support from the “Excellence Initiative (ZUK II)” of the German Research Foundation DFG. The research reported in this paper was approved by the Research Ethics Board of the University where the experiments were conducted.

[†]Department of Economics, Boğaziçi University; e-mail: mehmet.gurdal@boun.edu.tr

[‡]Faculty of Law, Social Sciences & Economics, University of Erfurt; e-mail: guererk@uni-erfurt.de

[§]Department of Economics, Koç University; e-mail: myahsi15@ku.edu.tr

1 Introduction

Culture is a concept relevant to many aspects of human social and economic interactions. Guiso, Sapienza, and Zingales (2006) define culture as “those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation.” While these beliefs and values are crucial for understanding human behavior and cooperation (Richerson & Boyd, 2005), the extent of the impact of culture on economic prosperity is the subject of an ongoing discussion. Some scholars highlight the predominance of institutions on economic development (Acemoglu & Robinson, 2013); others argue that culture and institutions co-evolve influencing each other (Bisin & Verdier, 2017).

With empirical (field) data, it is challenging to disentangle both the effects of institutions and the impact of culture on behaviors. There are some few notable exceptions of field studies documenting different cultural trajectories may play a role in whether a specific institution succeeds or not. Putnam, Leonardi, and Nanetti (1993), for example, report the case of governmental reform in Italy beginning in 1970 that succeeded very differently in different regions. While the adoption of the reforms concerning government performance was a success in North and Central Italian regions, similar institutional arrangements were not that successful in south Italian provinces. The authors explain the unequal impact of this institutional intervention with different cultural trajectories in Northern-Central and Southern Italy, particularly with differences in the historically routed civic engagement.

The experimental methodology can help to gain insights about relationships between culture and institutions that are difficult, if not impossible, to observe in the field. In controlled lab settings, we can implement the *same* institutional arrangement in different locations with different cultural trajectories, which enables us to study the culture’s causal influence on behavior. Most of the existing experimental studies follow this approach by keeping the institutional arrangement fixed and comparing the outcomes in different cultures.

On the other hand, not much is known about the interaction of culture and institutions when people face a *choice* between two different institutional arrangements. This ability to choose between institutions will allow individuals to obtain the first-hand experience across multiple institutions, update their choices and behavior, and potentially allow superior institutions to prevail. In this study, we investigate such a situation by utilizing a framework that gives individuals a choice between two communities with different institutional parameters while varying the actors’ cultural background. One critical element of our framework is that members of a community can monitor what happens both in their own and the other community.

We find that in Turkey, which has a substantially different cultural trajectory than Germany, a sanctioning community wins the *endogenous institutional competition* against a voluntary contribution mechanism, just as it does in Germany. Our results extend the findings of a previous study by Herrmann, Thöni, and Gächter (2008) that finds less cooperative behaviors in some cultures (among them Turkey) if the sanctioning institutions are *exogenous*. Indeed, we have some evidence that the subjects from two cultures exhibit a substantial variation during the initial stages of the experiment, yet, they manage to converge towards very similar outcomes. Our results thus show that in countries with substantially different cultural backgrounds, sanctioning institutions beneficial to societies can be adapted and prevail.

2 Related literature

Public good experiments, where subjects contribute towards a non-excludable pool, have been extensively studied. Under certain parameter restrictions, the usual design of these experiments induces an *n-person* prisoner's dilemma game where free-riding is the dominant strategy. In repeated plays, usually subjects start with positive contributions, but eventually, free-riding prevails, and contributions exhibit a continuous decline. Sanctioning institutions, however, can successfully resolve the problem of free-riding in public goods environments (Yamagishi, 1986; Ostrom, Walker, & Gardner, 1992; Fehr & Gächter, 2000). In addition to this, when given a choice, people overwhelmingly prefer sanctioning institutions over alternatives where this option is not available, and they effectively utilize sanctions to tame free-riding behavior (Güerker, Irlenbusch, & Rockenbach, 2006).

Previously conducted public good game experiments reveal striking cross-cultural differences between subject pools from countries like Germany, Switzerland, the USA, and those from Turkey, Greece, and Russia (Herrmann et al., 2008). There are also few studies that examine the cultural differences between different regions of a country.¹ Substantial differences also exist for other contexts, such as bargaining (Roth, Prasnikar, Okuno-Fujiwara, & Zamir, 1991), truth-telling (Abeler, Nosenzo, & Raymond, 2016; Gächter & Schulz, 2016), propensity to punish corrupt behavior (Cameron, Chaudhuri, Erkal, & Gangadharan, 2009), and betrayal aversion (Bohnet, Greig, Herrmann, & Zeckhauser, 2008). Besides, cross-cultural differences in social preferences (assessed via classical tools like ultimatum and dictator games) are prevalent across small scale societies (Henrich et al., 2005, 2010).

Given the stylized findings on sanctioning in public goods experiments, we ask whether *culture* can moderate the prevalence and the 'long-term' success of sanctioning institutions when individuals have a free choice to join and leave institutions as they wish. Motivated by this question, the current paper compares the findings of a public goods experiment reported in Güerker, Irlenbusch, and Rockenbach (2014), that was run in Germany, with a series of new experiments conducted among a pool of undergraduate students in Turkey.

2.1 Differences between Germany and Turkey when institutions are exogenous

We think that a comparison between Turkey and Germany provides an ideal setup for answering our main research question. As will be explained below, we were motivated by the data reported in the seminal paper by Herrmann et al. (2008). Extensive cross-country data from this study reveals that the sanctioning mechanism is highly effective in some countries, whereas it has no apparent positive effect on contributions in others. According to these results, Germany belongs to the former group and Turkey belongs to the latter.

¹Bigoni, Bortolotti, Casari, Gambetta, and Pancotto (2016) and Bigoni, Bortolotti, Casari, and Gambetta (2019) conduct a large scale lab-in-the-field experiment to find that Southern Italians cooperate less than Northerners and attribute this difference to the beliefs about others' cooperative behavior and the preferences toward social risk. On the other hand, Zhang (2018) studies the difference in the propensity to report the corruption between Northern and Southern Italians and observes that exposure to corruption enhances accountability norms merely in the presence of high-quality enforcement institutions.

Herrmann et al. (2008) data also reveal that there is a great deal of heterogeneity in terms of initial contributions. While in Copenhagen (Denmark), St. Gallen (Switzerland), and Boston (USA), the starting contributions tend to be high, in some other cities like Istanbul (Turkey), Riyadh (Saudi Arabia), and Athens (Greece), they are relatively lower. In the first three cities, without sanctioning mechanism, subjects' initial contributions amount to 65-70 percent, while in the latter three subjects contribute only 40-45 percent of their endowments initially. One observes a similar picture also with the sanctioning mechanism. Again, with 75 percent or more, Boston, Copenhagen, and St. Gallen display the highest initial contributions. In Istanbul, Athens, and Riyadh, subjects contribute only 30 percent, even less than without sanctioning mechanism. Also, concerning the average contributions across both treatments, Istanbul exhibits the 14th lowest average, and Bonn (Germany) exhibits the 7th lowest average contribution out of 16 cities. Finally, while the sanctioning mechanism leads to a significant improvement in contributions in Germany, it fails to do so in Turkey. In Germany, 9.2 out of 20 the endowed tokens are invested in the sanction-free mechanism versus 14.5 tokens in the sanctioning mechanism ($p = 0.001$), in Turkey it is 5.4 versus 7.1 ($p = 0.326$).

In a recent study, Gürdal, Torul, and Yahşi (2019) report the results of a replication of Herrmann et al. (2008) conducted at the same school where Herrmann et al. (2008) had conducted their experiment. Gürdal et al. (2019) find that contributions are significantly higher in the sanctioning institution (6.1 versus 10.9, $p = 0.019$).²

Table 1: Contributions and payoffs with exogenous institutions

	Contributions in period 1			Contributions over all periods			Payoffs	
	VCM	PUN	<i>p</i> -value	VCM	PUN	<i>p</i> -value	VCM	PUN
Istanbul - (Gürdal et al., 2019)	8.9	9.1	0.123	6.1	10.9	0.019	23.6	19.4
Istanbul - (Herrmann et al., 2008)	8.9	6.5	0.034	5.4	7.1	0.326	23.3	17.0
Bonn - (Herrmann et al., 2008)	10.9	12.1	0.012	9.2	14.5	0.001	25.5	24.1

Notes: The table reports the average contributions (out of 20 tokens) and the *p*-values for Mann-Whitney tests that use group average contributions as independent observations and test for the equality of distributions in the sanction free institutions with voluntary contribution mechanism (VCM) and sanctioning institutions with punishment possibilities (PUN).

