

Voigt, Stefan; von Jacobi, Nadia

**Working Paper**

## Mapping Informal Institutions – a Global Dataset

ILE Working Paper Series, No. 89

**Provided in Cooperation with:**

University of Hamburg, Institute of Law and Economics (ILE)

*Suggested Citation:* Voigt, Stefan; von Jacobi, Nadia (2025) : Mapping Informal Institutions – a Global Dataset, ILE Working Paper Series, No. 89, University of Hamburg, Institute of Law and Economics (ILE), Hamburg

This Version is available at:

<https://hdl.handle.net/10419/331244>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT  
FÜR RECHTSWISSENSCHAFT

INSTITUTE OF LAW AND ECONOMICS  
WORKING PAPER SERIES

# Mapping Informal Institutions – a Global Dataset

Stefan Voigt  
Nadia von Jacobi

Working Paper 2025 No. 89

November 2025



*Photo by UHH/RRZ/Mentz*

NOTE: ILE working papers are circulated for discussion and comment purposes.  
They have not been peer-reviewed.

© 2025 by the authors. All rights reserved.

# Mapping Informal Institutions – a Global Dataset

STEFAN VOIGT\* and NADIA VON JACOBI†

## *Abstract*

*The crucial importance of institutions for economic and social development has become part of the conventional wisdom in economics. In recent years, economists have become particularly interested in the potential role of informal institutions but measuring them remains a serious challenge. This paper collects more than three dozen variables depicting various informal institutions, meaning institutions whose non-compliance is sanctioned by members of society, in up to 180 countries. It draws on individual and collective measures elicited through different approaches such as surveys and lab experiments, but also family types and language traits. The resulting dataset can be used in many ways. We apply correlation network analysis to showcase which variables are more adequate for which type of empirical application.*

*Keywords: informal institutions, internal institutions, culture, conventions, social norms, customs, institutional economics, correlation network analysis, system.*

*JEL classification: A13, D90, K00, O10, Z10*

---

\* Institute of Law & Economics, University of Hamburg and CESifo, Munich, Germany; Email: [Stefan.Voigt@uni-hamburg.de](mailto:Stefan.Voigt@uni-hamburg.de)

† Department of Economics and Management, University of Trento, Italy; Email: [nadia.vonjacobi@unitn.it](mailto:nadia.vonjacobi@unitn.it).  
We thank Tim Schnelle for outstanding research assistance. Special thanks also go to Rafal Ramadan for truly excellent research support. We thank Abishek Choutagunta, Marco Faillo, Veronika Fikfak, Jerg Gutmann, Nourhan Hassan, Monia Idoudi, Dylan Johnson, Mahdi Khesali, Raphael Maesschalck, Hashem Nabas Roei Sarel, Betül Simsek, Nancy Wu, and Emmi Zaepernick as well as participants of the WINIR conference in Catania and the SIDE (Rome) for critical comments and constructive suggestions on a previous version of this manuscript and Caroline Graf for sharing some data with us.

## 1. Introduction

It has been claimed that “institutions rule” (Rodrik et al. 2004), or even that “informal institutions rule” (C. Williamson 2009). By now, a number of empirical studies provide evidence in favor of that claim. Gorodnichenko and Roland (2011, 2017, 2021) demonstrate that institutions promoting individualism—as opposed to collectivism—not only explain cross-country differences in growth and productivity, but also in innovativeness and the establishment of democracy. Informal institutions can explain within-country income variations across a number of European countries (Tabellini 2008a, b, 2010) but also differences in the extent of the welfare state (Gründler and Köllner 2020). Other recent studies show that informal institutions are closely aligned with corresponding moral systems (Enke 2019, 2023, Henrich 2020, Moscona et al. 2020, Schulz et al. 2019).

Gutmann et al. (2025) recently showed that constitutional compliance, i.e. the degree to which governments comply with the text of their constitutions, is a function of both individualism and the willingness to accept hierarchical relationships: constitutional compliance is higher in societies with a high degree of individualism, and smaller when there is a greater willingness to accept hierarchies. Both individualism and the acceptance of hierarchies are backed by dozens of informal institutions (Hofstede 1980). Governments complying with constitutional rules are an important aspect of the rule of law, implying that informal institutions are also important for the degree to which countries can be expected to actually implement the rule of law.

The degree to which entire societies are attributed different levels of “individualism”—and similar traits—is, however, a very coarse measure for informal institutions. According to most definitions, individualism is not an institution itself but can be thought of as a multitude of informal institutions. The same is true for collectivism. Coming up with precise measures for the underlying institutions constitutes a serious challenge. Furthermore, the prevalent use of coarse measures prevents our possibilities to study how the different informal institutions making up those measures actually relate to each other. Not considering the interdependence among different informal institutions may lead us astray as we over (or under) estimate the relevance of single institutions.

Ostrom (1996, 2008) pointed out that informal institutions “may be almost invisible to outsiders, especially when they are well accepted by participants who do not even see

them as noteworthy.”<sup>1</sup> Empirical research interested in identifying the relevance of informal institutions has, thus, been hampered by the lack of appropriate indicators.

Our paper aims at changing this and contains a collection of indicators that may capture different types of informal institutions—and their differences across countries. Institutions—loosely defined as commonly known rules endowed with a sanctioning mechanism—are always directed at a multitude of people: They serve to make the behavior of others more predictable. Institutions have therefore often been described as generally serving to increase predictability in the behavior of peers (horizontal) or of authority (vertical), which reduces transaction costs. Decisions regarding the inclusion of indicators into our dataset were therefore made on the basis of this function.

To our knowledge, this paper contains the most comprehensive dataset on informal institutions currently available. We build on previous work and combine informational sources that are usually assessed separately. In addition to surveys, we report results from experiments that have been implemented in various countries based on the hypothesis that a variety of standard experiments could serve to infer underlying institutions. Our dataset further relies on a variety of language traits, as the rules of languages are conventions, an important type of informal institutions.

Our study adds value as we systematize all these data in a way that facilitates comparison to other types of data. This study is, however, not a meta-study; our focus is on offering proxies for a number of informal institutions, as this is a precondition for empirically testing their often-claimed relevance. We intend to make this a “growing database” by regularly adding additional datapoints. Beyond offering a broad palette of proxies of different informal institutions, we go one step further and explore their interdependence—and compose entire systems of institutions.

We proceed as follows: In the next section, we briefly present our taxonomy of informal institutions. In Section 3, we discuss three issues: why both the individual and the collective level are necessary sources of information to properly grasp informal institutions, how granular our measures of institutions ought to be, and how best to delineate the groups or collectivities for which we want to provide data. The various measures—and types of institutions they capture—included in our dataset are briefly described in Section 4. Section 5 contains a first look at the data and some insights that

---

<sup>1</sup> Voigt (2018, 8f.) contains a list of seven reasons spelling out these difficulties in more detail. In that paper, a number of proposals on how informal institutions could be measured were made without, however, offering concrete data.

emerge using a system perspective, relying on correlation network analysis. We conclude in Section 6.

## 2. The Taxonomy of Informal Institutions

Institutions can be defined as commonly known rules used to structure recurrent interaction situations that are endowed with a sanctioning mechanism (see, e.g., Voigt 2013, 2019). North (1990) distinguishes between formal and informal institutions, using the rule component as the criterion. Since North bases his distinction on the formality of the rule, we typically refer to the distinction as formal versus informal rules.

In addition to distinguishing the type of rules that make up an institution, one can also distinguish who is expected to sanction the actor who has not complied with a rule. If the state sanctions rule-breaking, the enforcement is external to society. We suggest considering the institution “external”. If rule-breaking is sanctioned by members of society, we propose to consider the institution “internal” (Voigt 2019). In addition, some kind of rules and standards that surround us are not sanctioned by anybody apart from ourselves. Table 1 contains a proposal for a relatively fine-grained taxonomy of institutions that focuses on who does the sanctioning. In this paper, the focus is on the first three types of institutions.<sup>2</sup>

In game theory, a convention is defined as an equilibrium of a coordination game having two or more possible equilibria in pure strategies (Lewis 1969; Sugden 1986). This implies that no individual can make herself better off by unilateral defection; each individual has incentives to comply with the convention. The most frequently cited example of a convention are the rules of the road: If everybody drives on the right side of the road, I cannot make myself better off by driving on the left side of the road. The rules of language are another example: if my aim is to be understood by others, I better follow the pertinent rules of grammar. Since conventions are accepted as exogenous by most individuals, a collective measure on the entire group sharing the convention is apt (type 1 in Table 1). We thus include some basic grammatical rules in our database.

---

<sup>2</sup> Although non-compliance with type 4 institutions is not sanctioned by representatives of the state, they frequently rely on highly formalized rule systems. Identifying and measuring the relevant institutions is thus fairly straightforward and of less interest in our context.

Table 1: Types of Internal versus External Institutions

Rule	Form of Enforcement	Type of Institution	Examples
1. Conventions	Self-enforcement	Internal type 1	Grammatical rules
2. Personal norms	Self-commitment	Internal type 2	The Ten Commandments; the categorical imperative; honesty
3. Social norms	Spontaneous informal societal enforcement	Internal type 3	Rules of social conduct; family ties
4. Formal private rules	Organized private enforcement	Internal type 4	Private arbitration courts
5. Rules made by parliament	Organized state enforcement	External	Private law; criminal law

Source: Voigt 2019

Type 2 institutions (i.e., personal norms) are unique as they deal with the behavioral constraints that individuals impose upon themselves. Non-compliance with one's own standards is not sanctioned by others but can lead to guilt and loss of self-esteem (Schwartz 1973). But, of course, the type of restrictions an individual puts on herself will be heavily influenced by her environment. Personal norms are best identified at the individual level, but they are likely to be affected by and co-evolving with collective features, including social norms.

Social norms (i.e., type 3 institutions) are informal rules; non-compliance with them is sanctioned by members of one's group and they are, hence, both informal (with regard to the rule part) and internal (with regard to the sanctioning) institutions. They are of central interest here, as they rely on the explicit interaction of a norm-breaker and one (or more) people sanctioning the norm-breaking.

Type 4 institutions (i.e. explicitly agreed upon private rules) are the outcome of explicit/deliberative choices of a subgroup of society. External institutions are laws and non-compliance with them is sanctioned by representatives of the state. Since the rules with regard to both type 4 internal institutions as well as with regard to external institutions are usually written down, ascertaining them implies less of a challenge than with regard to the other types of institutions. This is why we do not include them in our dataset.

### **3. Choosing the Appropriate Unit of Observation**

In this Section, we discuss three issues, namely in which cases the individual and in which cases the collective level is a better source of information for informal institutions. We argue that the two levels are complementary and should, hence, both be taken into consideration. Second, we discuss how granular our measures of institutions ought to be. Third, we discuss how best to delineate the groups or collectivities for which we want to provide data.

#### **3.1. Individual and Collective Measures**

Research on the importance of institutions for social and economic development has often been embedded in a broader literature focusing on the potential relevance of culture (e.g. Guiso et al. 2006, Spolaore & Wacziarg 2013, Alesina & Giuliano 2015 or Nunn 2020). Hundreds of definitions of culture have been proposed. Here, we follow Guiso et al. (2006) in defining culture as “those customary beliefs and values that ethnic, religious and social groups transmit fairly unchanged from generation to generation” (Guiso et al. 2006, p. 23).

According to social psychology, culture can be understood as the collectively shared responses of a group to the challenges posed by the specific environment of that group. As the environment is shared by all members of the group, a series of cognitive processes tied to overcoming environmental challenges are experienced by multiple individuals simultaneously. In such standardization then lies the root of a collective phenomenon (Amendolagine and von Jacobi 2023; Greif and Mokyr 2017; Oyserman 2017) that reinforces collective identities and fast-thinking (Tversky and Kahneman, 1983).

Geographical psychology has further corroborated this hypothesis, finding that values tend to cluster geographically (e.g. Rentfrow 2010). The collective nature of culture resides, therefore, in a common determinant (geographical conditions) and in the response to it that is consolidated by being shared (intra- and intergenerationally) by more than one person.

Culture is thus situated on the collective level but has a direct impact on individual behavior. Yet, culture needs to be transmitted from generation to generation. This needs to be done by both individuals (parents to children) and society (previous generation to new)—through vertical and oblique transmission, respectively (Bisin and Verdier 2022). In sum, culture as a collective phenomenon impacts individual behavior. Simultaneously, individual behavior sustains culture, but it is also the driving force



behind cultural change. This suggests that cultural traits may be transmitted by two channels—a collective and an individual one.<sup>3</sup>

To make different cultures comparable with each other, measures relying on the individual level, as well as those relying on the collective level, are desirable. Institutions are always situated at the collective level as they are to structure the behavior of multiple individuals. They can only serve their function—namely, to reduce uncertainty—if others are also subject to them.

Precisely ascertaining the content of an informal institution is more challenging. It is the defining trait of social norms that the involved rules are informal and often not written down anywhere. One way to ascertain social norms is then to observe how people behave: if in a given situation, most people do x but refrain from doing y, we may infer that x is a valid social norm for the respective situation. The behavior could either be a convention or based on a personal norm (i.e., a type 1 or type 2 institution). If, moreover, we observe that at least some of the people who do choose y are sanctioned by others, we have identified a type 3 institution. In this case, the internal institution has been inferred based on individuals' behavior, hence at the individual level.

We refer to a measure as “individual level” if individuals reveal their position on a specific issue as individuals. This could be the endorsement of a specific norm or the claimed or effective behavior in a specified situation. The most common way of producing individual level measures are surveys which have been complemented by experiments, which has the benefit of participants all in the same precisely delineated situation.

There are many surveys collecting responses that are subsequently aggregated in order to construct collective measures.<sup>4</sup> The World Values Survey is probably the best-known example following such an approach. Hofstede's (e.g. 2003) well-known dimensions used to identify national cultures, or House's (2004) Global Leadership and Organizational Behavior Effectiveness (GLOBE)-project are other examples.