However, as Table 1 shows, the average contribution in Gürdal et al. (2019) is still lower than those reported for Germany (Bonn) in Herrmann et al. (2008), suggesting robust differences across two cultures. Gürdal et al. (2019) point out the extensive heterogeneity across groups in the presence of the sanctioning institution for relatively lower levels of contribution. Both Gürdal et al. (2019) and Herrmann et al. (2008) state that groups with higher level of contribution in the first period have a comparably higher level of contribution in the subsequent periods.

²Herrmann et al. (2008) implement a within-subject design where sanctions treatment follows no sanctions treatment while Gürdal et al. (2019) employ a between-subject design, hence subjects either take part in the sanction-free institution or in the sanctioning institution.

3 Theoretical considerations

The prevalence of cultural differences is often attributed to cultural group selection hypothesis. According to this hypothesis, human beings are endowed with tools to acquire skills, beliefs, attitudes, and values from other individuals (Soltis, Boyd, & Richerson, 1995) and these skills enable individuals to adapt complex social norms. A high level of cultural variation among different cultural groups sets the stage for the evolution of group-beneficial behavior (Richerson et al., 2016). Henrich (2004) proposes several cultural transmission mechanisms that sustain the variation between different groups. These mechanisms facilitate cultural learning by leading individuals to copy high-frequency behaviors and imitate successful skills. Punishment of norm violators enhances this process further by increasing the fitness of prestigious and norm-abiding individuals, who in turn will be imitated at a higher frequency by the rest of the population.

One other relevant factor that facilitates the institution choice and convergence in our experiment is likely to be the common knowledge of the relative dynamic performances of sanctioning and non-sanctioning institutions. Weber and Dawes (2005) stress that it is often important what happens before decision-making as the decision problem itself and consider the ignorance of trajectories as one incompleteness of traditional economic modeling. Experimental evidence indeed shows that trajectory may affect behavior. Güreker (2013), for example, reports that if participants of an experiment are informed about the complete and detailed course of a previous run of the same experiment with different subjects, the observed initial behavior changes significantly. In an experiment that uses a similar setup as in this study, a higher fraction of the *informed* subjects join the sanctioning institution, compared to *uninformed* subjects. Moreover, informed subjects start with higher contributions to the public good.

3.1 On the process of endogenous choice: differential migration

Differential migration is one form of intergroup competition (Henrich, 2015). In particular, this peaceful form of intergroup competition suggests that people from less successful groups may migrate to more thriving societies that display higher cooperation and harmony. Boyd and Richerson (2009) provide a theoretical model on this. One can observe such movements in small-scale societies (Knauff, 1985; Tuzin, 2013), as well as in actual migration patterns from North Africa to Europe, or from Middle and South America to North America. Today, the internet and mobile devices make it easy than ever for individuals from less developed countries to compare his or her personal living conditions with other places.

In our setting, we could observe a similar differential migration, as described above. In an endogenous choice setting with two options, namely a sanctioning institution and a sanction-free institution, members of groups that start with lower average contributions in the sanctioning institution may be inclined to switch to the sanction-free institution, while individuals from groups with higher average contributions in the sanction-free institution may be inclined to switch to the sanctioning institution. The dynamics of institutional choice can make endogenous institutions perform differently than corresponding exogenous institutions. Thus, we hypothesize that an endogenous sanctioning institution in Turkey is likely to be as successful in inducing high levels of contributions to the public good as in Germany.

3.2 On the cultural differences between Germany and Turkey

There are well-documented cultural differences between Germany and Turkey. In this subsection, we will refer to Hofstede (2001) to measure cultural variations across countries, and to Falk et al. (2018) which investigates the country-level differences in terms of economic preferences.

Hofstede (2001) lists power distance, individualism, masculinity, uncertainty avoidance, long term orientation, and indulgence as the six dimensions of culture. The score for each scale ranges between 0 to 100, with 50 being a mid-level. The cultural differences between Germany and Turkey on these indexes can be seen in Figure A5 in the Appendix. We think that the reported differences between Germany and Turkey in terms of uncertainty avoidance and long term orientation might be relevant in our context. Based on the long term orientation index (Germany: 83 vs. Turkey: 46), one can expect Germans to contribute at relatively higher levels in the early periods in order to sustain high contribution levels in subsequent periods. Similarly, based on uncertainty avoidance index (Germany: 65 vs. Turkey: 85), subjects from Turkey might contribute at lower levels initially to avoid the risk brought about by others' potentially low contributions.

Falk et al. (2018) study the global variation in economic preferences and report cross-country differences for the following: time preference, risk preference, positive and negative reciprocity, altruism, and trust. Each preference measure is standardized at the individual level; hence, each preference has a mean of zero and a standard deviation of one in the individual-level world sample. The respective comparison of Germany and Turkey can be seen in Figure A6 in the Appendix.

In terms of time preferences, Germans are documented to be more patient than the Turks in this survey (0.624 vs. -0.047). In our context, patience might enhance subjects' persistence to contribute relatively higher amounts in order to sustain cooperation in their respective communities. Furthermore, we also observe that Germany scores higher in altruism (-0.051 vs. -0.279), which might also be supportive of expecting higher initial contributions in Germany.

Cross-country differences, which might have potential effects on cooperative behavior, were also reported in Herrmann et al. (2008). In the current study, we report the relevant differences between Germany and Turkey in Table A1. According to the World Value Survey, Germans trust larger a share of the population (45 percent) than Turks do (12 percent), which might be an explanation for the initial level of contributions. Norms of civic cooperation reveal no striking difference across these two countries, and hence, we expect not to observe significant variation in punishment attributes. Lastly, the difference in the rule of law index across two countries hints that Germans are more inclined to accept the norms/rules of the institution and act accordingly.

4 Experimental design and procedures

The experiment is based on the setup used in Gülerk et al. (2006, 2014), therefore we use the same terminology. In the experiment, a group of individuals simultaneously choose between two communities (institutions) before playing a public good game with each other who have chosen the same community. The two communities that subjects could choose among are termed as the VCM community and the PUN

community, respectively. While, in the VCM community, subjects merely decide on the contribution level towards the public good; in the PUN community, the players may punish other players in their community at an additional cost to themselves. The experiment has three stages; namely a voting with feet stage, a voluntary contribution stage and a punishment stage.

4.1 The voting with feet stage

In the voting with feet stage, all the members of an independent group choose one of the two communities simultaneously. Accordingly, the players decide whether to be part of the VCM or the PUN community. After all players made their decisions, each player is informed about the number of players who have chosen the same community. A player’s identity and her history of the play is concealed from the other players.

4.2 The voluntary contribution stage

In this stage, a player i is in a relation only with the players who have chosen the same community. Each player is given 20 monetary units (tokens) and may decide to contribute an integer amount of g_i , with $0 \leq g_i \leq 20$ to a joint project. The players simultaneously choose their own contribution. The amount not contributed stays in the respective player’s private account. The sum of all contributions, denoted by G , in the respective community is multiplied by 1.6, and then allocated to each player in this community, irrespective from the player’s individual contribution g_i . Consequently, the marginal per capita return depends on the community size n (see Table 2). Once all players make their contribution choices, they are informed about the individual contributions of each member in their respective community without the identity of players being revealed.

Table 2: Marginal per capita return

Community size n	2	3	4	5	6	7	8	9	10	11	12	13
MPCR	0.80	0.53	0.40	0.32	0.27	0.23	0.20	0.18	0.16	0.15	0.13	0.12

4.3 The punishment stage

In the punishment stage, played only by those choosing the PUN community, each player is endowed 20 additional monetary units independent from its own decision in the contribution stage. We provide 20 additional tokens to the players in the VCM community in order to eliminate incentives for the players to choose the punishment community just to receive extra monetary units. Since there is no punishment in the VCM community, the public good game ends after the contribution stage for those subjects choosing this community. The total monetary payoff for player i in the VCM community is given by

$$\pi_i = (20 - g_i + 1.6G/n) + 20.$$

On the other hand, in the PUN community, all players must simultaneously decide whether to punish other members of their community or not. Each player may allocate up to 20 tokens for punishment. Player i can punish player j of the PUN community by assigning punishment tokens t_{ij} . If player i assigns 1 token to player j , it will reduce the payoff of player i by 1 and the payoff of player j by 3 tokens. Let $\tau_i = \sum_{j \neq i} t_{ij}$ and $\tau_{-i} = \sum_{j \neq i} t_{ji}$. Accordingly, the total monetary payoff of player i in the punishment community is given by

$$\pi_i = (20 - g_i + 1.6G/n) + (20 - \tau_i - 3\tau_{-i})$$

The expressions in parentheses denote the stage payoffs from the contribution stage and the punishment stage, respectively. Once the punishment stage is completed, we inform all the players in the PUN community about contributions, sent punishment tokens to others, received punishment tokens and the resulting total payoffs for all players in that community. We also inform the members of both communities, the VCM and the PUN community, about the individual contributions and payoffs in both communities. With the availability of such open information flow between both communities, subjects may decide to change their community in the subsequent periods. The above sequence of interactions is repeated for consecutive 20 periods during given session.