---

<sup>3</sup> Chen (2013) finds supporting evidence for such transmission in the case of language—a collective internal institution of type 1. Yet his findings only hold if specific levels of aggregation are considered (Roberts et al. 2015).

<sup>4</sup> Usually, simple means giving all respondents equal weight are calculated. Smith & Bond (2019) assume that higher status groups may have more influence on the values and norms held by society as a whole and therefore suggest attaching greater weight to members of high-status groups in calculating this type of “individual based collective measure” (Bisin & Verdier 2022 refer to *prestige-based cultural transmission*).

Beyond such well-known variables, we explicitly include less-used collective measures of informal institutions, such as the grammatical rules of a language, and a series of aspects of family formation and maintenance. Collective-level measures are likely to convey expectations about desirable/accepted behavior.

Regarding individual-level measures, experiments are useful to capture concrete behavior, whereas survey responses can be thought of as individual positioning on a spectrum of possibly accepted behaviors. Some collective measures can easily be derived from individual-level data, as the latter can be used to inform on the range and shape of a distribution. The mean, median and mode of such distribution are only measurable at the collective level, not the individual one. Dispersion around such centrality measures is likely to reflect the degree to which deviation is tolerated (as in “tight” vs “loose” cultures; Gelfand et al. 2011). While dispersion around the socially expected behavior can only be captured by relying on individual measures, the resulting distribution is a collective property—that can only be seen and recognized by combining individual behaviors. The distribution of individual behaviors is therefore a collective measure.<sup>5</sup>

In summary, we argue that in order to ascertain prevalent institutions, relying on information derived from both the individual as well as the collective level can be useful.

### 3.2. On the Granularity of Institutions

In principle, it appears desirable to have information on how very specific interaction situations are structured, implying that one would have hundreds or even thousands of relevant institutions that could be compared across groups. Although this may be desirable in principle, it is unfeasible from a practical point of view (see also Robinson 2013). The question then becomes: how thematically precise should the variables used to measure informal institutions optimally be?

Acemoglu and Robinson (2025) propose a surprising answer: namely, that we need both, abstract and precise measures of institutions. Their argument stresses how more abstract (coarse) and more specific norms co-exist within a society and form configurations.

---

<sup>5</sup> Inferences from the individual to the collective level—and vice versa—need to be treated with care as the ecological as well as the atomistic fallacy loom large. The ecological fallacy implies transposing average values observed at the group level to the individuals that compose it. The fallacy is due to *within-group* inequality (Hofstede 1980; Robinson 1950). The atomistic fallacy implies that mere aggregation of individual-level information may not correctly represent a group-level phenomenon (Hox 2002, Richards et al. 1990), violating cross-level isomorphy, mainly because the interconnections between individual-level measures provide additional dimensionality and content to the group-level phenomenon.

They hypothesize that more abstract elements—which are valid across different interaction situations—can more easily be combined in different ways with other, more specific norms. This is consistent with findings in social psychology, in which broad concepts—like individualism—that can fluidly be activated in different situations play an important role for the activation of associative knowledge networks in the human brain (Oyserman 2017).

Translating the framework of Acemoglu and Robinson (2025) into our perspective centered on informal institutions, we can summarize their conjecture as follows: More abstract institutions are more easily combinable with a multitude of other informal institutions because their interpretation can vary. As a consequence of their thinking, more specific informal institutions—that systematically appear in conjunction with other informal institutions—are less amenable to change. We expect that the use of a multitude of informal institutions, both abstract and specific, allows for additional insights. The inclusion of multiple measures permits to adopt a systems approach in which the study of interdependencies between informal institutions becomes key (see Section 5).

### 3.3. Geographical Aggregation

Before collecting datapoints for our dataset, another question must be answered: which is the appropriate delineation of the group for which internal institutions are to be documented? Today, most available statistics focus on the nation-state level, implicitly assuming that this is the most appropriate delineation. There are, however, good reasons to assume that this is not always the case.

For example, very large nation states may be populated with quite diverse groups following vastly different cultural traditions. Additionally, many nation-state borders were not drawn in congruence with homogenous ethnic or linguistic groups.

Yet, most available indicators offer information only for the nation-state level. This implies that some aggregation rule is necessary for cases in which there are clear differences within nation states. An obvious example is a country in which different languages are spoken in different regions.<sup>6</sup> Of particular importance is the identification of possible criteria for identifying one's belonging to a group. While this can be based on language in some contexts, it may be based on ethnic, religious, or class affiliation

---

<sup>6</sup> For our treatment of language variables, and the combination of different languages coexisting in the same country, see the dedicated section in the codebook included in the Supplementary Material.

in another (Henrich 2016; Kinzler and Dautel 2012; Kinzler 2021). In-group versus out-group considerations result from such identification processes and can influence the way individuals interact with others, depending on whether they belong to one's group or not (Akerlof and Kranton 2010; Hahm et al. 2023; Whitt and Wilson 2007).

Coding institutions at the subnational level is definitely a desideratum. But given the scarcity of more disaggregated data, most of the data contained in our dataset is aggregated at the nation-state level.

## **4. Variables Included in the Global Dataset**

### **4.1. Criteria Used for Including Variables**

This section serves to briefly describe the variables included in our dataset. The descriptions themselves can be found in subsections 4.2 through 4.4 below, and more technical information on the construction of each variable is detailed in the codebook included in the Supplementary Material. Here, we briefly spell out the criteria used for including some variables but not others. There are theory-based considerations as well as pragmatic ones. We are particularly interested in including institutions that are conjectured to be conducive to economic development.

Economic growth and development are facilitated if there are institutions enhancing impersonal exchange. This is the first group of institutions we aim to include in our dataset. Moreover, in many situations, individuals are better off if public goods are provided. We therefore also seek to include informal institutions that are conducive to the ability of a group to provide itself with public goods.

These are two important functions that institutions can have for economic development. Some institutions may be conducive to both functions. Think of institutions encouraging cooperation. To make cooperation sustainable, it seems crucial to sanction those who do not cooperate or contribute their fair share ("pro-social punishment" e.g. in Fehr & Schmidt 2006).

The pragmatic considerations guiding the choice of variables to be included in the global dataset are more mundane. To ensure comparability of results across countries, survey questions should have been formulated identically across relevant countries and be statistically representative at the chosen level of analysis. In the case of experiments, most experiments to date have been run with student samples. Noting this, if these data are included in our dataset, we refrain from insisting they are representative of the population of an entire country. We only include experiments that are similar in their main traits, including the sampling process. We aim to include variables only if they are

available for at least 30 countries, as that seems to be the threshold for doing any econometric analysis.<sup>7</sup> Often, the starting point in our search for variables were surveys or meta-studies. In some cases, we updated them to include studies published until 2022.<sup>8</sup>

This is the first exercise to map informal institutions on a global scale. To keep it tractable, we do not report results for specific decades or even single years. Experiments run and surveys conducted over the last 20 or 30 years are all classified as contemporaneous, but information on family-based institutions as historic. Each of the following subsections describes available data for each type of internal institution as delineated above (Section 2). Within subsections, we begin with experiments (both lab and field), followed by surveys.

While surveys are often designed to be run in a large number of countries, this is typically not the case in experiments. Their main purpose is to see if differences in treatment can induce participants to behave in different ways. By comparing results of (standard) experiments run in different countries, we use them in a different way, namely to elicit differences in behavior across the various nations where they have been played. Although we rely on standard experiments only, we rely on many different sources and contexts in which the experiments have been implemented. We therefore expect measures based on experiments to be noisier or less precise (see our distributional analysis in Figure 1 in the Appendix).

For results from experiments to be included, they need to fulfill a number of criteria: Following Engel (2011, 586), a peer-reviewed publication is not necessary, as including working papers may be a guard against publication bias. In case a particular game has been played multiple times in a single country, we calculate the mean across different observations within the same country. Where available, we also reconstruct the confidence interval for the upper as well as the lower bound (see figure A1 in the Appendix). If there are various treatments for the same game (say the dictator game is played by males and females), each treatment is included separately. Larney et al. (2019) found little effects caused by differences in stake size in both the Dictator and the Ultimatum Game, which is why we deem it unnecessary to explicitly control for.

In citing survey and experimental data, we propose to adopt a common perspective that puts the centrality (average) of an institution and, where available, its dispersion

---

<sup>7</sup> The only exception are two variables on the Public Good Game, which have less than 30 observations as we intentionally split simple versions from 2nd-stage versions.

<sup>8</sup> Thöni (2019) is an overview of both the potential as well as the challenges of cross-cultural experiments.

(variance) at the center: while averages capture the expected behavior, dispersions describe their precision.

## 4.2. Conventions (Type 1 Institutions)

### *Language*

Languages are highly complex instruments that allow us to transmit large chunks of knowledge from generation to generation. Thus, they are not only a substantial part of culture, but also a means for transmitting it. They further allow us to create abstract realities and thereby to establish meaningful order. Language is a type 1 institution because no speaker can make herself better off by not complying with the (implicitly agreed upon) rules that structure it. Languages can induce effects on behavior because their grammatical details enshrine differences in the way we encode time (Chen 2013), flag differences between men and women (Givati and Troiani 2012; Hill and Mannheim 1992), between certain and possible scenarios (Kovacic et al. 2016), but also in expressing collectively shared interpretations of feelings (such as ‘*saudade*’, ‘*Weltschmerz*’, or ‘*hygge*’).

Language structures have been used as instrumental variables for culture (Licht et al. 2007; Tabellini 2008a), but a separate literature stresses how language may directly affect cognition, adding a channel of influence to the traditional one of culture (Mavisakalyan and Weber 2018). According to the Linguistic Relativity Hypothesis (LRH; also known as the *Sapir-Whorf Hypothesis*), language influences thought. While contested by Chomsky (1957) and Pinker (1994), recent research has produced new evidence showing that language affects cognition, and thereby indirectly also behavior (Elpers et al. 2022, Mavisakalyan et al. 2018, Regier and Kay 2009, Slobin 2003, Wei et al. 2023).<sup>9</sup> But the recent revival of the LRH suggests that universal mechanisms of cognition concur with local linguistic conventions that rely upon other, non-universal factors (Regier and Kay, 2009).

The language we speak implicitly enacts selective attention and memory, because it puts emphasis on different aspects of reality. Language in this way affects people’s perception of reality – and of their place within it. We treat language and some of its features as conventions (type 1 institutions), without assuming a position regarding the precise channels through which its effect may unfold. Beyond representing a typology

---

<sup>9</sup> It still remains unclear whether the effect occurs at the level of thought or at the moment in which thought is enacted through speaking/expressing oneself (Slobin 2003).

of internal institutions that can have effects on economic outcomes (Chen, 2013; Mavisakalyan and Weber, 2018), we expect language to play a role in attributing more or less weight to certain rules or choices over others. We focus on four linguistic features that have been shown to affect individual behavior and economic outcomes.

**Grammatical gender.** The presence of grammatical gender captures a convention according to which men are distinct from women.<sup>10</sup> While this does not automatically imply discrimination, the empirical evidence suggests that when grammatical gender characterizes local languages, men tend to be perceived as having a more legitimate access to jobs, to health and to education (Bhalotra et al., 2015; Givati and Troiano, 2012; Mavisakalyan and Weber, 2018; Santacreu-Vasut et al., 2014).

**Future tense.** Another interesting linguistic trait is the presence of a compulsory future-tense. If a language obliges its speakers to mark the future as separate from the present, this stands for a convention to displace the future farther away (Chen, 2013). Such temporal displacement has important implications for intertemporal preferences and forward-looking behavior: the future is not in continuity with the present, and therefore implies higher social discount rates (Mavisakalyan et al., 2018). In other words, where language asks to mark the future explicitly, it is less real than the present, leading to reduced long-term orientation. Future time-reference (FTR) in a language can be weak (no difference between present and future) or strong. Weak FTR has been found to positively associate with saving-rates, health behaviors and retirement assets (Chen, 2013), environmentally responsible behavior (Mavisakalyan et al., 2018), and to some preferences such as patience, positive reciprocity, trust and altruism (Falk et al., 2018).

In attempts to document causality regarding these findings, a number of experiments have been run which shed some doubts on the so-called linguistic savings hypothesis (e.g. Chen et al. 2019 or Angerer et al. 2021). Given the potentially far-reaching economic effects of this language feature, we include it into our database nevertheless.

**Pronoun-use.** Another convention we record among language features is the use of pronouns. In languages in which the personal pronoun “*io*” in “*io canto*” (I sing) can be dropped without compromising the meaning of the sentence, language enshrines a convention according to which it is not necessary to stress one’s individual role separately (in English, in contrast, this would not be possible as “*I sing*” cannot be substituted by “*sing*” without loss of meaning). For this reason, pronoun-drop in a language has been associated with more collectivist cultures (Davis and Abdurazokzoda

---

<sup>10</sup> Such distinction can manifest in many different ways, example manifestations are „der“ vs. „die“ in German, or „actors“ vs. „actresses“ in English.

2016; Davis and Williamson 2016; Kashima and Kashima 1998; Licht et al. 2007; Tabellini 2008a).

**Mood.** Utopian thinking is another element we can detect in language conventions. To the extent that a language foresees different grammatical solutions to introduce non-indicative moods (that do not describe the world as is, but as it could be) it can influence the perception of reality of the speaker. Subjunctive moods accommodate ambiguity (Gironde et al., 2019). A greater use of such subjunctive moods may associate with greater risk-aversion (Kovacic et al., 2016), possibly because the convention alerts to multiple possible states of the world, which increases uncertainty (Mavisakalyan and Weber, 2018). “Se tiver comida” – *if tomorrow there may (eventually) be food...* is a grammatical construct that stresses the uncertainty regarding the future situation. While this is a very common construct in Portuguese, it is much more difficult to verbally express the same scenario in German, or Italian, or English. As risk-aversion clearly has economic consequences, we include this language feature in our dataset.