4.4 Experimental procedures

The experiment was conducted with 85 subjects in 2016 at the Economics Laboratory of the Boğaziçi University; 6 of the sessions had each 12 subjects, the remaining one session had 13 subjects. An e-mail was sent to subjects who previously indicated an interest in joining economics experiments, and subjects could register online for a date and time they choose. Subjects were informed about the experimental procedure upon their arrival, and the instructions were read aloud. No subject participated more than once, and the sessions lasted on average 50 minutes. Subjects were paid in cash at the end of the experiment. We did not pay the subjects any show-up fee. Instead, we provide them with a starting capital of 1000 tokens to prevent bankruptcy due to excessive punishment expenditures. Random reshuffling of the presentation order on the computer screens ensured that the identity of the players could not be traced over periods.

The current experiment uses the same software code based on z-Tree (Fischbacher, 2007), and instructions as the treatment PUN reported in (Gürrer et al., 2014) that was conducted at the University of Erfurt, Germany. The experiments in Germany and Turkey share the same design and procedures except the following two differences. First, in Gürrer et al. (2014), the experiment lasts for 30 periods, where this duration was 20 periods in our case. The reason for this was that the original experiments lasted more than 2 hours. Participants are not used to such a long duration for experiments at Boğaziçi University. In addition, from previous studies we know that much of the dynamics happen in the first 10-15 periods. Thus, we decided to shorten the number of rounds to attract a sufficient number of participants. Second, in Turkey, one of the observations has 13 players. This has no effect on the constant productivity of the public good. There is a marginal decrease in MPCR from 0.13 to 0.12, if a community reaches the maximum capacity with 13 individuals instead of 12.

5 Results

We start by looking at the overall picture before going more in detail. First we report some descriptive statistics. Table 3 summarizes the mean values and the standard errors of contribution decisions and realized payoffs, separately for the VCM and the punishment PUN communities in Turkey (present study), and those conducted in Germany by Güreker et al. (2014). Compared to PUN, the average contribution levels are quite low under VCM, and the standard errors of the contributions are arguably lower as well. In the PUN community, the mean contribution start at levels slightly above 10 and keep increasing for the subsequent periods. In the VCM community, we observe the opposite as the contributions approach to zero towards the end of the experiment. The average payoff attains values close to 40 and the variations is highly limited for the VCM community. This situation can be explained by the stable and low levels of average contributions. We also observe that the standard error of payoff is almost negligible. For the PUN community, compared to VCM, the payoff takes low values in the beginning of the experiment, but keeps increasing for the subsequent periods.

Table 3: Contributions and payoffs over time

	Average Contribution				Average Payoff			
	VCM		PUN		VCM		PUN	
	TUR	GER	TUR	GER	TUR	GER	TUR	GER
Period 1	3.14 (0.72)	7.39 (0.69)	10.05 (1.38)	13.07 (0.77)	41.88 (0.43)	44.44 (0.41)	21.25 (3.28)	23.04 (6.45)
Period 1-10	2.46 (0.39)	3.55 (0.59)	16.88 (0.85)	16.53 (0.85)	41.48 (0.23)	42.13 (0.36)	39.57 (2.50)	35.58 (2.80)
Period 11-20	1.21 (0.83)	2.07 (0.55)	19.54 (0.24)	19.19 (0.40)	40.73 (0.50)	41.24 (0.33)	49.14 (0.86)	47.95 (1.33)
Period 21-30	- -	1.06 (0.77)	- -	19.67 (0.11)	- -	39.63 (1.19)	- -	48.65 (0.93)

TUR and GER stand for the subjects in Turkey and Germany, respectively.

The unit observation is a subject's own contribution to the public good for the mean statistic.

Standard errors are clustered across groups.

In the VCM community, the contribution over all periods to the public good is on average 2.23 for Turkey, while it is on average 2.92 for Germany ($p = 0.203$). In the PUN community, the contribution over all periods is on average 18.48 for Turkey and 18.83 for Germany ($p = 0.643$).

A comparison of the payoff levels reveals no significant difference across countries for both communities. In particular, in the VCM community, the payoff over all periods is on average 41.34 for Turkey, whereas it is 41.75 for Germany ($p = 0.203$). In the PUN community, the average payoff over all periods is 45.33 for Turkey and 45.64 for Germany ($p = 1.000$).

Figure 1 shows the evolution of contributions in Turkey and Germany for each period. In both samples,

contributions remain much lower under VCM community compared to the PUN community. Number of subjects in the PUN community gradually increase in both samples, with almost nobody remaining in the VCM community towards the end of the game. For early periods, the percentage of subjects in the PUN community is generally higher in Turkey compared to Germany.

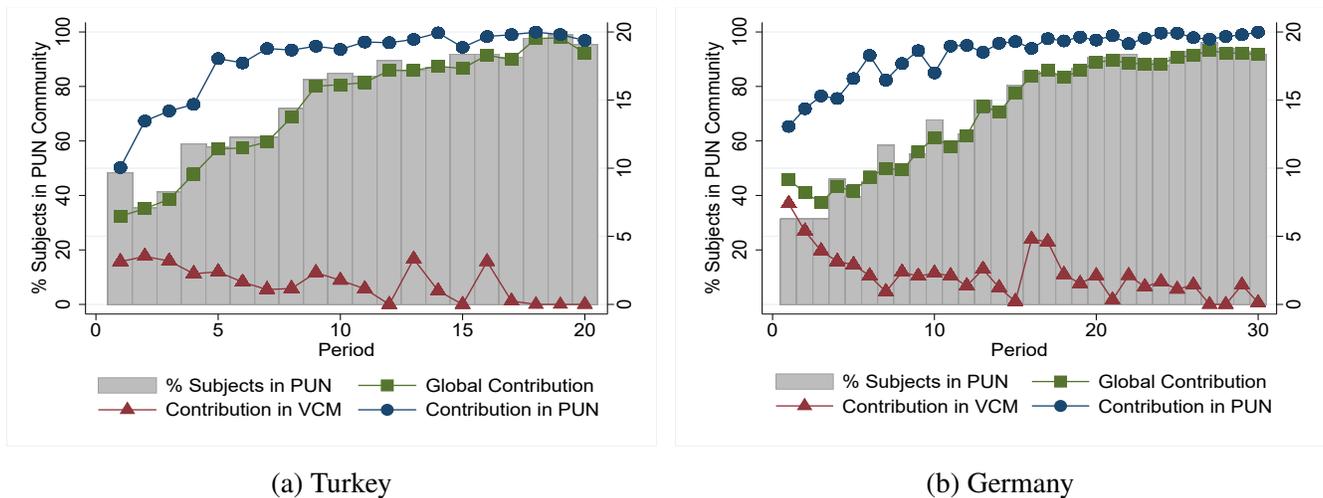


Figure 1: Evolution of Contributions

5.1 First period behavior

In this section, we compare the first period behavior of subjects from Turkey with those from Germany. For this comparison, the data indicate significant differences in contributions, however, not in the initial institutional taste: in the VCM community, the number of subjects is on average 2.95 (24%) for Turkey, whereas it is 3.41 (28%) for Germany ($p = 0.418$).³ Additionally, in the PUN community, the number of subjects is on average 9.19 (76%) for Turkey, whereas in the PUN community in Germany, the number of subjects is on average 8.59 (72%) for Germany ($p = 0.298$). In terms of the initial popularity of the institutions, our data do not indicate a significant difference.

In the VCM community, the first period contribution to the public good is on average 3.14 for Turkey, whereas this value is on average 7.39 for Germany ($p = 0.003$). In the PUN community, the contribution in the first period is on average 10.05 for Turkey and 13.07 for Germany ($p = 0.073$). To summarize, subjects in Turkey start with lower contributions in both communities.

A comparison of the payoff levels for two samples point out a difference in the VCM communities but not in the PUN communities. In particular, in the VCM community, the first period payoff is on average 41.88 for Turkey, whereas it is 44.44 for Germany ($p = 0.003$). This is not surprising given the higher contributions in Germany for the initial periods. On the other hand, in the PUN community, the first period payoff is on average 21.25 for Turkey and 23.04 for Germany ($p = 0.565$). In PUN

³All non-parametric statistical tests reported in this paper are two tailed and take total population, or the respective communities as units of observations. While comparisons within a treatment are tested with the Wilcoxon matched-pairs tests, comparisons across treatments are done with the Mann-Whitney U tests.

community, while subjects assign 0.98 sanction points on average to other subjects in Turkey, it is 1.52 in Germany ($p = 0.949$).

Result 1 *Initial preferences for both communities are similar in Turkey and Germany. In the VCM community, contributions in Germany are considerably higher, which leads to higher payoffs. In the PUN communities, contributions in Germany are only weakly significantly higher than contributions in Turkey, while neither average punishment nor payoffs differ significantly between Germany and Turkey.*

5.2 Mean punishment expenditures

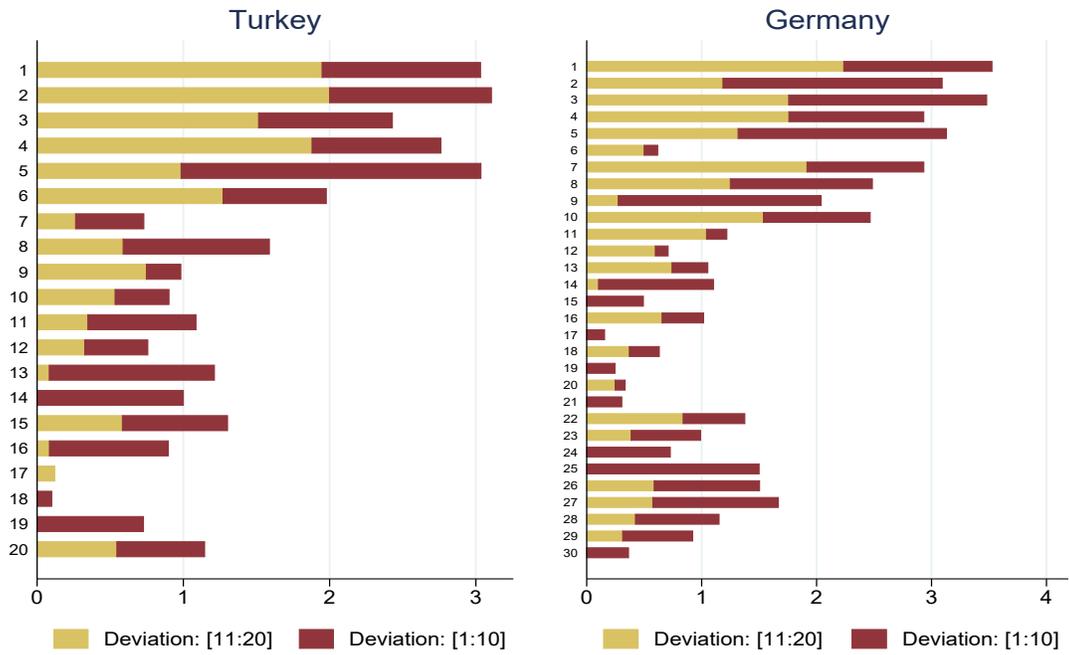
Figure 2 shows the mean punishment expenditures in Turkey and Germany per period. To begin with, the upper portion of Figure 2 indicates the average sanctioning points assigned to the subjects who has contributed lower number of tokens to the public good compared to the punisher himself/herself (free-rider punishment). For this case, per period averages are between 0.01 and 1.47 in Turkey while they are between 0.04 and 1.75 in Germany. Free-rider punishment expenditure is on average 1.61 for Turkey, and it is 1.54 for Germany ($p = 0.908$). Disregarding the cases for punishment expenditure was zero, the free-rider punishment expenditures is on average 2.82 for Turkey and 3.32 for Germany ($p = 0.418$). Free-rider punishment is strictly greater than 0 for 57.24% of all possible cases in Turkey; whereas it is for 46.46% of all possible cases in Germany. We do not find any difference between the two samples in terms of the free-rider punishment frequency ($p = 0.355$).

The lower portion of Figure 2 shows the mean punishment expenditures that fall into the category of anti-social punishment. To put it simply, anti-social punishment is observed when a subject assigns sanctioning points to another subject who has contributed at least as much as the punishing subject. We notice that the mean anti-social punishment expenditure varies across periods and takes values between 0.01 and 0.44 in Turkey and values between 0.00 and 0.62 in Germany. Interestingly, in both countries, most of the anti-social punishment executions are directed towards high contributors, as indicated by the green columns in the lower portion of Figure 2 that occupy the most space. The overall anti-social punishment expenditures is on average 0.035 for Turkey and 0.037 for Germany ($p = 0.817$). Considering only the anti-social punishment expenditures that are strictly greater than zero, the average is 1.57 for Turkey, while it is 1.83 for Germany ($p = 0.037$). The frequency of anti-social punishment does not differ between both countries. Anti-social punishment is executed in 2.24% of all possible cases in Turkey; whereas it is for 2.05% of all possible cases in Germany ($p = 0.355$).

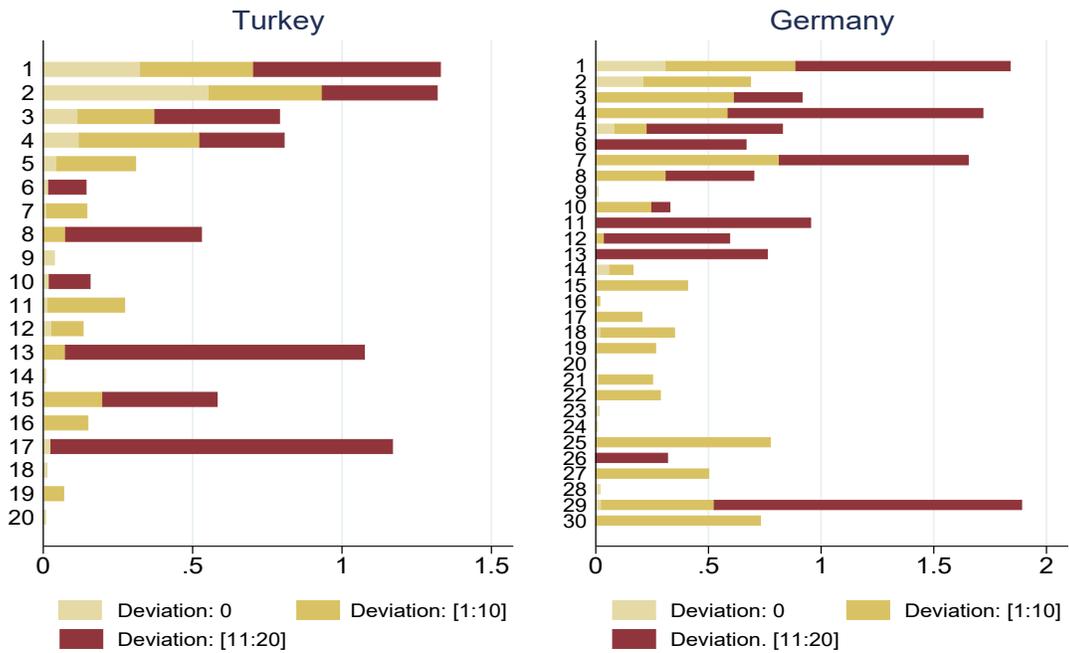
Result 2 *In Turkey and Germany, free-riders are punished similarly in terms of punishment frequency and severity. The frequency of anti-social punishment is similar in Turkey and Germany, while the severity of anti-social punishment executions is significantly higher in Germany.*

5.3 Contribution levels by different types of subjects

We define types of subjects according to the level of contribution in the first period. We label subjects who contribute at least 15 tokens to the public good but do not punish other subjects as cooperators. We



(a) Free-rider punishment



(b) Anti-social punishment

Figure 2: Sanctioning Expenditures

label participants who are cooperators and also punish others as strong cooperators. Finally, we say that subjects who contribute 5 or less tokens are free-riders. Accordingly, Table 4 indicates the mean level of

contribution made by cooperators, strong cooperators and free-riders respectively.

Table 4: Contribution levels by different types of subjects

Statistic	Free-rider		Cooperator		Strong cooperator	
	TUR	GER	TUR	GER	TUR	GER
Mean	14.52	11.53	-	19.17	18.95	18.44
SE	(0.63)	(1.48)	(-)	(0.24)	(0.16)	(0.53)

TUR stands for the subjects in the respective category (e.g. free riders in Turkey)

GER stands for the subjects in such category (e.g. cooperators in Germany)

Averaged over all periods, free-riders contribute 14.52 in Turkey, while they contribute 11.53 in Germany ($p = 0.197$). Moreover, the contribution by strong cooperators is on average 18.95 for Turkey, while it is on average 18.44 for Germany ($p = 0.886$).⁴

Result 3 *Average contributions of the subjects who were categorized as free-riders or strong cooperators do not differ between Turkey and Germany.*

5.4 Individual change/migration behavior

A closer examination of the average number of switches per subject from one community to another, from PUN to VCM, and vice versa, reveals that German subjects on average switch their communities more often than the Turkish subjects do. The average number of switches per subject is 2.64 on average in Turkey, while, in Germany, it is 4.21 ($p = 0.011$).⁵

The average number of switches per free-rider is 3.38 on average in Turkey, while, in Germany, this value is 9.17 ($p = 0.092$).⁶ Also, the average number of switches per strong cooperator is 1.71 on average in Turkey, while, in Germany, this value is 2.57 ($p = 0.943$).⁷

Result 4 *Overall, individuals in Turkey change the communities significantly less often than subjects in Germany. If we look at types, we observe that free-riders in Germany change more often than Turkish free-riders.*

⁴There were not any subjects who contribute at least 15 tokens to the public good but do not punish other subjects in the first period for Turkey; therefore, the comparison of contribution levels made by cooperators across Turkey and Germany is not available. Note that the contribution by cooperators is on average 19.17 for Germany.