To sum up: language features can be thought of as conventions. In future extensions of the database, including other conventions (such as medium of exchange, adequate physical distance to others, tipping conventions, dress code etc.) is definitely a desideratum.

### 4.3. Personal Norms (Type 2 Institutions)

Under type 2 institutions (i.e., self-imposed personal norms), an individual not conforming to her own behavioral standards sanctions herself—most likely by some feelings of guilt or loss of self-esteem. *Prima facie*, one could therefore expect huge variance regarding these institutions across individuals. These considerations may, however, overestimate the degree of variance found among members of single societies.

In his writings, Henrich has repeatedly made the point that internalized social norms are likely to reflect the prevalent local norms. In our terminology: type 2 institutions (personal norms) are likely to reflect type 3 institutions (social norms), at least partially. Henrich (2016, 197) argues that natural selection might have favored norm internalizers: “Such motivations may help us avoid short-term temptations, reduce cognitive or attentional loads, or more persuasively communicate our true social commitments to others.” Further (ibid., 199): “In short, to survive in a world governed by social rules enforced by third parties and reputations, we became norm learners with prosocial biases, norm adherers internalizing key motivations, norm-violation spotters, and reputation managers.”



Type 2 institutions reduce decision making costs: For example, “internalized social norms help guide us through complex social environments, allowing people to automatically – without conscious reflection or complex mental calculations of the reputational consequences – do the “right thing” (i.e., comply with local norms)” (Henrich, 2016:188).<sup>11</sup>

Personal norms held by individuals cannot be “observed” by others. Instead, they need to be derived from observable features, such as behavior in the lab, in the field, from claimed behavior, or simply from stated convictions in surveys. Our dataset comprises variables relying on all of these elicitation methods.<sup>12</sup>

### *Fairness Norms*

Widely shared fairness norms make mutual exchange more attractive. When people believe that others will act fairly, they are more willing to cooperate with each other, leading to higher levels of economic activity. Following numerous economists (e.g. Camerer 2011), we propose to rely on the dictator game as an indicator for fairness norms.

The dictator game is a simple one-person game. The decider (“dictator”) is asked to divide a pie (usually a sum of money) between herself and some other person, and the experimenter then allocates the money accordingly (Forsythe et al. 1994). There is no second stage. Neither the anonymous other player nor the experimenter intervene in any way, implying that it is solely the personal norms of the “dictator” who are decisive. Engel (2011) is a meta-study on the dictator game which analyzes 131 papers on the topic. We are adding additional results for the dictator game that were run and published between 2010 and 2022.<sup>13</sup>

---

<sup>11</sup> The degree to which people follow norms “automatically” has even been tested in the lab. Rand et al. (2012, 2014) studied how long it took people to decide upon the size of their contribution in a Public Goods Game.

<sup>12</sup> In psychology, the distinction between descriptive and injunctive norms has played an important role. Cialdini et al. (1990) propose to think of injunctive norms as “what most others approve or disapprove” and of descriptive norms as “what most others do.” In a sense, observed behavior—no matter whether in the lab or the field—reflects descriptive norms, whereas survey answers are closer to reflect injunctive norms.

<sup>13</sup> Experimentalists are often challenged by questions doubting the external validity of their results. With regard to fairness norms, there is some evidence that behavior in the lab does mirror behavior in the real world (Franzen & Pointner 2013).

### *Moral Universalism*

Moral universalism is the belief that there are several moral principles that apply to all individuals regardless of ethnic or religious differences (Enke 2019). If moral universalism is widely shared, it reduces transaction costs and thereby fosters cooperation both domestically and across nation-states.

Recently, Cappelen et al. (2025) had survey respondents from 60 countries play a variant of the dictator game. In it, respondents were asked to (hypothetically) divide the equivalent of \$1,000 between a member of their in-group and a stranger. The “stranger” was from the same country as the respondent in what the authors refer to as “domestic universalism” and from another country in “foreign universalism”. We categorize the elicited degree of universal moralism as an internal institution of type 2 as the respondents cannot be sanctioned by anyone for whatever division of the money they propose.

### *Cooperate for the Common Good*

Many public goods, such as roads or public education, are preconditions for the efficient production of private goods. The voluntary production of public goods is considered precarious, but may be important for economic development, particularly in situations in which the state does not provide the necessary public goods. To recognize individuals’ propensity to contribute to the public good, we include the public goods game into our dataset.

Public good games (PGGs) are usually played in groups of four. Each participant receives an initial endowment and then chooses how much of the endowment to put in a joint fund with the other players and how much to keep for herself. While the amounts in the private funds do not change, the amount in the joint fund is subject to a positive multiplier that is greater than one. After the multiplication, the joint fund is divided equally between all players, independently of the specific amount each player put into it. This means that it would be collectively rational that all players put their entire endowment in the joint fund. Individually, however, it is rational to keep one’s endowment in the private fund and hope that others will put at least part of their endowment in the joint fund (Ledyard 1995).

The game ends after all participants have made their decision on how to split their endowment and the experimenter has paid them accordingly: There is no sanctioning by others. This is why we interpret this simple version of the PGG as a proxy of the individuals’ willingness to contribute to public goods. The contributions can be

interpreted as the degree to which a group of people (a society) is capable of providing themselves with public goods.

As a starting point for our search for data, we relied on Chaudhuri (2011). Here, too, we are adding additional results for this game that were run and published between 2010 and 2022. We report the mean proportion of the original endowment that participants contribute to the joint fund. As we are interested in personal norms implied, we do not take experiments with a sanctioning stage into account. Our focus here is on the different countries in which the public goods game has been played.

### *Honesty*

If most members of society are honest, then it is rational to trust others. Trust has been shown to be highly correlated with both economic and political development (e.g. Knack & Zak 2003, Bjørnskov 2012). Diffused honesty reduces transaction costs by promoting trust. This is why we propose to include measures of honesty into our dataset. Honesty can be interpreted as a specific personal norm and, therefore, as a type 2 institution. In case my personal norms demand me to be honest but I am not, I could sanction myself with feelings of guilt.

Experimental economists have developed ways to infer into the honesty of their participants (Fischbacher & Föllmi-Heusi 2013). Usually, participants are asked to think of some number between 1 and 6 and are then shown a randomly chosen number between 1 and 6 on their monitor. In case the two numbers match, the participant gets some reward; in case they do not, there is no reward. However, experimenters cannot know what number the players are thinking of. This implies that it is impossible to know whether a participant has honestly claimed a match in a single round of the game or not. Yet, if the game is repeated many times, inferences can be made on the basis of probability theory. Abeler et al. (2019) ran a meta-study on the preferences for truth-telling in which they relied on 90 experimental studies that were ran in 47 countries. The results reported in that paper are incorporated into our database.

Recently, Graf et al. (2024) proposed an additional measure for honesty. The authors rely on questions contained in the PISA tests that measure students' self-reported familiarity with several mathematical functions (e.g., "exponential function") on a scale from 1 ("never heard of it") to 5 ("know it well, understand the concept"). Three non-existing concepts ("proper number", "subjective scaling", and "declarative fraction") were included into the test to elicit students' level of overclaiming. The authors then operationalize honesty as the degree to which students honestly report their unfamiliarity with the fake concepts. We rely on the aggregate score over all three fake concepts.

### *Helpfulness*

Societies in which people are likely to help each other benefit from several advantages. As with regard to honesty, a higher level of general trust is likely to make mutually beneficial cooperation more likely. We propose to include a measure of helpfulness in the dataset that was elicited in the field—and not in the lab. The experimental variables described until now have all been elicited via lab experiments. This is different in the case of the lost wallet experiment, which takes the question to what degree people are likely to help from the lab to the field. The authors of this study (Cohn et al. 2019) visited 355 cities in 40 countries and “lost” more than 17,000 wallets. As this study was conceived of as a cross-country study from the outset, the data produced by it can be directly plugged into our global dataset.

### *Caring for Others: Observed Altruism*

Helping others selflessly is one kind of altruism. Another way to behave altruistically is to donate blood. Both behaviors can be interpreted as holding a personal norm to help others: altruistic individuals strive to be kind, which is why they help others. The World Health Organization (2022) publishes a global report that contains data on unpaid voluntary donations at the country level. We include this variable to our dataset as it reflects real behavior of millions of people as recorded by official statistics.

### *Claiming to Help Others: Declared Altruism*

Yet another way to elicit altruistic behavior is to survey people and ask if they have behaved altruistically in the past. The U.K. based Charities Aid Foundation (2022) has run its World Giving Index (WGI) annually since 2009. The Index is composed of the answers to three survey questions inquiring into different aspects of altruism, namely (1) helping a stranger, (2) donating money, and (3) volunteering time. The 2022 edition of the survey is available for 119 countries.<sup>14</sup>

---

<sup>14</sup> There is, however, evidence that charitable giving is also affected by social pressure (della Vigna et al. 2012) and would, therefore, not be solely based on ethical rules.

#### 4.4. Social Norms (Type 3 Institutions)

In type 3 institutions, a rule breaker may be sanctioned by another individual or group of peers in case she does not comply with the respective rule. Since the act of sanctioning is costly in some manner, it is never guaranteed to occur.<sup>15</sup>

##### *Sharing Norms*

In section 4.3, we introduced the dictator game as a variable that proxies fairness as a personal norm. In this section, we deal with interactions and the possibility to sanction non-rule complying behavior by others. These are therefore social norms, or type 3 institutions. The ultimatum game is a game that creates such a situation. It is played by a proposer and a responder. The proposer gets to propose the division of a sum of money between himself and a responder. The responder can either accept or reject the proposal. In case she accepts, the experimenter divides the sum accordingly. If she rejects, neither player gets any money (Güth et al. 1982).

In this game, the proposer, who is interested in securing a large share of the pie, needs to anticipate whether the responder will agree to her proposal or not. In case the proposer asks for too large a share, she will be sanctioned by the responder. In order to compare the behavior in this game with that in the dictator game, we record the mean offer as we did in the dictator game. In this case, the mean offer is to indicate the collectively accepted sharing norm as perceived by the proposer.

##### *Sanctioning norms*

According to standard economic theory, voluntary contributions to the provision of public goods are unlikely to occur. And if they do, the amount contributed will be sub-optimally low. But if societies manage to establish an institution that sanctions those who do not contribute their fair share, cooperation may be sustainable (see, e.g. Fehr and Gächter 2002). We refer to such institutions as sanctioning norms. Since they refer to an interaction between (at least) two persons, we think of them as type 3 institutions.

By now, many experimenters have added a second stage to the public goods game already described above. The second stage gives participants the possibility to sanction the other members of the group, usually at a cost to themselves. By now, the PGG with the option to punish has been played in many countries. Out of the reported results, we include the average share proposed under the threat of being sanctioned.

---

<sup>15</sup> To counteract that possibility, many societies have developed meta-norms that are to be applied if someone who is supposed to sanction defectors does not do so.

### *Third Party Sanctioning*

Large complex societies are more likely to thrive if its members are prepared to sanction those who have treated others unfairly, as this increases the cost of non-compliance with general rules and will make deviating less attractive. This is commonly referred to as altruistic punishment (see, e.g., Boyd et al. 2003 or Fehr & Gächter 2002) and can be interpreted as *res publica* orientation. Compare this to societies whose members will only sanction others when they themselves have been treated unfairly. This might be referred to as revenge or second-party punishment. Whereas the PGG with the second stage refers to the possibility to sanction those one has previously interacted with, the third-party sanctioning variable asks which actors hitherto noninvolved are ready to sanction wrongdoers. Falk et al. (2018) contains a variable capturing these options via survey responses, which we include in our dataset.

### *Particularistic Loyalty*

Loyalty can be an asset for societies if its members are loyal to its constitution, its laws, its government and so on. However, it may also be a problem if agents are loyal to small in-groups whose behavior is in conflict with general rules. We refer to this as “particularistic loyalty”. Particularistic loyalty can be based on a social norm if it is generally expected—and accepted—that loyalty to a small group ought to trump loyalty to the general rules of society. Particularistic loyalty can be thought of as an extension of 'amoral familism' (Banfield 1958) to a broader network of personal contacts. Stouffer and Toby (1951, 396) produced a vignette that encapsulates the tradeoff between either following some universal rules or being loyal to friends. The vignette has been used as survey question and in the meantime, responses from participants in 43 countries have been recorded (Trompenaars & Hampden-Turner 2020).

### *Whistle Blower Norms*

Corruption is conventionally defined as the abuse of power for private gain. Corruption only works if at least two actors cooperate: the giver of a bribe and its recipient. By agreeing on a deal both actors indicate that they expect to be better off as a consequence. In all likelihood, the successful cooperation between these actors would be to the detriment of others, for example, others bidding for a contract or taxpayers who will eventually have to make up for higher government expenditures. Corruption is therefore an instance of successful cooperation between parts of society that is to the detriment of others: in other words, involving negative externalities. This also serves as a reminder that successful cooperation does not necessarily benefit the society as a whole.

Corruption is tricky and is likely to depend on a number of internal institutions. Rather than trying to capture all of these in our database, we are interested in norms that deal with reporting instances of corruption, like the possible sanctioning of corruption. Such a norm can be interpreted as increasing the likelihood of criminal behavior being sanctioned. If it is widely accepted and followed, this reduces the expected utility from corrupt behavior, which should therefore occur less frequently.

The World Values Survey contains a question inquiring how high the risk of being held accountable for giving or receiving a bribe, gift or favor in return for public service. While this is not equivalent to the whistle blower norm, the likelihood of being held accountable is certainly related.

### *Family Institutions*

What we refer to as family institutions here are sanction-based rules regulating intra-family relationships. These institutions had been highly time-invariant for centuries, so we think of them as historic collective-level measures. According to Schulz et al. (2019): “... anthropological research suggests that kin-based institutions represent the most fundamental of human institutions and have long been the primary framework for organizing social life in most societies.”