⁵Note that subjects play the public good game for 30 periods in Germany while they play the same public good game in Turkey for only 20 periods. If we consider only the first 20 period for Germany, the number of switches becomes 3.61 on average for this sample, still indicating a significant difference in between two samples ($p = 0.032$).

⁶If we consider only the first 20 period for Germany, the number of switches per free-rider becomes 7.50 on average for this sample, still indicating a significant difference between two samples ($p = 0.092$).

⁷If we consider only the first 20 period for Germany, the number of switches per strong cooperator becomes 2.29 on average for this sample ($p = 0.943$).

5.5 Community change behavior and next period contributions

We examine the relationship between the received sanction points and the switching behavior from PUN to VCM community, and the contribution adjustments when subjects stay in the punishment community for both countries. First two columns (models) of Table 5 show probit models to identify the one to one relation between received sanction points and the movement from PUN to VCM community for Turkey and Germany. The third and the fourth columns demonstrate the determinants of contribution adjustments for the subsequent periods when subjects remain in PUN, for Turkey and Germany respectively. In the fifth and sixth columns, we only consider the contributions for periods that subjects switch from PUN to the VCM community.

The dependent variable `Switched from PUN to VCM` is a binary variable that takes the value of 1 if a subject switches from PUN community to VCM community for the next period. The independent variable affecting the community movement is the received sanction points. The first two columns of Table 5 show the marginal effects for the percentage point changes when the respective independent variable increases by one unit. Accordingly, the first column of Table 5 is the switch from the punishment community to VCM community in Turkey while the second column stands for the switch from PUN to VCM in Germany.

Table 5: Community change behavior and next period contributions

	(1)	(2)	(3)	(4)	(5)	(6)
	Switched from PUN to VCM		Next cont when staying in PUN		Cont after switch from PUN to VCM	
	TUR	GER	TUR	GER	TUR	GER
Sanction Points	0.099*** (0.013)	0.097*** (0.010)	0.264*** (0.029)	0.037** (0.018)	-0.067 (0.115)	-0.099 (0.081)
Contribution			0.693*** (0.030)	0.426*** (0.022)	0.012 (0.134)	-0.131 (0.116)
Constant	-1.832*** (0.066)	-1.767*** (0.071)	5.849*** (0.592)	11.182*** (0.451)	5.295** (2.367)	7.283*** (2.227)
N	1615	2784	1114	1801	92	173
r ²						

(1) and (2): Probit Regressions, (3) – (6): GLS random effects models, standard errors in parantheses

* $p < .1$, ** $p < .05$, *** $p < .01$

The received sanction points increase the likelihood of switching from the PUN community to the VCM community in both countries. Specifically, one point increase in the received sanction points in the punishment community leads to 9.9 percent increase in the likelihood of switching from the PUN community to the VCM community in Turkey while it leads to 9.7 percent rise in the probability of switching in Germany. We conclude that the received sanction points has a similar effect on the switching behavior from PUN to VCM community in both countries.

In the third and fourth columns, we analyze the contribution adjustments for the subjects who were in the PUN community at time $t - 1$ and remain in PUN at time t . We use subject's own contribution in addition to the independent variable in the first two columns to explain the adjustments to the contribution

for the next period. Both independent variables have a significant impact on the next round's contributions for the subjects that remain in the punishment community. Accordingly, an increase on any of these two variables leads to an increase in the next period's contribution for the subjects that choose to stay in the PUN community (0.264 and 0.037 points in Turkey and Germany, respectively, $p = 0.121$). Apparently, subjects, who do not leave the PUN community, increase their contributions to avoid future sanctions. Furthermore, for subjects staying in the PUN community, one point increase in the subjects' own contribution results in an average increase of 0.693 in the next period's contribution in Turkey, and 0.426 in Germany ($p = 0.366$).

In the fifth and sixth columns, for subjects who switch to VCM community, we observe that neither contribution level nor received sanction points are significant to affect the next period contributions.

Result 5 *Received sanctions have a similar impact on subjects to leave PUN for VCM in Turkey and Germany. For the switching subjects to VCM, sanctions do not affect their next period contributions, neither in Germany nor in Turkey. For subjects who stay in PUN after being sanctioned, sanctions have a significant effect on their next round contributions, with subjects in Turkey showing a stronger reaction than Germans, though the difference is not significant.*

5.6 Community choice and the norms of civic cooperation

In this section, we investigate the relationship between the community choices and the norms of civic cooperation. To that end, we ask subjects in Turkey three different questions to measure their score on civic norms.⁸ We set score of 1 as the lowest level of cooperation and the score of 10 as the highest level of cooperation (free-rider) on each question. Then, we sum the scores on these three questions to obtain a total score of the norms of civic cooperation. We observe that the average score of civic norms do weakly significantly differ between subjects who choose the VCM community and those who join the PUN community in the first period (25.57 vs. 24.21, $p = 0.052$).

The relationship between the first round contributions and the civic norm score can be found below. There is a positive relationship between civic norm score and first round contributions in the PUN community. This means subjects valuing civic norms higher contribute more. This relation is not observed in the VCM community.

Result 6 *In Turkey, subjects who choose the PUN community initially value civic norms of cooperation somewhat higher. For those choosing the PUN community in the first period, we observe a positive relationship between the valuation of civic norms and first period contributions.*

5.7 Evolution of payoffs by communities

Figure 3 shows the evolution of payoffs by communities in Turkey and Germany, respectively. In the VCM community, the payoffs vary less across periods for Turkey than for Germany. Furthermore, we observe

⁸In the original experiments conducted in Germany, this questionnaire was not asked, so we can analyze only the subjects' behavior in Turkey.

Table 6: Civic norms

	(1)	(2)	(3)
	Contribution	Contribution VCM	Contribution PUN
totalcivic	0.183 (0.119)	-0.027 (0.106)	0.325* (0.167)
Constant	2.040*** (2.701)	3.863** (2.349)	2.987*** (4.102)
N	72	42	30
r2	0.025	0.001	0.077

Standard errors in parentheses

OLS Regressions, standard error in parantheses

* $p < .1$, ** $p < .05$, *** $p < .01$

that the payoffs are more volatile in the PUN community in the first periods for Germany. Overall, the evolution of payoffs in both communities follow a similar pattern for Turkey and Germany.

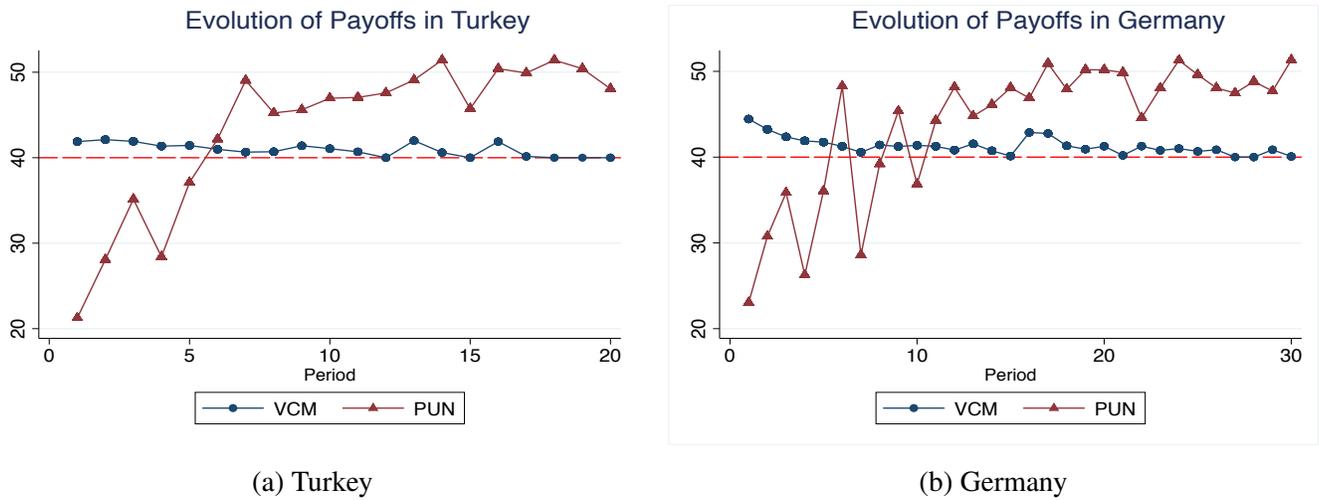


Figure 3: Evolution of payoffs by communities

In Turkey, in the VCM community, the payoff is 41.34 on average while, in the PUN community, it is 45.33 ($p = 0.063$). Considering the first 10 periods in Turkey, the payoffs in the VCM community do not significantly differ from the payoffs in the PUN community (41.48 versus 39.57, $p = 0.311$); however, the payoffs between 11th and 20th period in the VCM community are significantly lower than the payoffs between 11th and 20th period in the PUN community (40.73 versus 49.14, $p = 0.018$).