With regard to marriage, institutions may regulate who can marry whom, who pays for the marriage ritual, or who gets the children in case one of the parents dies, and so on. Failure to comply may be sanctioned in various ways, such as a loss of status (Henrich et al. 2012, 658f.). This is why family institutions qualify as type 3 institutions.

French anthropologist Frederic Le Play (1895) proposed that the institutions determining the relationship between father and son indicate a society’s concept of liberty, whereas the institutions determining the relationship between brothers indicate the prevalence of equality. The first aspect is covered by the place where newly-wed couples live: do they move into their own home (neolocality) or do they stay with the parents of either spouse? The latter would indicate a lower degree of freedom, as parents are likely to interfere and encourage their children to stick to time-honored practices (“we have always done it like that”). This could lead to low levels of innovation. The second aspect (equality) is captured by inheritance rules: primogeniture (or ultimogeniture) would indicate inequality as heirs are not treated equally.

In addition to the two aspects proposed by Le Play, Todd (1985) asks whether cousin marriage is allowed or even encouraged. This could also be indicative of collectivism as marrying cousins fortifies connections to near others, and increases the perception that one is in a relative—not absolute—position with respect to others.

The different family types resulting from combining these three traits are depicted in Table 2. Gutmann & Voigt (2022) tested some of Todd’s hypotheses and found a number of noteworthy effects. The communitarian family type (and in particular its endogamous version) is associated with more racist attitudes, higher state fragility, and weaker civil society, for example.

Table 2: Schematic representation of family types

		Liberty		
		Low: married son stays with parents		High: married son moves out
Equality	Low: unequal treatment of brothers	Authoritarian (e.g., Norway, Sweden, Germany, Ireland)		Absolute nuclear (e.g., England, Canada, US)
	High: equal treatment of brothers	Endogamous community (e.g., Pakistan, Morocco)	Exogamous community (e.g., Russia, Mongolia, China)	Egalitarian nuclear (e.g., Spain, Italy, Poland)

Source: Gutmann & Voigt (2022)

In addition to the three traits recognized by Todd (locality, inheritance, endogamy), we also include a variable testing for matrilinearity, the tracing of kinship through the female line. Most of the African matrilineal societies practice matrilineal residence, they are less likely to have the custom of bride-price payments, and in some, women inherit land directly. It has been argued that matrilinearity directly affects the position of women in society (e.g. by Lowes 2020). As gender roles—and the possibility of gender discrimination—are potentially relevant for economic development, we include them here.

We also use some contemporaneous family variables to gather information on gender-discriminatory practices which rely on respective social norms. Based on the Gender, Institutions and Development Database (GID-DB) (OECD 2019), we extrapolate some information on women's reproductive and financial autonomy, and on women's standing in terms of having equal rights to divorce and inheritance as men. Beyond the law, these measures take actual customary, traditional, and religious practices explicitly into account.

#### *In addition: Trust*

We do not consider trust as an institution. We rather envisage it as an emergent feature of a system of (other) informal institutions. As trust and trustworthiness are closely associated with some of the institutions described in this paper (e.g. honesty, bilateral and multilateral cooperation), we include both survey and experimental trust measures: next to the general trust question of the World Values Survey, we include a “radius of



trust” measure that informs whether trust in others is confined to one’s in-group (family, neighborhood, people you know personally) or extends to one’s out-group (people you meet for the first time, of another religion, of another nationality). This measure was first proposed in Delhey et al. (2011).

The trust game as proposed by Berg, Dickhaut and McCabe (1995) has become the standard in eliciting trust and trustworthiness in the lab. Johnson and Mislin (2011) is a meta-analysis including experiments conducted across 35 countries. Their analysis includes papers available until around 2010. By extending the coverage until 2022, we are able to increase country coverage to 50 countries.

#### 4.5. Bundles of Norms

Many studies have not used single measures of informal institutions but have instead relied on bundles of norms. Hofstede et al. (1980; 2003) and Schwartz (2004), for example, have used a multitude of measures of informal institutions to classify nations in terms of their culture. Relying on principle component analysis and similar methods, the authors proposed a small number of dimensions central to the cross-country comparison of cultures. Hofstede (1980) suggested the following aggregated dimensions: power distance (acceptance vs. rejection of authority), uncertainty avoidance (acceptance vs. rejection of incurring risk), individualism versus collectivism (individual vs. shared achievements), and masculinity versus femininity (earnings, success vs. good relations, cooperation, security).

Schwartz (2004) went one step further, combining measures at the individual level with dimensions framed at the collective level. Using a smallest space analysis, he plotted responses to the question: “Rate the extent to which each of 56 values (proposed) are a guiding principle in your life”. This palette of individual values was organized along two continua (self-transcendence vs. self-enhancement; openness to change vs. conservation). Schwartz sought to transform the values of individuals into a collective cultural feature. He therefore reorganized the same values along three different, yet potentially combinable, continua; namely autonomy-embeddedness, hierarchy-egalitarianism, and mastery-harmony (of the human being with respect to nature).<sup>16</sup>

Both, Hofstede’s and Schwartz’s measures represent bundles of norms that define content hidden behind a latent factor like culture. In our analysis, we include their aggregated dimensions—the four original dimensions of Hofstede (1980), plus

---

<sup>16</sup> Schwartz’s approach is one that underlines that there may be no clear isomorphy between measuring values at the individual and at the collective level.

“indulgence” and “long-term orientation” introduced later. We include Schwartz’s “intellectual autonomy”, “egalitarianism” and “mastery”.<sup>17</sup>

## 5. First Insights Derived from the New Global Dataset

The primary purpose of this paper is to serve as a database for scholars interested in empirically ascertaining the relevance of informal institutions. We believe that it can help improve our knowledge regarding origins, interplay, and effects of informal institutions simply because it offers considerably more detail than more coarse measures, like individualism versus collectivism. As mentioned in the introduction, there are many studies that claim informal institutions are highly relevant for both economic and political development. However, many of these claims have either been insufficiently tested or never at all because data on informal institutions has been scarce or non-existent. The data assembled here can be used as a first step toward the testability of some claims. Our goal in this section of the paper is, however, different. Rather than trying to establish causal relationships between informal institutions and some outcome variables, we run a small number of analyses that serve to highlight some important aspects of the database here introduced.

We use descriptive statistics and correlation network analysis (Horvath 2011; von Jacobi, 2018) to shed light on aspects of informal institutions that typically remain unseen. As the use of correlation network analysis is still uncommon within the field of economics, we dedicate the next subsection to introducing it in more detail.

### 5.1. Correlation Network Analysis

Measuring and analyzing informal institutions in isolation entails the risk of over (or under-) estimating the relevance of single institutions. With regard to econometric analyses, this can also mean that relevant interaction effects are not taken into account. This is why we propose to analyze the data that are included in Section 4 of this paper by relying on a tool that has been devised to make interactions not only visible, but also at the forefront of analysis, allowing for a systematic approach. Correlation network analysis is such a tool.

---

<sup>17</sup> Schwartz’s measure of autonomy is built on two components (affective and intellectual) that are highly correlated with each other. To avoid a double counting, we only include one of them in section 5, namely intellectual autonomy. Both dimensions are included in the dataset.

A correlation network visualizes a correlation matrix computed over a pool of variables; in our case, informal institutions. Each variable is represented by a node, and each pairwise correlation coefficient by a link between the nodes. The links between nodes, known as edges, have greater weight when the correlation coefficient is stronger. The exact location of the nodes within the network is the outcome of an algorithm that puts nodes with more and higher correlation coefficients at the center of the system.<sup>18</sup> In a correlation network, one can immediately see whether variable A correlates with many other variables. If this is the case, A assumes a more central role in the system. Network analysis has produced a series of precise measures that capture such centrality (Freeman 1977). For informal institutions, we expect those assuming a more central role to be highly relevant for the context in which individuals interact. Contrarily, informal institutions remaining on the outskirts of a correlation network are less connected to the relevant context of values and norms in which people interact.

While a correlation matrix only makes bivariate correlations visible, correlation networks make connections between various bivariate correlations visible. If A correlates with B and B with C, one can visualize a concatenated correlation between A and C in the network. In case of exogenous shocks, concatenated correlation contributes to the diffusion of new impulses in the system. When public policies target a specific informal institution, the number of direct correlations of that institution allows one to better anticipate the potential spill-over effects of the intervention. Neglecting the concatenated network structure entails the danger of experiencing unintended side-effects. Therefore, correlation network analysis applied to informal institutions is likely to provide a more in-depth ground for policy recommendations.

Yet another layer of insight can be found when we inspect a correlation network for *motifs* (Fagiolo 2007, Onella et al., 2005) – for example a closed triplet in which A correlates with B, which correlates with C, which correlates with A in turn. Variables that form such a closed triplet reinforce each other and are therefore less amenable to change (Dey et al. 2019).

While we do not employ all of the three aspects mentioned here with equal depth, they enable insights beyond those attainable by relying exclusively on descriptive statistics.

---

<sup>18</sup> Correlation networks are weighted, undirected networks (Horvath 2011).

## 5.2. Correlation Networks Within Single Types of Informal Institutions

We start by analyzing the bivariate correlations between proxies for institutions of the same type (see Section 2). Within the typology of conventions, it is noteworthy that languages disposing of grammatical gender and those containing the possibility to express different epistemic possibilities are very highly correlated ( $r=.82$  based on 92 countries, significant at the 1% level). Given that the possibility of pro-noun drop has been associated with higher levels of collectivism (based on the Hofstede dimension of individualism/collectivism), we report that the correlation between the two is confirmed, with  $r=0.58$  derived from 38 observations, significant at the 1% level.

Among personal norms (type 2 institutions), it is interesting to look at the correlation between observed altruism (as proxied by the percentage of people who voluntarily donate blood) and the share of survey respondents who claim to have helped someone else recently, either by helping a stranger or by having donated time or money. The respective bivariate correlations are all insignificant, possibly adding another empirical example for what has been discussed as self-deception for many years (see, e.g., von Hippel and Trivers 2011). Yet, the low correlations may also indicate that we are dealing with two very different—and non-correlating—types of altruism. In the appendix, we report all bivariate correlations and statistical significance grouped by type of institution (Tables A3-A5).

Among social norms (type 3 institutions), it turns out that the historical endogamous communitarian family type is highly correlated with some contemporaneous variables relating to gender discrimination: the data show that women living in countries that have been categorized as endogamous communitarian (1) do not have the same rights over divorce as men, (2) do not enjoy the same inheritance rights as widowers and sons, and (3) are significantly less likely to have an account at a financial organization (correlation coefficients are 0.53, 0.62, and -0.57 respectively, all significant at the 1% level). Figure 2 displays correlation networks of personal norms (top) and of social norms (bottom). Any edge displayed represents a statistically significant correlation coefficient at the 95% level. For simplicity, we do not distinguish positive from negative correlations in the visualization. Nodes starting with SUR are variables derived from survey data, whereas experimental data have labels starting with LEXP (lab) or FEXP (field). For a detailed overview of all variable labels used in Section 5, see table A1 in the Appendix.

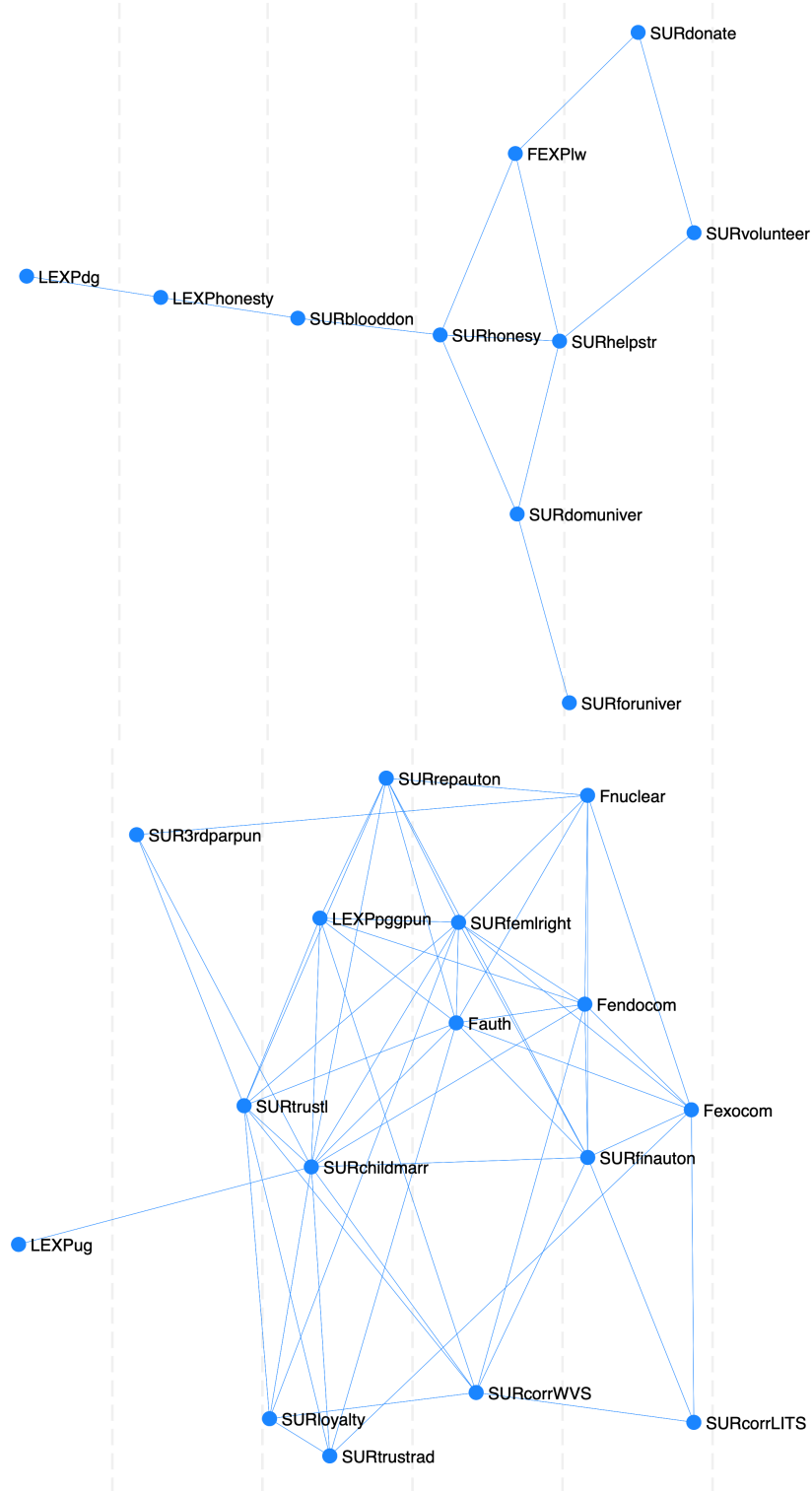
Figure 2 shows that social norms (bottom) have a higher density (0.46) in their correlation network than personal norms (0.26). This suggests that personal norms do not as often correlate among each other, whereas social norms tend to be more tightly interwoven. The top figure shows that survey-based variables regarding both honesty as well as helping strangers assume a central position within the network—which can also

be identified as being part of multiple motifs—which connect different measures regarding altruism and universalism. While helping a stranger (*SURhelpstr*) is part of two triplets and one quadruplet,<sup>19</sup> the results of the field experiment “Lost Wallet” (*FEXPlw*) form one quadruplet and one triplet. This is noteworthy because the included lab experimental variables (*LEXPdg* and *LEXPhonesty*) are much more marginal, not belonging to any motif and only being sparsely connected to the rest of the network. Since there is no significant correlation between blood donations (*SURblooddon*) and declared altruism (*SURdonate*, *SURvolunteer*, *SURhelpstr*) there are no edges connecting the respective nodes.

The bottom figure refers to social norms. In this correlation network, the most central node is a measure of historical family types, namely the authoritarian family (*Fauth*, newly-wed sons stay with their parents and brothers are treated unequally in their inheritance). It is part of multiple motifs that, for example, tie this long-term informal institution to contemporaneous family institutions (e.g., female legal rights, *SURfemright*). While the 2-stage public goods game (with punishment, *LEXPpggpun*) assumes a relatively central position within the network, this is not the case for results of the Ultimatum Game (*LEXPug*), which remains marginal and sparsely connected to other proxies of social norms.

---

<sup>19</sup> *SURhelpstr*, *SURvolunteer* and *SURdonate* are all variables belonging to the same survey, so we assume they have a higher ex-ante probability of correlating due to a common sample structure.

Figure 2: Correlation networks among personal norms (top) and social norms (bottom)<sup>20</sup>

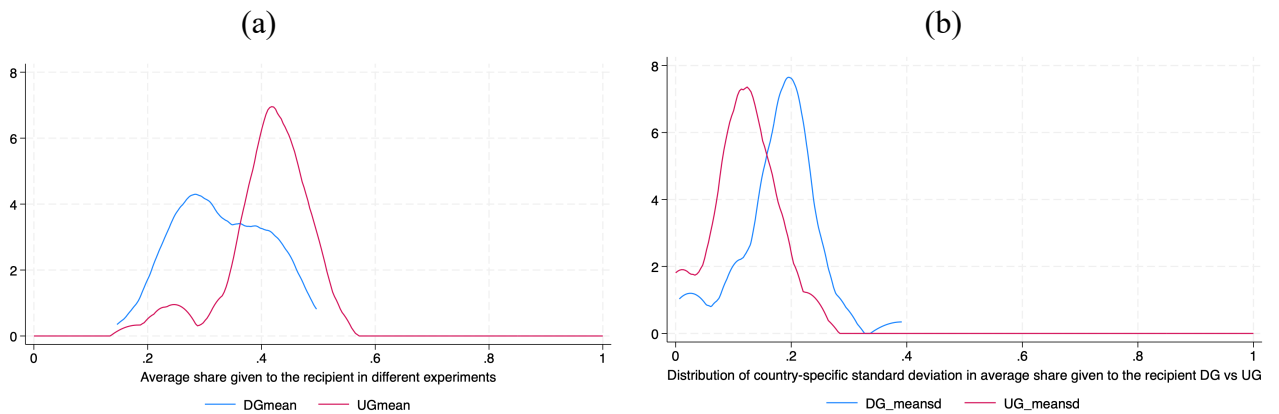
Source: elaboration by the authors based on Informal Institutions Data

<sup>20</sup> We use the STATA command *ymbionet* (von Jacobi and Joyez 2025) for all network visualizations.

### 5.3. Correlation Patterns Between Personal and Social Norms

As a first step, we take up the argument by Henrich (see Section 4.3), which claims that personal norms (i.e., type 2 institutions) have co-evolved with social norms (i.e., type 3 institutions) and tend to re-enforce each other. We first focus on norms that deal with the sharing of some pie. We compare the amount offered in the Ultimatum Game (UG, a type 3 institution) with the amount offered in the Dictator Game (DG, a type 2 institution). We suggest analyzing the differences in the offers made in the two games: the offer made in the DG can be attributed to internalized personal norms whereas the proposer in the UG is likely to take the anticipated reaction of the responder into account when making her offer. If the social norm is completely internalized and has also become a personal norm, the offers made in the two games should be identical. If, however, there is not complete internalization, the mean offer in the UG is expected to be higher than that in the DG. The difference between the two offers would then contain information on the effect of (anticipated) sanctions by others. Cochard et al. 2021 already offers results of such comparison for 29 countries based on studies published until 2014. We updated their approach and thereby increased the number of included countries to 46. Figure 3 compares the distribution of the average share given in the DG and the UG in 46 countries for which we have data on both games.

Figure 3: Kdensity distributions of average share given to the recipient (panel a), and the standard deviation of such share (panel b) in a game activating personal norms (DG) and one activating social norms (UG)<sup>21</sup>



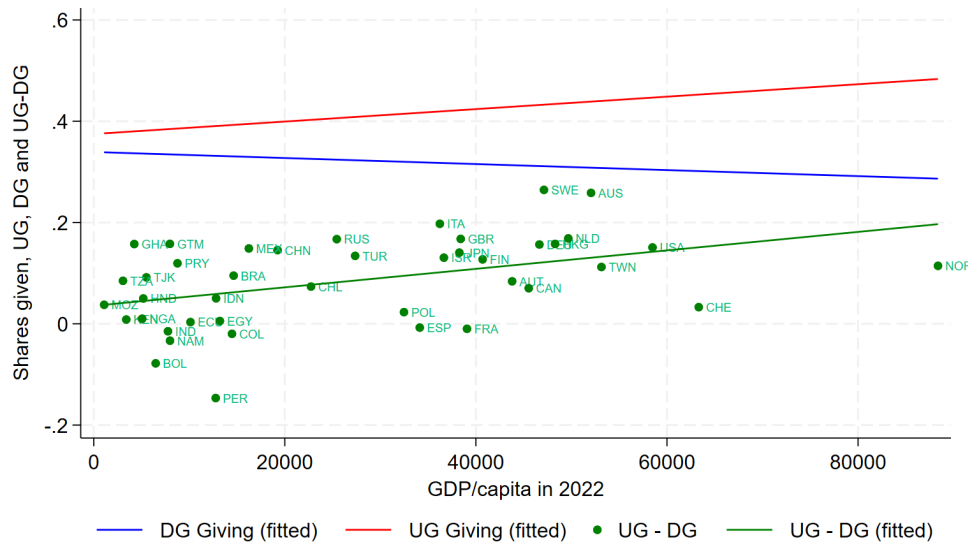
Source: elaboration by the authors based on Informal Institutions Data. Panel (a) is computed over 46 countries, panel (b) over 23 countries due to data availability.

<sup>21</sup> A kdensity distribution (kernel density distribution) is a smoothed estimate of a variable's probability distribution. A curve not "touching" the zero probability arises as each kdensity computes its own range, and one predominates in the visualization.

Figure 3 (panel a) shows that the amount of the pie shared with the recipient in a lab experiment tends to be higher when the design of the experiment foresees a potential sanction by the other side. When individual players only rely on their personal norm of what they think is fair, they share less. We also look at the dispersion of shares given. For 23 countries, we can reconstruct a standard deviation measure across experiments and treatments played within the same country. Panel b in figure 3 shows that the variance in shares offered to the recipient are similarly shaped, but systematically higher in the DG than in the UG. The potential ‘threat’ of being sanctioned increases the amount offered by individuals and reduces variability in such offering in similar ways across countries.

We next observe that subjects in poorer countries are making slightly higher offers in the DG than participants in richer countries (see Figure 4). The inverse is true regarding the UG: there, mean offers slightly increase in income. As expected, UG shares consistently tend to be higher than DG shares.<sup>22</sup> Subtracting the mean offer of the DG from that of the UG leads to very small differences in poorer countries and increasing differences in rising income. These findings can be interpreted with two lenses. On the one hand, it seems that fairness norms are internalized (*DG*) slightly more in poorer countries. In richer countries, on the other hand, the threat of being sanctioned seems to be more effective.

Figure 4: UG, DG shares given and their difference, plotted against GDP / capita<sup>23</sup>



Source: elaboration by the authors based on Informal Institutions Data

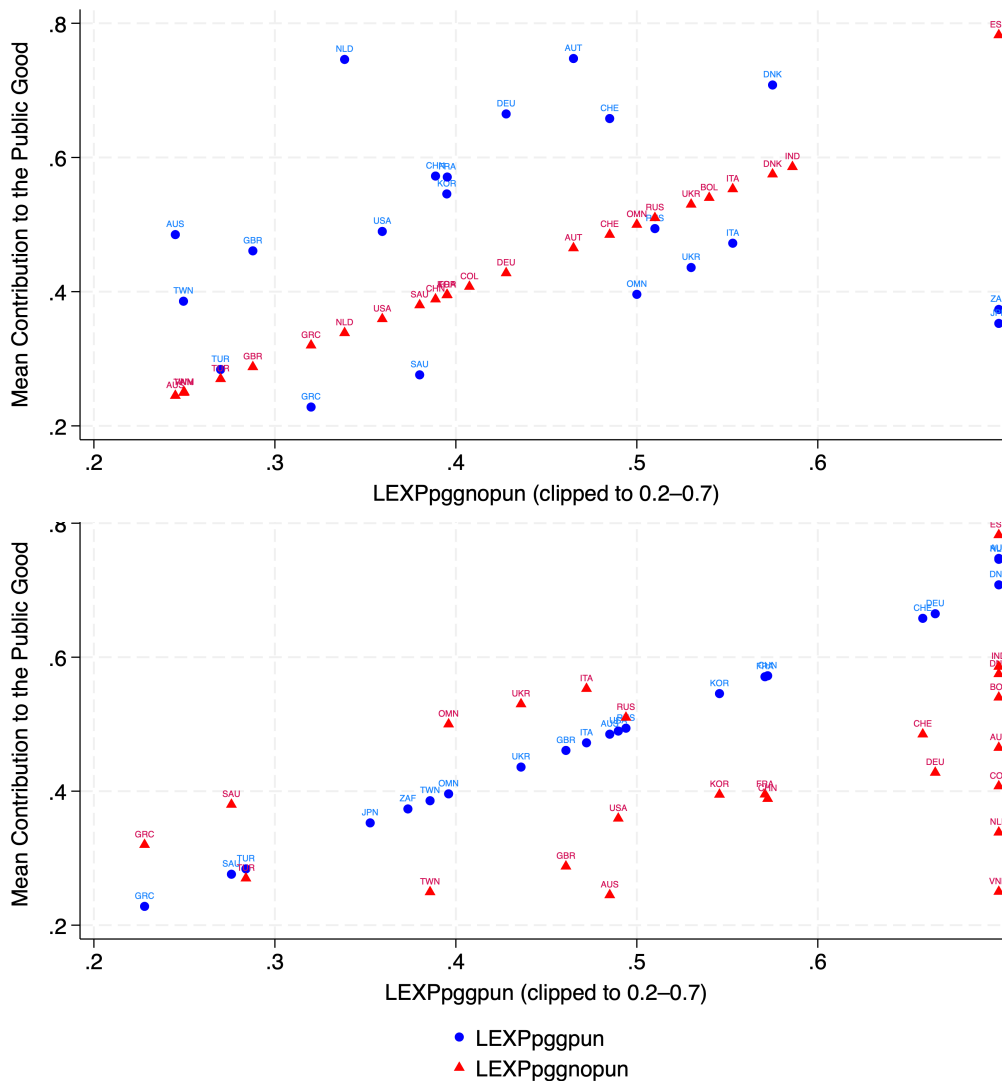
<sup>22</sup> Exceptions—DG shares more than 0.02 higher than UG shares—are Peru, Bolivia and Namibia.

<sup>23</sup> The figure only contains incentivized games. In a test in which we combine hypothetical and incentivized games, standard errors are almost the same 2.37 (all) vs. 2.34 (only incentivized) games.



In a similar fashion, we next focus on the willingness to cooperate for the common good. We use differences in mean contributions offered in public good games (PGG) and analyze different versions of the PGG. The offers made in the single stage PGG reflect the personal norms of the participants, whereas the offers made in the two-stage PGG (which includes the possibility of sanctioning others) reflect the (anticipated) effect of being sanctioned by others in case of non-compliance with the norm. The results are comparable with those just reported on the DG and the UG. Figure 5 compares the mean contributions to the public good when sanctioning is possible (*LEXPpggpun*) and when it is not (*LEXPpggnopun*).