In Germany, in the VCM community, the payoff is 41.75 on average while, in the PUN community, it is 45.64 ($p = 0.036$). Considering the first 10 periods in Germany, different than what we observe in Turkey, the payoffs in PUN are significantly lower than the payoffs in VCM (35.63 and 42.13, $p = 0.036$). Between 11th and 20th periods, the payoffs in PUN, however, catch up the payoffs in VCM (47.95 and

41.24, $p = 0.017$). Finally, the payoffs between 21th and 30th period reveal that the average profit is significantly higher in PUN community (48.65 and 40.63, $p = 0.028$).

Over all periods, average payoffs in Turkey and Germany are very similar. In the VCM community, average payoff is 41.34 for Turkey while it is 41.75 for Germany ($p = 0.203$). In the PUN community, average payoff is 45.33 for Turkey while it is 45.64 for Germany ($p = 1.000$).

Result 7 *Averaged over all periods, payoffs in Turkey and Germany are very similar, both in the VCM and PUN communities. “in the long run”, i.e., in the last 10 periods, the sanctioning community PUN induces significantly higher payoffs than in the VCM community, both, in Turkey and Germany.*

6 Concluding remarks

Summing up our main findings, in the initial phase, we observe a similar reluctance in both cultures to join the sanctioning institution, with roughly one in four individuals choosing it. The initial contributions, however, differ between Turkey and Germany. Initially, subjects in Turkey contribute lower amounts in both institutions. This observation is in line with the findings of Herrmann et al. (2008), in the case of the exogenous sanction-free (VCM) and sanctioning (PUN) institutions.

Looking at sanctioning behavior, in contrast to Herrmann et al. (2008), we observe similarities in Turkey and Germany. In both cultures, people discipline free-riders similarly, in terms of *frequency* of punishment, and its *severity* when executed. Interestingly, not the frequency but the severity of anti-social punishment is higher in Germany.

Regarding payoffs, participants in Turkey and Germany obtain very similar earnings, both in the VCM and the PUN institution. In the “long-run”, the sanctioning institutions are the clear winner of the institutional competition in both cultures, both in terms of contributions and payoffs. This result extends the findings of Gächter, Renner, and Sefton (2008) who show the long-run advantage of exogenous sanctioning institutions over the VCM.

We believe that experimental methodology has a remarkable potential to contribute to the cross-cultural comparison of the success and the prevalence of different institutions. This is mainly due to the ability to exert control over potential confounds and the ability to fine-tune the degree of cultural variation via access to various combinations of subject pools. Previous studies mainly employ a design where the same institutional arrangement is tested across different subjects pools, and the results are taken as evidence of the effect culture on social behaviors such as altruism, cooperation, and truth-telling. We think that a more accurate picture regarding the interaction of culture and institutions can be obtained by employing experimental designs where (i) knowledge about the past performance of institutions, (ii) endogenous choice between institutions, or both of these are available so that subjects engage in a better-informed decision-making process. In the experiment reported here and in Gülerk et al. (2014), endogenous choice of institutions is the critical factor that allows the eventual convergence of subjects from two distinct cultures.

Placing our results in a broader context, with this study, we demonstrate that previously observed uncooperative behaviors in some cultures may be the result of a specific trajectory experienced in those

cultures. These (uncooperative) trajectories may be overcome if the right institutions can be chosen by the members of the society. This is in the spirit of the argument made by Acemoglu and Robinson (2013) in the context of the prosperity of nations and institutions: it is not geography or culture which is decisive for the prosperity of a nation but the availability (and the quality of) its institutions.

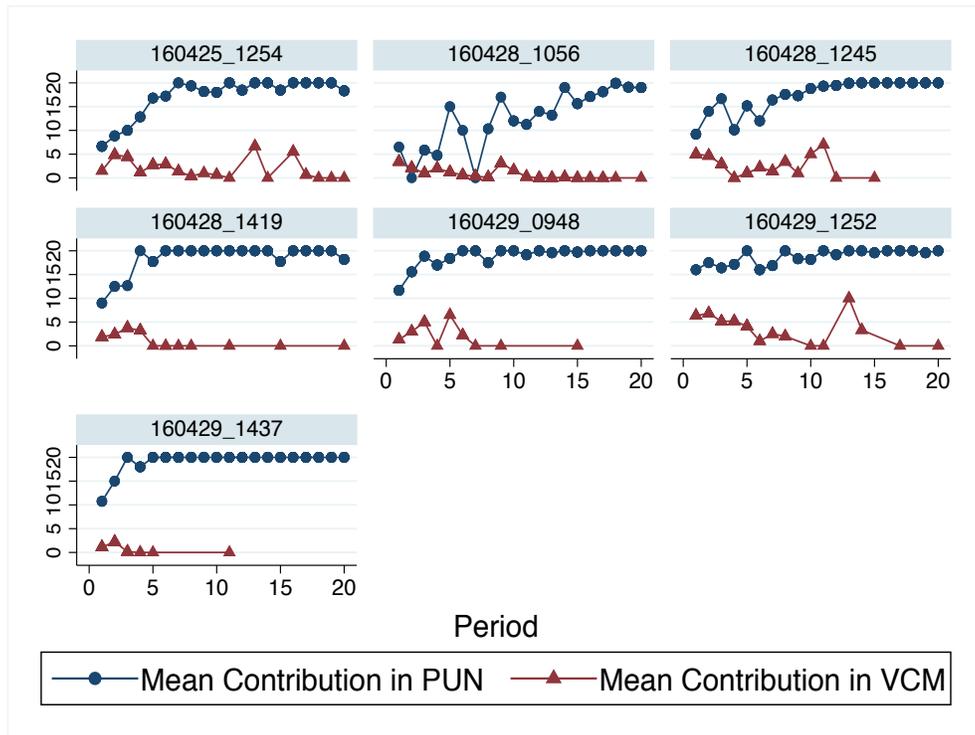
References

- Abeler, J., Nosenzo, D., & Raymond, C. (2016). *Preferences for truth-telling* (IZA Discussion Papers No. 10188). Institute of Labor Economics (IZA).
- Acemoglu, D., & Robinson, J. A. (2013). *Why nations fail: The origins of power, prosperity, and poverty*. Broadway Business.
- Bigoni, M., Bortolotti, S., Casari, M., & Gambetta, D. (2019). At the Root of the North–South Cooperation Gap in Italy: Preferences or Beliefs? *Economic Journal*, 129(619), 1139–1152.
- 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. *Economic Journal*, 126(594), 1318–1341.
- Bisin, A., & Verdier, T. (2017). *On the joint evolution of culture and institutions* (Tech. Rep.). National Bureau of Economic Research.
- Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *American Economic Review*, 98(1), 294–310.
- Boyd, R., & Richerson, P. J. (2009). Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior. *Journal of Theoretical Biology*, 257(2), 331–339.
- Cameron, L., Chaudhuri, A., Erkal, N., & Gangadharan, L. (2009). Propensities to engage in and punish corrupt behavior: Experimental evidence from australia, india, indonesia and singapore. *Journal of Public Economics*, 93(7–8), 843–851.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018, 05). Global Evidence on Economic Preferences*. *The Quarterly Journal of Economics*, 133(4), 1645–1692.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4), 980–994.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.
- Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *Science*, 322(5907), 1510–1510.
- Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531(7595), 496.
- Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2), 23–48.
- Gürdal, M. Y., Torul, O., & Yahşi, M. (2019). *Heterogeneity in sanctioning behavior and public goods*.
- Gürerk, Ö. (2013). Social learning increases the acceptance and the efficiency of punishment institutions in social dilemmas. *Journal of Economic Psychology*, 34, 229–239.
- Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, 312(5770), 108–111.
- Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2014). On cooperation in open communities. *Journal of Public Economics*, 120, 220–230.