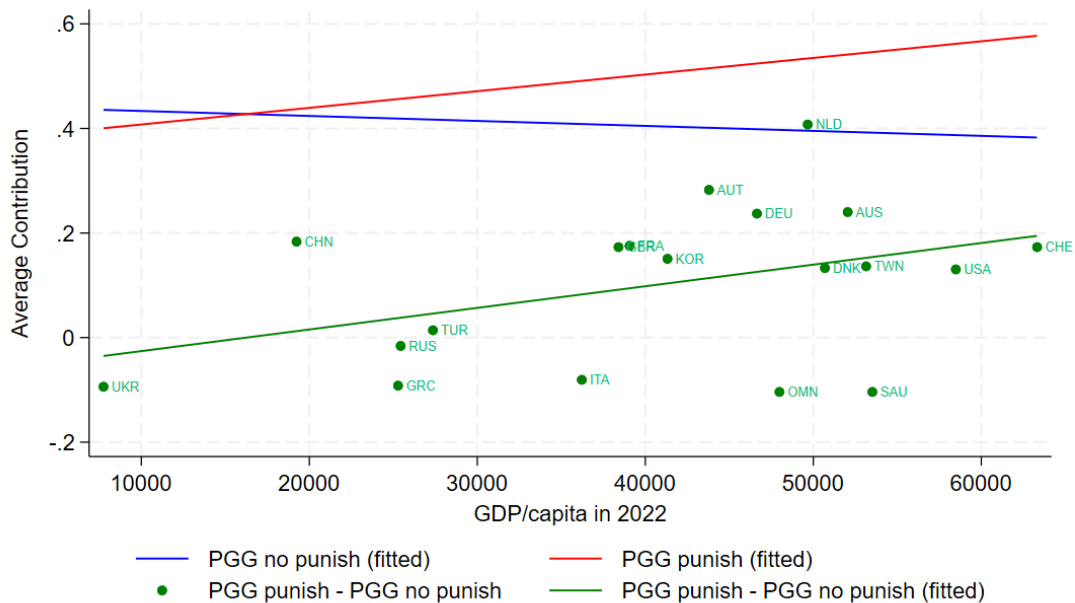
Figure 5: Country averages in mean contributions to simple and 2-stage PGGs



As expected, the introduction of a sanctioning stage raises mean contributions. Spain is an outlier for freely offered shares to the public good. When sorting on the contributions given under a sanction-threatening, most countries display higher contributions than when sanctions are absent. Exceptions to this rule include Russia, Italy, Ukraine, Oman, Saudi Arabia, and Greece.

We further observe that if the public goods game is played without punishment, offers in poorer countries are slightly higher than in richer ones (Figure 6)—in line with our previous findings comparing DG and UG results. Sharing also appears to be slightly more internalized in poorer countries. As soon as the punishment stage is added, the picture reverses. Offers in richer countries are now significantly higher than in poorer ones. These are correlations and no causation can be inferred from them.

Figure 6: Differences in mean contributions in a Public Goods Game (PGG) with and without punishment plotted against GDP per capita



Source: elaboration by the authors based on Informal Institutions Data.

#### 5.4. Networks of Informal Institutions

We now proceed with the computation of correlation networks making use of Moral Foundations Theory (*MFT*, see Graham and Haidt 2012; Haidt, 2012), which proposes that certain moral values emerge as evolutionary solutions across societies. We do so by grouping our variables thematically. In MFT, the moral foundations ‘care’, ‘fairness’, ‘authority’, ‘loyalty’ and ‘sanctity’ emerge to face challenges, such as the need to (1) secure the survival of vulnerable children, (2) be able to reap benefits from cooperation, (3) see the benefits in the establishment of social hierarchies, (4) be able to reap the benefits from forming networks, and (5) survive the exposure of pathogens, respectively. In other work, we confirm that moral foundations are in fact universally present across societies (von Jacobi, Khesali, and Voigt 2025).

We loosely group our variables belonging to all types (conventions, personal norms, social norms, bundles of norm) in line with four out of these five moral foundations.<sup>24</sup> Informal institutions are attributed to ‘care’ when their content enshrines care for others and avoidance of risk or harm. We group informal institutions under ‘fairness’ when they deal with equal sharing and defending justice. Any informal institution evoking some kind of hierarchy (men over women, elderly over youth, man over nature) is attributed to ‘authority’, and ‘loyalty’ groups all informal institutions dealing with horizontal network formation patterns. All candidate variables included for correlation network computation are listed in Table A2 in the Appendix.

Figure 7 shows correlation networks that only contain statistically significant correlations at the 95% level. Within the thematic area ‘care’ (panel a), we find that donating (as measured through the World Giving Index) locates rather centrally. The two bundles of institutions we include (long-term orientation and uncertainty avoidance proposed by Hofstede) form motifs with some personal norms. Honesty, measured through the field experiment “lost wallet” and the self-declared provision of volunteering (*SURvolunteer*), both form a quadruplet, each with long-term orientation (*Hlongt*), self-declared donations (*SURdonate*), and availability to help a stranger (*SURhelpstr*). Uncertainty avoidance (*Huncertav*) forms a triplet with volunteering (*SURvolunteer*) and with donating money (*SURdonate*). Our proxy for observed altruism, blood donations (*SURblooddon*), assumes a rather marginal location in the “care” network as it is connected to all other variables only through long-term orientation. We notice our dataset does not include social norms in the “care” area.

---

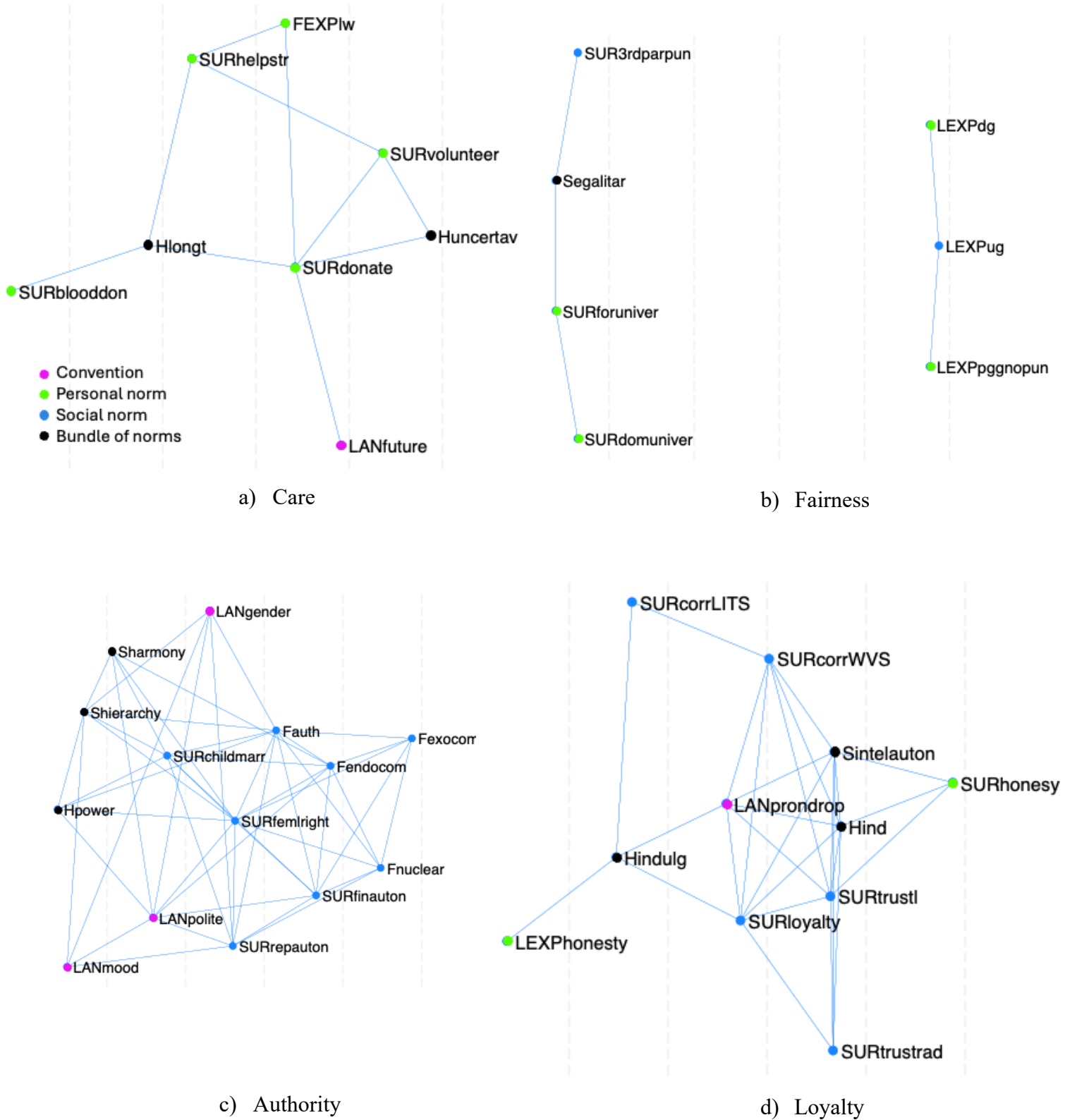
<sup>24</sup> Our dataset does not contain a sufficient number of proxies that could be attributed to the foundation ‘sanctity’.

Within the thematic area of ‘fairness’ (panel b), we find correlated personal and social norms. Instead of a single network, we find two non-connected components. Lab experiments form one cluster, while survey questions and bundles of norms form another. Schwartz’ measure of egalitarian value orientations (*Segalitar*) connects to our variables on third-party punishment (*SUR3rdparpun*) and on moral universalism involving foreigners (*SURforuniver*). Egalitarianism and moral universalism towards foreigners locate most centrally within the component made of survey questions. Within the component of lab experiments, the social norm measure (*LEXPug*) is more central than the two personal norms (*LEXPdg* and *LEXPpggnopun*).

The system of informal institutions belonging to the moral foundation of ‘authority’ (panel c) is the densest network, meaning that the number of computed ties between variables is high with respect to all possible ties. The correlation network is populated by conventions and by social norms. The variable measuring whether women enjoy fewer rights than men (*SURfemlright*) assumes the most central position in the network, showcasing that the position of women in society is crucial to how that society deals with authority. The authoritarian family type (*Fauth*) also locates very centrally. The bundles of norms on power distance (*Hpower*), and on hierarchy (*Shierarchy*) and harmony (*Sharmony*) proposed by Hofstede and Schwartz, respectively, are connected to the system, but not in a central way.

Within the system of informal institutions dealing with ‘loyalty’ (panel d), we find that proxies for individualism and collectivism emerge as central. Hofstede’s measures of individualism (*Hind*), and Schwartz’s measure on intellectual autonomy (*Sintelauton*) are the most central when paired with the grammatical feature of pronoun drop (*LANprondrop*). We find our proxy for particularistic loyalty (*SURloyalty*) locates well at the center of the network. Both our measures of honesty (*SURhonesty* and *LEXPhonesty*), both our measures of trust (*SURtrustl* and *SURtrustrad*) and both our measures of corruption (*SURcorrLITS* and *SURcorrWVS*) are also part of the ‘loyalty’ network. We find that our experimental measure of honesty only connects to the network through its correlation with indulgence (*Hindulg*).

Figure 7: Correlation networks among conventions, personal and social norms grouped by moral foundations (a) ‘care’, (b) fairness, (c) authority and (d) loyalty



Source: elaboration by the authors based on Informal Institutions Data.

### 5.5. Assessing the Centrality of Variables Within the Complete Correlation Network

In the preceding subsections, we analyzed correlations both between institutions of the same (e.g. personal norms or type 2 institutions) and different types. We now conclude the analysis with a correlation network that comprises statistically significant correlations for all variables across types (conventions, personal norms and social norms) and methodological approaches (language data, surveys, experimental data, and bundles of norms), presented in Figure 8. The Figure shows that lab experiments tend to be positioned on the outskirts of the network, hinting that these variables are less connected to the broader context of informal institutions. Interestingly, language data adopt a bridging position between the broader context and the measures based on lab experiments. Language may thus directly affect cognition as suggested in Mavisakalyan and Weber (2018) and in Wei et al. (2023). Our only variable capturing field experimental evidence is situated remarkably centrally. This suggests the “lost wallet” experiment is rather well suited for capturing multiple aspects of the social norms context in which individuals interact.

We now compute a series of network statistics to gather more precise, quantitatively estimated, measures of the location of each proxy within the system of informal institutions. Table 3 reports the ranks of our informal institutions measures according to different centrality measures—all computed on statistically significant correlation coefficients. Variables assume the role of nodes. *Unweighted degree centrality* counts the number of correlations of a variable, *weighted degree centrality* computes the weighted sum of correlation coefficients. *Closeness degree centrality* computes the sum of shortest paths (concatenated correlations) needed to connect a variable to every other variable in the network.<sup>25</sup> *Betweenness degree centrality* computes the percentage of shortest paths between any two pairs of variables on which another variable is located. It captures how often variables are likely to display a mediating role. Finally, *eigenvector degree centrality* is a recursive measure combining the centrality of a variable with the centrality of its neighboring (directly connected) variables. In table 3, the highest (1-5) and lowest (36-41) ranking variables are reported, Table A7 in the Appendix reports all results. Table 4 computes averages of the centrality measures grouped by the type of variable (experimental, survey, language data or bundles of institutions).

Measures elicited through lab experiments tend to rank lowest in centrality. Their average eigenvector degree centrality is 0.05 (and only 0.025 when excluding the 2-

---

<sup>25</sup> Closeness degree centrality captures the potential of a node to having high diffusion within the network if changed.

stage PGG which in our networks tends to be better connected to other informal institutions than other experimental variables). In contrast, the field experiment “Lost Wallet” displays an eigenvector degree centrality of 0.30. Survey variables on average have a centrality of 0.15, and language variables that capture slow-moving conventions on grammar have a centrality of 0.09. Bundles of institutions, which are derived from multiple survey questions, locate more centrally—expected in a correlation network, where their multiple components are likely to correlate with more of the other informal institutions of the network.

The result is robust to a restriction to statistically significant correlation coefficients with values  $>|0.40|$ .<sup>26</sup> Under such restriction, the centrality of survey and language variables decreases slightly, while the centrality of bundles of institutions remains rather stable. Centrality of experimental variables increases slightly, but only if the 2-stage PGG is included. In any case, the average centrality of this group of variables remains the lowest in the network (see Table 4).