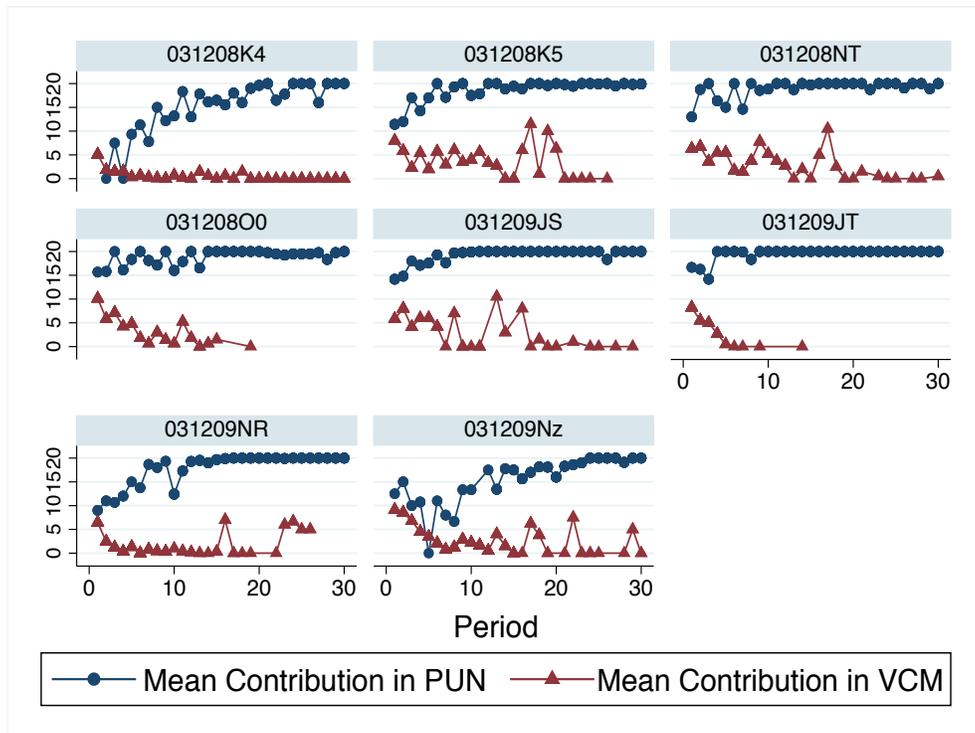
- Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, 53(1), 3–35.
- Henrich, J. (2015). *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., . . . others (2005). “economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6), 795–815.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., . . . others (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327(5972), 1480–1484.
- Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868), 1362–1367.
- Hofstede, G. (2001). *Culture’s consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage Publications.
- Knauff, B. M. (1985). *Good company and violence: Sorcery and social action in a lowland new guinea society* (No. 3). Univ of California Press.
- Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, 86(2), 404–417.
- Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1993). *Making democracy work: Civic institutions in modern italy*. Princeton University Press, Princeton.
- Richerson, P. J., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., . . . others (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, 39.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone : How culture transformed human evolution*. University of Chicago Press.
- Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market behavior in jerusalem, ljubljana, pittsburgh, and tokyo: An experimental study. *The American Economic Review*, 1068–1095.
- Soltis, J., Boyd, R., & Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural group selection?: An empirical test. *Current Anthropology*, 36(3), 473–494.
- Tuzin, D. (2013). *Social complexity in the making: A case study among the arapesh of new guinea*. Routledge.
- Weber, R., & Dawes, R. (2005). Behavioral economics. *The handbook of economic sociology*, 90–108.
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51(1), 110.
- Zhang, N. (2018). Institutions, norms, and accountability: A corruption experiment with northern and southern italians. *Journal of Experimental Political Science*, 5(1), 11–25.

A Appendix

Figure A1 illustrate the evolution of contributions per session in Turkey and Germany respectively.



(a) Groups in Turkey



(b) Groups in Germany

Figure A1: Evolution of contributions

Figure A2 shows the average punishment tokens for each period in Turkey and Germany. The statistic denotes the mean punishment token with respect to all possible sanctioning decisions.

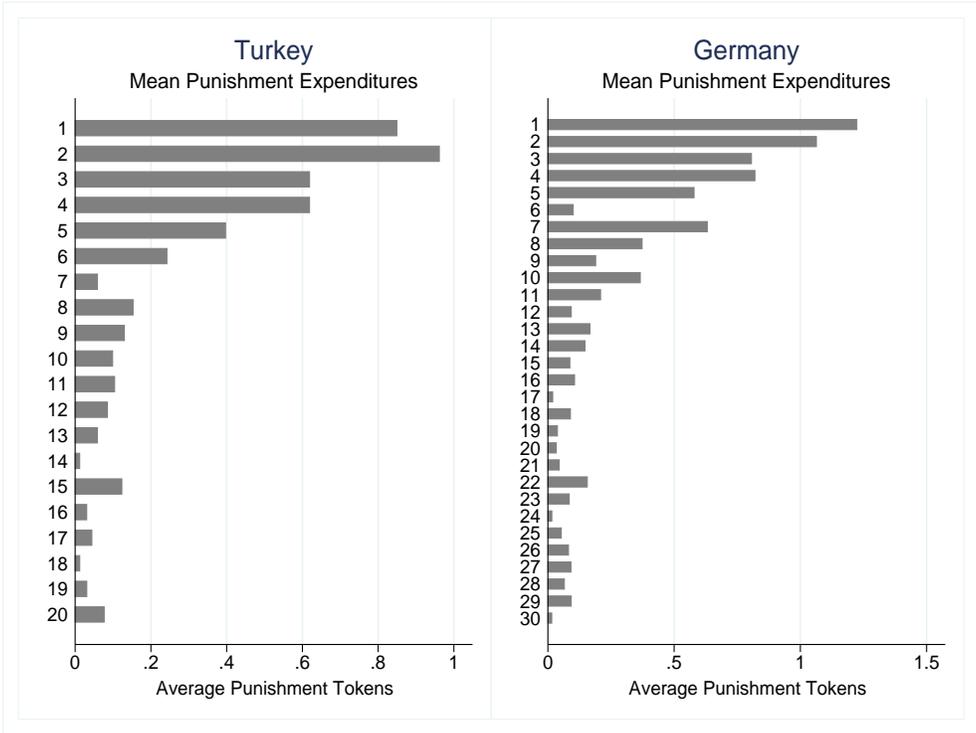


Figure A2: Average punishment tokens

Figure A3 shows the average frequency of anti-social punishment for each period in Turkey and Germany.

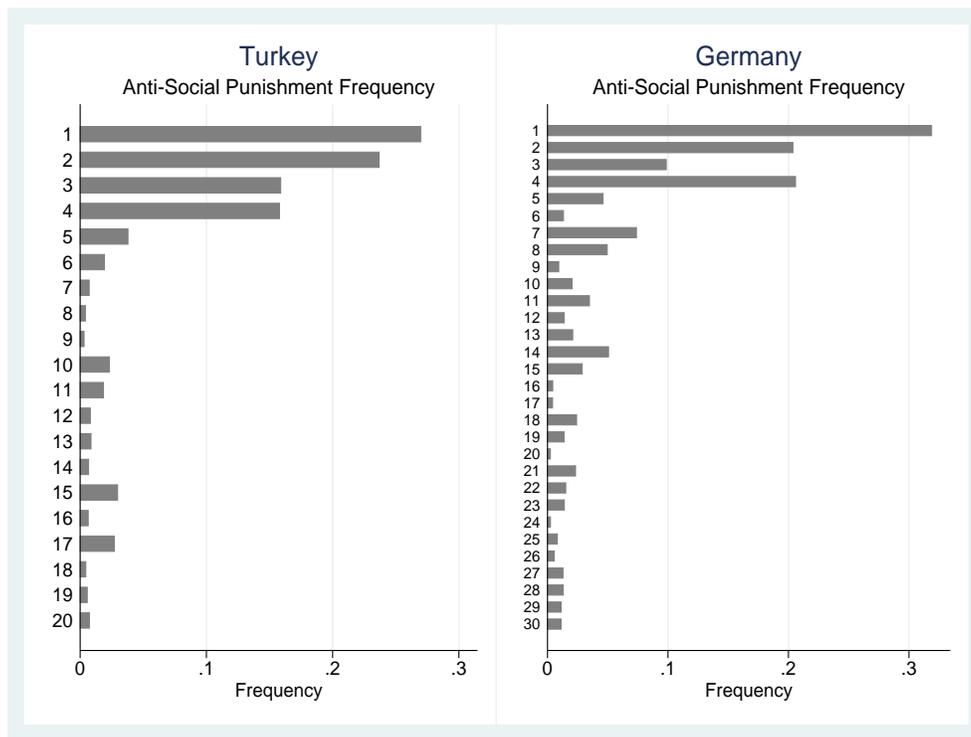


Figure A3: Mean frequency of anti-social punishment

Figure A4 shows the average anti-social punishment expenditures for each period in Turkey and Germany. The statistic denotes the mean value with respect to all possible anti social punishment decisions (e.g. anti-social punishment expenditure can be 0).