---

<sup>26</sup> The corresponding network is included in the Appendix Section 5.

Table 3: Degree centrality of informal institutions within the overall correlation network, ranking highest (1-5) or lowest (37-41)

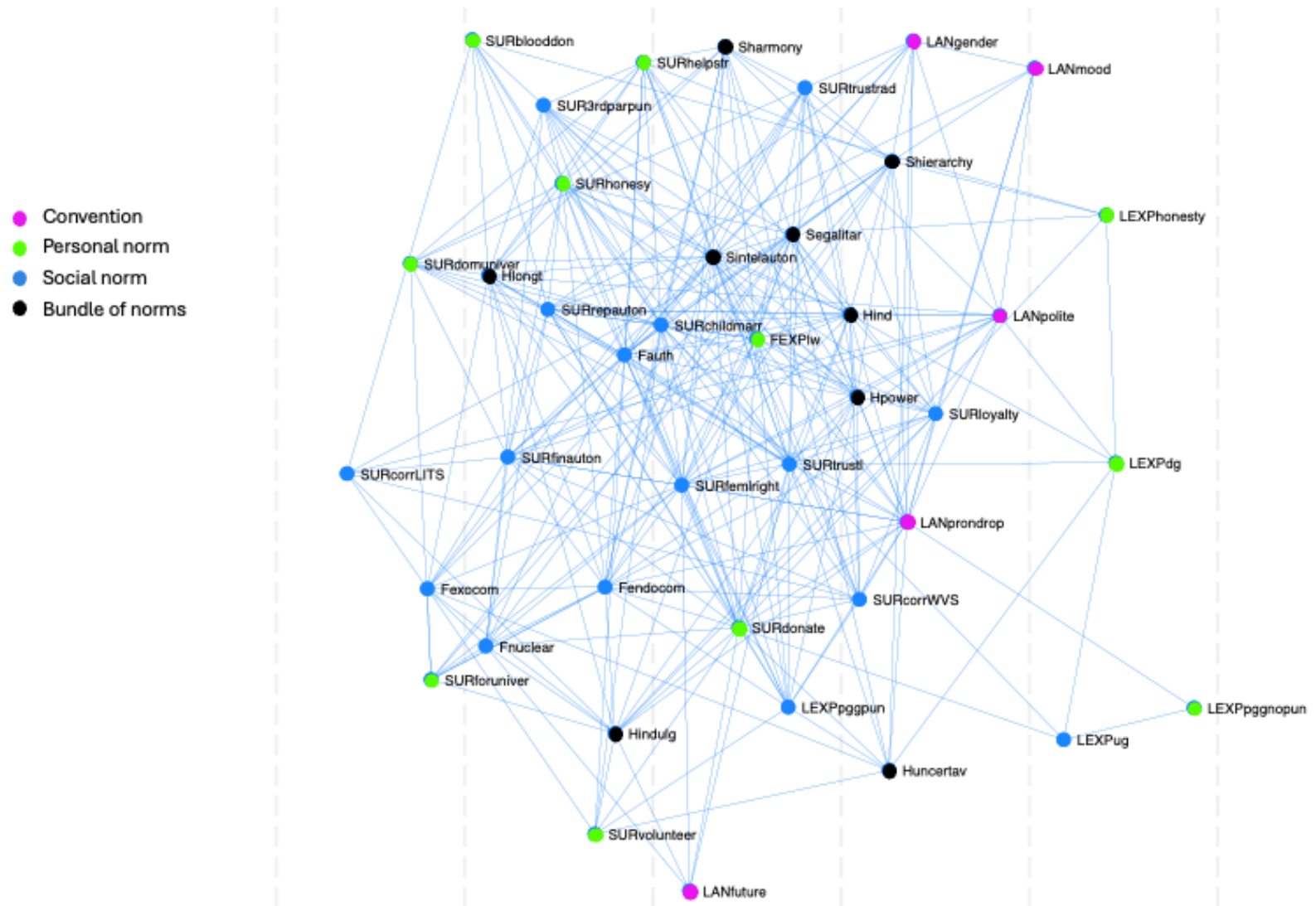
Rank	Unweighted Degree		Weighted Degree		Closeness Degree		Betweenness Degree		Eigenvector Degree	
1.	Sintelauton	0.625	Sintelauton	0.291	FEXPlw	0.342	SURloyalty	0.072	FEXPlw	0.302
2.	SURchildmarr	0.625	FEXPlw	0.287	Sintelauton	0.335	Sintelauton	0.065	Sintelauton	0.287
3.	Fauth	0.600	Hind	0.247	SURloyalty	0.318	FEXPlw	0.064	Hind	0.257
4.	SURfemlright	0.600	SURchildmarr	0.247	Hind	0.311	SURblooddon	0.059	SURchildmarr	0.237
5.	SURtrustl	0.575	SURtrustl	0.239	SURchildmarr	0.307	LEXPpggpun	0.056	SURtrustl	0.237
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
37.	LEXPdg	0.150	LEXPdg	0.056	LANmood	0.202	SURhelpstr	0.001	LANfuture	0.040
38.	LEXPhonesty	0.125	LEXPhonesty	0.052	LANfuture	0.201	LEXPpggnopun	0.001	LEXPdg	0.035
39.	LANfuture	0.125	LANfuture	0.044	Fnuclear	0.199	LANfuture	0.000	LEXPhonesty	0.029
40.	LEXPug	0.100	LEXPug	0.039	LEXPpggnopun	0.187	Fnuclear	0.000	LEXPug	0.023
41.	LEXPpggnopun	0.050	LEXPpggnopun	0.024	LEXPug	0.187	Sharmony	0.000	LEXPpggnopun	0.015

Source: elaboration by the authors based on Informal Institutions Data. For an explanation of variable labels please see table A1 in the Appendix.

The complete table (all variables ranking 1-41) is included in the Appendix, Table A7.



Figure 8: correlation network of all conventions, personal and social norms included in the Informal Institutions Dataset



Source: elaboration by the authors based on Informal Institutions Dataset. Only correlations with statistical significance of at least 95% retained.

Table 4: Eigenvector Degree Centrality of Lab experimental, Survey, Language data, and of bundles of institutions (Hofstede and Schwartz variables) with and without restriction by weight of correlation coefficients in the Complete Correlation Network

	Eigenvector Degree Centrality (no corr. size restriction)	Eigenvector Degree Centrality (with corr. size restriction)
LEXPall	0.0589	0.0618
LEXP_no_PGGpun	0.0253	0.0143
SURall	0.1475	0.1240
LANall	0.0944	0.0779
BUNDLESsall (H & S)	0.1685	0.1621

Source: elaboration by the authors based on Informal Institutions Data. LEXP\_no\_PGGpun computes centrality of all experimental variables excluding the results of the Public Goods Game with punishment—this is the experimental variable locating most centrally in our network.

## 5.6. Looking Back and Summing Up: Recommendations for Data Usage

After having analyzed a series of network properties over the previous sections, we now proceed with discussing the usefulness of various sources of information regarding informal institutions. A key advantage of our analysis comes from its ability to compare different proxies of the same informal institution. This allows us to identify some systematic issues related to their measurement.

(1) It does matter whether a norm is inferred in the lab or in the field. As previously alluded to, real-life context is naturally present in field experiments. We have argued throughout that social norms always display their effects within specific contexts and that lab experiments are explicitly designed to abstract from that context. This is, in a sense, one reason to rely on them in a cross-country setting. As the respective context is likely to differ widely across countries, lab experiments help to focus on one particular aspect, implying high comparability but uncertainty regarding external validity. Yet, to understand how social norms channel behavior in the real world, explicitly isolating from all context may not be the straightest path to enlightenment.<sup>27</sup>

---

<sup>27</sup> A series of studies have engaged with the external validity of lab experimental results, and there is evidence that behavior in the lab does mirror behavior in real life. Franzen & Pointner (2013) document this for fairness games and Rustagi et al. (2010) do so for conditional cooperation; these are just two examples. An important discussion regarding the comparative advantages and disadvantages of lab vs. field experiments has been popularized by Levitt and List (2007), who stress

(2) Experiments—whether in the lab or field—elicit behavior (although possibly in a highly artificial setting), whereas surveys elicit declarations about behavior. One way to make differences between the two elicitation methods plausible is to point to the social desirability bias, which appears here in the form of an interviewer demand effect. Of course, the experimenter demand effect is also well known, but the former is likely to be stronger as surveys are not incentivized while experiments are. Self-deception may also play a role: the surveyed simply deceive themselves to create a positive self-image. But the differences between survey and experimental results may also be useful in increasing our knowledge regarding social norms. Injunctive norms refer to how one ought to behave to ensure social approval, whereas descriptive norms describe how most others behave in a given situation (Cialdini et al. 1990). Surveys are more likely to prompting injunctive norms while experiments may be closer to descriptive ones. A promising trend relies on experimentally validated survey questions (see, for example, Falk et al. 2018).

(3) Both surveys and lab experiments often disregard possibly relevant context: for example, the World Giving Index asks whether the surveyed person has helped a stranger in the past month. Both “help” and “stranger” need to be delineated by every surveyed person, and it is by no means certain that this will be done in an identical fashion. In a small village, the likelihood of even meeting a stranger is significantly lower than in a large city. Then again, a stranger may be far easier to detect in a village than in a large city. “Helping” could mean anything between showing a stranger her way or offering him money to pay for a hotel or similar. Our correlation network analysis and the focus on centrality we proposed shows which survey or experimental variables are likely in doing a better job at properly capturing parts of the relevant context.

(4) We notice that granularity of measurement matters: to rely on a single variable for eliciting the presence of a general norm may not be sufficiently precise. Norms can vary across various dimensions, such as their content, intensity, scope, stability, and so on (see already Schwartz 1973). Honesty norms, for example, may specify under which exceptional conditions one is allowed to lie (e.g., in an experimental game vs. vis-à-vis your mother). To the degree that we are dealing with personal norms, the way in which the norm of honesty is made concrete is likely to be extremely heterogenous across individuals and contexts in which they are embedded. This can explain why we

---

the advantages of field over lab experiments. Camerer (2011) responded to their assertion, and stresses the potential complementarity between the two approaches.

sometimes detect unexpectedly low correlations between proxies meant to represent the same institution (e.g., altruism).<sup>28</sup>

Based on our findings, some recommendations follow. In collecting our data, we made an effort to bridge different measurement approaches to provide a more comprehensive information base for those interested in working with data on informal institutions. The correlation network analysis shows that despite some missing links (i.e. unexpected absence of significant correlations), there is multicollinearity between many of the measures included. While multicollinearity typically represents a problem to econometric estimations, our analysis has sought to highlight how much can be learned through engaging with it. When informal institutions are meant to enter a regression analysis as an independent variable, it may be preferable to select a single measure—possibly the one most central to the overall system of informal institutions. In this way, the variable is likely to capture a broader (partially latent) informal institutional factor.

When multiple variables are to be included in the same analysis, careful attention should be paid to their correlational structure not being too similar (i.e. it is preferable to select variables here placed on the outskirts of the networks). Where informal institutions are intended to enter the analysis as instrumental variables, we suggest to use language features, which are less endogenous than other measures. In our correlation networks, language features tend to locate on the outskirts, apart from ‘pronoun drop’ which correlates with many other variables proxying collectivism. Given that language features are more slow-moving than other measures, we suggest they are good candidates for instruments.

We further notice that the survey variable used to measure ‘particularistic loyalty’ has high potential for becoming an important measure for the informal institutional context. In our correlation network analysis, it exists very centrally with important ties to ‘trust’ ‘corruption’ and ‘honesty’. This variable is likely to be valuable in better understanding how informal institutions shape network formation patterns.

---

<sup>28</sup> Tables with all bivariate correlations are available in the [Online Material](https://www.dropbox.com/scl/fo/39dytoexs7bh75j0c601z/ALrcfTAnjximj9QVpxgCt-0?rlkey=wt3yie87kcbpso438wog3yqn8&dl=0) (https://www.dropbox.com/scl/fo/39dytoexs7bh75j0c601z/ALrcfTAnjximj9QVpxgCt-0?rlkey=wt3yie87kcbpso438wog3yqn8&dl=0).

## 6. Conclusions and Outlook

To our knowledge, the database on informal institutions presented in this paper is the most encompassing to date. We hope that even more data will become available in the future, both regarding currently missing country data points, and with variables proxying additional informal institutions.

Measurement of personal norms is particularly tricky, as non-compliance with a personal norm is not sanctioned by other observable actors, but by the non-complier herself. To separate personal norms from preferences or simple habits is, therefore, a serious challenge. Being able to ascertain the degree of guilt or loss of self-esteem that a person experiences when she does not comply with her own personal norms would be an important step towards making type 2 institutions measurable in a more precise way. This is definitely a desideratum for future research.

In our empirical section, first steps towards analyzing institutions as components of institutional systems were taken. To do so, we relied on correlation network analysis. The insights gained by analyzing entire systems of institutions can be highly policy relevant: if institutions have a firm place in a broader system of institutions, attempts to change single ones may fail because their deep embeddedness activates resilience of the system, such as shunning official legislation. Policies targeting single informal institutions without taking the entire systemic context into account could either fail outright or have unintended side effects on interdependent institutions.

## References

- Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, 87(4), 1115-1153.
- Acemoglu, D., & Robinson, J. A. (2025). Culture, institutions, and social equilibria: A framework. *Journal of Economic Literature*, 63(2), 637-692.
- Akerlof, G. A., & Kranton, R. E. (2010). Identity economics. In *Identity Economics*. Princeton University Press.
- Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4), 898-944.
- Amendolagine, V. & von Jacobi, N. (2023). What feeds on what? Networks of interdependencies between culture and institutions. *Economia Politica*, 1-42.
- Angerer, S., Glätzle-Rützler, D., Lergetporer, P., & Sutter, M. (2021). The effects of language on patience: An experimental replication study of the linguistic-savings hypothesis in Austria. *Journal of the Economic Science Association*, 7, 88-97.
- Banfield, E. C. (1958). *The moral basis of a backward society*. New York: The Free Press.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1), 122-142.
- Bhalotra, S., Clarke, D., Gomes, J., & Venkataramani, A. (2015). Maternal mortality and female life expectancy: The importance of gender inequality. *Unpublished manuscript*.
- Bisin, A., & Verdier, T. (2022). Advances in the economic theory of cultural transmission. *Annual Review of Economics*, 15.
- Bjørnskov, C. (2012). How does social trust affect economic growth?. *Southern Economic Journal*, 78(4), 1346-1368.
- Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6), 3531-3535.
- Camerer, C. (2011). The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List. Available at SSRN 1977749.
- Cappelen, A. W., Enke, B., & Tungodden, B. (2025). Universalism: global evidence. *American Economic Review*, 115(1), 43-76.
- Charities Aid Foundation (2022). World Giving Index 2022 – A global view of giving trends.
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental economics*, 14(1), 47-83.
- Chen, M. K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *American Economic Review*, 103(2), 690-731.
- Chen, J. I., He, T. S., & Riyanto, Y. E. (2019). The effect of language on economic behavior: Examining the causal link between future tense and time preference in the lab. *European Economic Review*, 120, 103307.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6), 1015.

- Cochard, F., Le Gallo, J., Georgantzis, N., & Tisserand, J. C. (2021). Social preferences across different populations: Meta-analyses on the ultimatum game and dictator game. *Journal of Behavioral and Experimental Economics*, 90, 101613.
- Cohn, A., Maréchal, M. A., Tannenbaum, D., & Zünd, C. L. (2019). Civic honesty around the globe. *Science*, 365(6448), 70-73.
- Davis, L. S., & Abdurazokzoda, F. (2016). Language, culture and institutions: Evidence from a new linguistic dataset. *Journal of Comparative Economics*, 44(3), 541-561.
- Davis, L. S., & Williamson, C. R. (2016). Culture and the regulation of entry. *Journal of Comparative Economics*, 44(4), 1055-1083.
- Delhey, J., Newton, K., & Welzel, C. (2011). How general is trust in “most people”? Solving the radius of trust problem. *American sociological review*, 76(5), 786-807.
- Dey AK, Gel YR, Poor HV. (2019). What network motifs tell us about resilience and reliability of complex networks. *Proceedings of the National Academy of Science. USA.*, 116(39):19368-19373.
- Elpers, N., Jensen, G., & Holmes, K. J. (2022). Does grammatical gender affect object concepts? Registered replication of Phillips and Boroditsky (2003). *Journal of Memory and Language*, 127, 104357.
- Engel, C. (2011). Dictator games: A meta study. *Experimental economics*, 14(4), 583-610.
- Enke, B. (2019). Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics*, 134(2), 953-1019.
- Enke, B. (2023). Market exposure and human morality. *Nature Human Behaviour*, 7(1), 134-141.
- Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 76(2), 026107.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4), 1645-1692.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137-140.
- Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism—experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1, 615-691.
- Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3), 525-547.
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3), 347-369.
- Franzen, A., & Pointner, S. (2013). The external validity of giving in the dictator game: A field experiment using the misdirected letter technique. *Experimental Economics*, 16, 155-169.
- Gelfand, M.J., Raver, J.L., Nishii, L., Leslie, L.M., Lun, J., Lim, B. C., Yamaguchi, S. (2011). Differences between tight and loose cultures: a33-nation study. *Science*. 332 (6033), 1100–1104.
- Givati, Y., & Troiano, U. (2012). Law, economics, and culture: Theory of mandated benefits and evidence from maternity leave policies. *The Journal of Law and Economics*, 55(2), 339-364.
- Gorodnichenko, Yuriy & Gerard Roland (2011), Which dimensions of culture matter for long-run growth?, *The American Economic Review*, 101(3), pp. 492-498.

- Gorodnichenko, Yuriy & Gerard Roland (2017), 'Culture, Institutions, and the Wealth of Nations', *Review of Economics and Statistics*, 99(3), pp. 402-416.
- Gorodnichenko, Yuriy & Gerard Roland (2021), Culture, institutions and democratization, *Public Choice*, 187(1-2), 165-195.
- Graf, C., Pondorfer, A., Schulz, J. (2024). Culture and gender differences in honesty [Unpublished draft].
- Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach.
- Greif, A., & Mokyr, J. (2017). Cognitive rules, institutions, and economic growth: Douglass North and beyond. *Journal of Institutional Economics*, 13(1), 25-52.
- Gründler, K., & Köllner, S. (2020). Culture, diversity, and the welfare state. *Journal of Comparative Economics*, 48(4), 913-932.
- Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes?. *Journal of Economic perspectives*, 20(2), 23-48.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, 3(4), 367-388.
- Gutmann, J., Lewczuk-Czerwińska, A., Lewkowicz, J., & Voigt, S. (2025). *Culture and constitutional compliance* (No. 85). ILE Working Paper Series
- Gutmann, J., & Voigt, S. (2022). Testing Todd: family types and development. *Journal of Institutional Economics*, 18(1), 101-118.
- Hahm, H., Hilpert, D., & König, T. (2023). Divided by Europe: affective polarisation in the context of European elections. *West European Politics*, 46(4), 705-731.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.
- Henrich, J. (2016). The secret of our success. In *The Secret of Our Success*. Princeton University Press.
- Henrich, J. (2020). *The WEIRDest people in the world: How the West became psychologically peculiar and particularly prosperous*. Penguin UK.
- Henrich, J., Boyd, R., & Richerson, P. J. (2012). The puzzle of monogamous marriage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1589), 657-669.
- Hill, J. H., & Mannheim, B. (1992). Language and world view. *Annual review of anthropology*, 381-406.
- Hippel von, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and brain sciences*, 34(1), 1-16.
- Hofstede, G. (1980). *Culture's Consequences – International Differences in Work-Related Values*. Sage: Beverly Hills et al.
- Hofstede, G. (2003). *Cultures and Organizations. Software of the Mind*. London: Profile Books.
- Horvath, S. (2011). *Weighted network analysis: applications in genomics and systems biology*. Springer Science & Business Media.
- House, R. J. (2004). An overview of the GLOBE research program. *Transformative organizations: A global perspective*, 483(2003).
- Hox, J. (2002). *Multilevel Analysis: Techniques and Applications*. Taylor & Francis.



- Jacobi von, N. (2018). Institutional interconnections: understanding symbiotic relationships. *Journal of Institutional Economics*, 14(5), 853-876.
- Jacobi von, N., & Joyez, C. (2025). SYMBIONET: Stata module to build a symbiotic (or symmetric) correlation network from a list of variables.
- Jacobi von, N., Khesali, M., & Voigt, S. (2025). *Did they live happily ever after? Moral foundation narratives in folklore*. Mimeo, Universities of Hamburg and Trento.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of economic psychology*, 32(5), 865-889.
- Kashima, E. S., & Kashima, Y. (1998). Culture and language: The case of cultural dimensions and personal pronoun use. *Journal of cross-cultural psychology*, 29(3), 461-486.
- Kinzler, K. D. (2021). Language as a social cue. *Annual review of psychology*, 72, 241-264.
- Kinzler, K. D., & Dautel, J. B. (2012). Children's essentialist reasoning about language and race. *Developmental science*, 15(1), 131-138.
- Knack, S., & Zak, P. J. (2003). Building trust: public policy, interpersonal trust, and economic development. *Supreme court economic review*, 10, 91-107.
- Kovacic, M., Costantini, F., & Bernhofer, J. (2018). Risk attitudes, investment behavior and linguistic variation. *University Ca'Foscari of Venice, Dept. of Economics Research Paper Series*, (34).
- Larney, A., Rotella, A., & Barclay, P. (2019). Stake size effects in ultimatum game and dictator game offers: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 151, 61-72.
- Ledyard, O. (1995). Public goods: some experimental results.
- Le Play, F. (1895). *L'organisation de la famille selon le vrai modèle signalé par l'histoire de toutes les races et de tous les temps*. Mame.
- Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic perspectives*, 21(2), 153-174.
- Lewis, D. (1969). *Convention: A philosophical study*. John Wiley & Sons.
- Licht, A. N., Goldschmidt, C., & Schwartz, S. H. (2007). Culture rules: The foundations of the rule of law and other norms of governance. *Journal of comparative economics*, 35(4), 659-688.
- Lowes, S. (2020). Kinship structure & women: Evidence from economics. *Daedalus*, 149(1), 119-133.
- Mavisakalyan, A., & Weber, C. (2018). Linguistic structures and economic outcomes. *Journal of Economic Surveys*, 32(3), 916-939.
- Moscona, J., Nunn, N., & Robinson, J. A. (2020). Segmentary Lineage Organization and Conflict in Sub-Saharan Africa. *Econometrica*, 88(5), 1999-2036.
- North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.
- Nunn, N. (2020). The historical roots of economic development. *Science*, 367(6485), eaaz9986.
- Onnela, J. P., Saramäki, J., Kertész, J., & Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 71(6), 065103.
- Organization for Economic Cooperation and Development (2019). Gender, Institutions and Development Database. <https://stats.oecd.org/Index.aspx?DataSetCode=GIADB2019>

- Ostrom, E. (1996), Incentives, Rules of the Game, and Development, M. Bruno (ed.), *Annual World Bank Conference on Development Economics*, Washington, D.C.: The World Bank, 207-34.
- Oyserman, D. (2017). Culture three ways: Culture and subcultures within countries. *Annual review of psychology*, 68, 435-463.
- Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language*. New York: Harper.
- Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416), 427-430.
- Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature communications*, 5(1), 1-12.
- Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in cognitive sciences*, 13(10), 439-446.
- Rentfrow, P. J. (2010). Statewide differences in personality: toward a psychological geography of the United States. *American Psychologist*, 65(6), 548.
- Richards, J.M., Gottfredson, D.C. and Gottfredson, G.D. (1990). Units of analysis and item statistics for environmental assessment scales. In: *Current Psychology* 9 (4), pp. 407-413. url: <http://dx.doi.org/10.1007/BF02687196>.
- Roberts, S. G., Winters, J., & Chen, K. (2015). Future tense and economic decisions: Controlling for cultural evolution. *PloS one*, 10(7), e0132145.
- Robinson, J. A. (2013). Measuring institutions in the Trobriand Islands: a comment on Voigt's paper. *Journal of Institutional Economics*, 9(1), 27-29.
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review* 15, pp. 351-357.
- Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of economic growth*, 9(2), 131-165.
- Rustagi, D., Engel, S., & Kosfeld, M. (2010). Conditional cooperation and costly monitoring explain success in forest commons management. *science*, 330(6006), 961-965.
- Santacreu-Vasut, E., Drori, I., Manos, R., & Shoham, A. (2014). Do Female/Male Distinctions in Language Influence Microfinance Outreach to Women?. In *Academy of Management Proceedings* (Vol. 2014, No. 1, p. 16074). Briarcliff Manor, NY 10510: Academy of Management.
- Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, 366(6466), eaau5141.
- Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of experimental social psychology*, 9(4), 349-364.
- Schwartz, S. H. (2004). Mapping and interpreting cultural differences around the world. In: Vinken, H., Soeters, J., Ester, P. (Eds.), *Comparing cultures: Dimensions of Culture in a Comparative Perspective*. The Netherlands, Leiden, Brill, pp. 43-73.
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In Gentner and Goldin-Meadow (eds.) *Language in mind: Advances in the study of language and thought*, chapter 7. MIT Press.
- Smith, P. B., & Bond, M. H. (2019). Cultures and persons: Characterizing national and other types of cultural difference can also aid our understanding and prediction of individual variability. *Frontiers in Psychology*, 10, 2689.

- Spolaore, E., & Wacziarg, R. (2013). How deep are the roots of economic development?. *Journal of economic literature*, 51(2), 325-369.
- Stouffer, S. A., & Toby, J. (1951). Role conflict and personality. *American Journal of Sociology*, 56(5), 395-406.
- Sugden, R. (1986). *The Economics of Rights, Co-operation and Welfare*. Xxx
- Tabellini, Guido (2008a), Institutions and culture, *Journal of the European Economic Association*, 6(2-3), 255-294.
- Tabellini, Guido (2008b), The scope of cooperation: Values and incentives, *The Quarterly Journal of Economics*, 123(3), 905-950.
- Tabellini, Guido (2010), Culture and institutions: Economic development in the regions of Europe, *Journal of the European Economic Association*, 8(4), 677-716.
- Thöni, C. (2019). Cross-cultural behavioral experiments: potential and challenges. Chapter 18 in: *Handbook of research methods and applications in experimental economics*, 349.
- Todd, E. (1985). *The explanation of ideology: Family structure and social systems*. Oxford: Blackwell.
- Trompenaars, F., & Hampden-Turner, C. (2020). *Riding the waves of culture: Understanding diversity in global business*. Hachette UK.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293.
- Voigt, S. (2013), 'How (Not) to measure institutions. *Journal of Institutional Economics*', 9(1): 1-26.
- Voigt, S. (2018). How to measure informal institutions. *Journal of Institutional Economics*, 14(1), 1-22.
- Voigt, S. (2019). *Institutional economics: An introduction*. Cambridge University Press.
- Wei, X., Adamson, H., Schwendemann, M., Goucha, T., Friederici, A. D., & Anwender, A. (2023). Native language differences in the structural connectome of the human brain. *Neuroimage*, 270, 119955.
- Whitt, S., & Wilson, R. K. (2007). The dictator game, fairness and ethnicity in postwar Bosnia. *American journal of political science*, 51(3), 655-668.
- Williamson, C. (2009). Informal institutions rule: institutional arrangements and economic performance. *Public Choice* 139(3/4):371-387.
- World Health Organization. (2022). Global status report on blood safety and availability 2021.