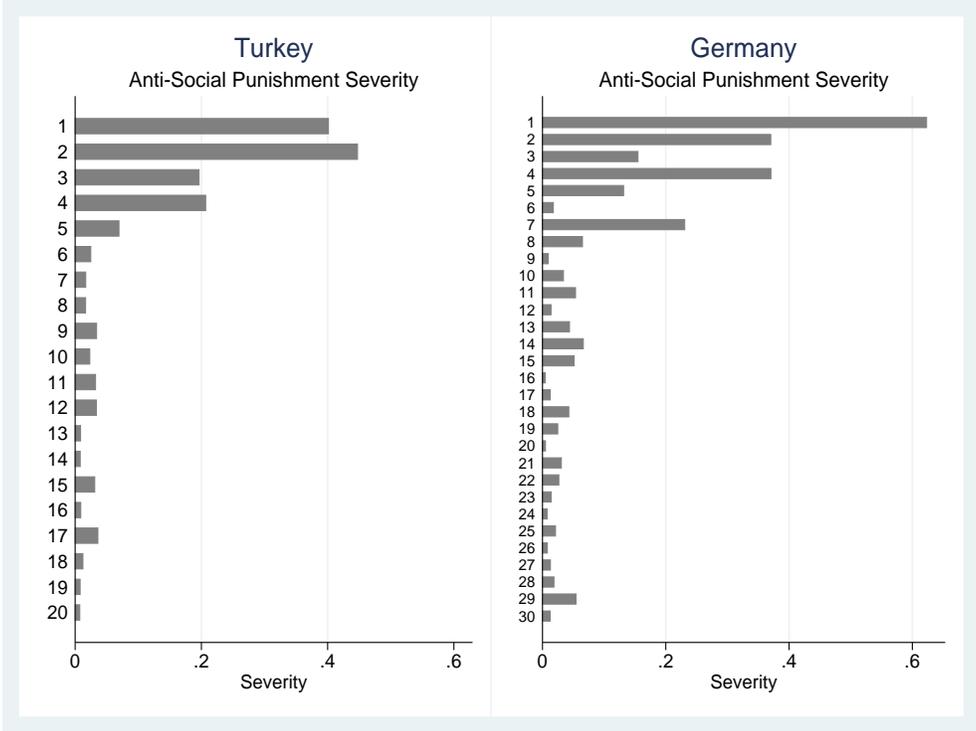


Figure A4: Average anti-social punishment

Figure A5 reflects Hofstede (2001)'s six cultural dimensions for Turkey and Germany.

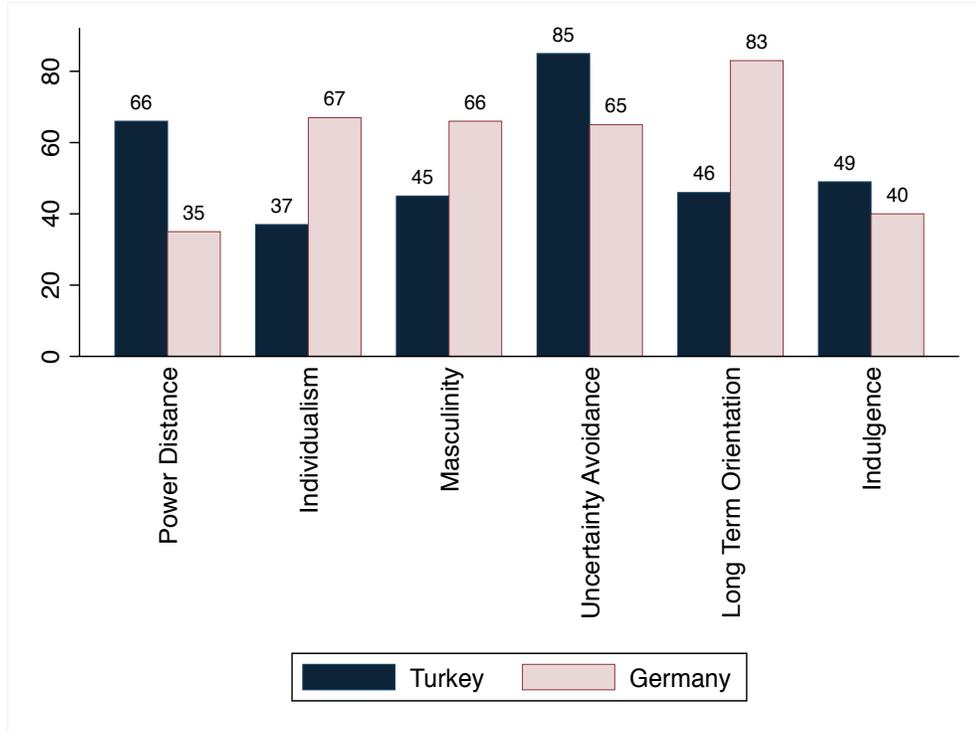


Figure A5: Hofstede's Cultural Dimensions for Turkey and Germany

Figure A6 shows the difference to the world mean in standard deviation of the six economic preferences for Germany and Turkey, retrieved from Falk et al. (2018).

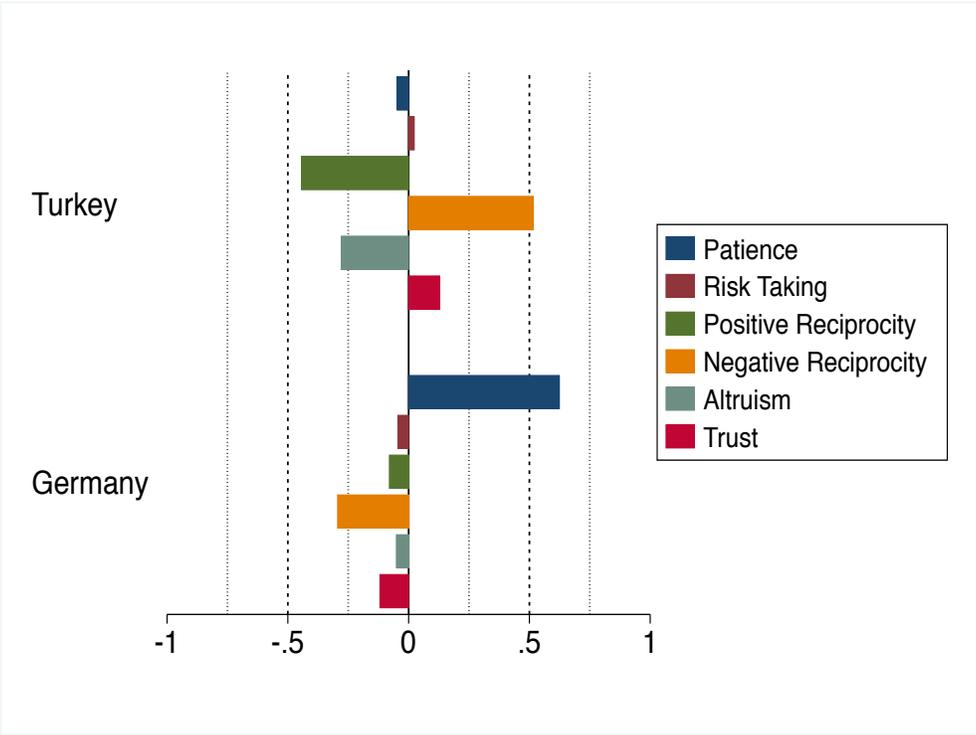


Figure A6: Global Preferences Survey for Germany and Turkey

Table A1 demonstrates the main cultural and societal background of the societies of our subject pools.

Table A1: Cultural and societal background of subject pools

Social capital variables		Norms of civic cooperation					Secularity &Autonomy		Law enforcement & Democracy		
Subject Pool	Country	Share of people who can be trusted	Avoiding fare on transportation	Claiming government benefits	Stealing Property	Cheating on taxes	Accepting bribe	Overall Secular Values	Autonomy Index	Rule of law	Importance of democracy
Erfurt	Germany	0.45	9.08	9.06	9.68	-	9.31	0.41	0.54	1.68	8.94
Istanbul	Turkey	0.12	9.16	9.33	9.74	9.73	9.80	0.24	0.00	-0.03	8.57

Rule of law data is retrieved from World Bank between the years of 2010 and 2017 while the rest of data are acquired from World Value Survey Wave 6

Norms of civic cooperation score ranges between 1 and 10 (1=very weak norms of civic cooperation; 10=very strong norms of civic cooperation)

Secular value lies between 0 and 1 (1=very secular)

Autonomy index ranges between -1 to 1 (-1=weakest autonomy,1=highest autonomy)

Rule of law ranges approximately between -2.5 and 2.5 (2.5=very strong rule of law)

Importance of democracy ranges between 1 and 10 (1=not important at all, 10= absolutely important)