

How to measure informal institutions

STEFAN VOIGT*

Institute of Law and Economics, University of Hamburg, Hamburg, Germany

Abstract. There have been repeated claims that informal institutions play an important or even crucial role in economic development. Yet measuring informal institutions entails serious challenges. This paper identifies the difficulties in measuring informal institutions, critically discusses some attempts hitherto taken and encourages the more intensive use of experiments to measure them.

... it is much easier to describe and be precise about the formal rules that societies devise than to describe and be precise about the informal ways by which human beings have structured human interaction. But although they defy, for the most part, neat specification and it is extremely difficult to develop unambiguous tests of their significance, they are important.

(North, 1990: 36)

1. Introduction

Many colonizers, imperial states, modernizers and reformers have failed in their efforts to substantially change the life of entire societies in the way they had intended, simply by changing formal legislation.¹ It has been conjectured (for instance by North, 1990: in particular, Chapter 5) that a common cause of this failure was the disregard of conventions, norms, mores and traditions commonly followed by members of these societies. Formulated differently: the informal institutions prevalent in a society might constitute a binding constraint on attempts to reform a society's formal institutions.

*Email: stefan.voigt@uni-hamburg.de

¹ Many examples of failed attempts to impose institutions from above could be given. Just think of the history of land rights in many African states. First, colonizers tried to impose institutions relying on titling as well as conflict resolution, which was administered by representatives of central government. After independence, newly sovereign governments – often with the support of international financial organizations – kept on pushing in the same direction. Yet to this day, informal institutions still play a very important role regarding the use of land in many African countries (see the introduction to the special issue of *World Development* by Sikor and Müller (2009) and the paper by Peters (2009), who approaches the issue as an anthropologist. Ensminger (1997) describes why many attempts to impose external institutions have failed). Allina-Pisano (2007) provides a very detailed analysis of the post-Soviet privatization of property rights regarding land in the Black Earth, i.e. Eastern Ukraine and the parts of Russia bordering on Ukraine. Her account pays particular attention to the political economy aspects of privatization.

It has also been conjectured that the divergence in development within countries endowed with a homogeneous set of formal institutions might be attributed to differences in informal institutions. Putnam (1993) famously made this case regarding the differences in informal institutions between Northern and Southern Italy. A similar phenomenon has been identified concerning many African countries, where the effects of formal institutions can be very limited beyond the nation's capital (see, for instance, Herbst, 2014).

These are only two examples of the ramifications arising from a general claim that is heard ever more frequently among economists: it is not just formal institutions that matter, informal ones matter too. When economists claim that something matters, they usually have income and economic development in mind. Some are even ready to argue that informal institutions are more important than formal ones.² To test these conjectures empirically, informal institutions need to be measurable. While the measurement of institutions in general lags behind the theoretical advances in institutional economics, the measurement of informal institutions definitely constitutes a weak spot in institutional economics and might even be called institutional economics' most serious challenge.

This paper presents some reasons why it is so difficult to measure informal institutions, critically discusses some proxies sometimes found in the empirical literature and finally proposes to make more use of experiments to ascertain informal institutions. It builds upon ideas developed by Voigt (2013) that focus on the potential pitfalls in measuring institutions more generally.

The rest of the paper is organized as follows. In [Section 2](#), I briefly define the central terms, including in particular, of course, 'institutions'. [Section 3](#) contains some of the arguments that demonstrate the potential relevance of informal institutions. More specifically, I distinguish between four important channels through which informal institutions might have an impact on economic outcomes. Some specific challenges that need to be resolved when trying to measure informal institutions are listed in [Section 4](#). [Section 5](#) critically discusses the most frequently used indicators of informal institutions. My own proposals on how to achieve improved measures of informal institutions are developed in [Sections 6](#) and [7](#). Possible lessons for research as well as possible next steps are sketched in [Section 8](#).

2. How to delineate institutions?

Institutions can be defined as commonly known rules used to structure recurrent interaction situations that are endowed with a sanctioning mechanism. North (1990) distinguishes between formal and informal institutions, using the rule component as the criterion. Since North bases his distinction on the rule

² Echoing the claim that 'institutions rule' made by Rodrik et al. (2004), Claudia Williamson (2009) claims that 'informal institutions rule'.

Table 1. Types of internal and external institutions

Rule	Kind of enforcement	Type of institution	Example
1. Convention	Self-enforcement	Type 1, internal	Grammatical rules of language
2. Ethical rule	Self-commitment	Type 2, internal	Ten Commandments, categorical imperative
3. Custom	Informal societal sanctioning	Type 3, internal	Norms of etiquette
4. Formal private rule	Organized private enforcement	Type 4, internal	Rules created by merchants (e.g. arbitration courts)
5. State law	Organized state enforcement	External	Civil and criminal law

component, I 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 responsible for sanctioning when a rule has been renegeged upon. If the state sanctions rule-breaking, the enforcement is external to society and I propose to call the institution ‘external’. If rule-breaking is sanctioned by members of society, I propose to call the institution ‘internal’. One can also think of this as ‘public’ *versus* ‘private’ sanctioning. Table 1 contains a proposal for a more fine-grained taxonomy of internal institutions that focuses on who does the sanctioning.³

In type 1 institutions, an individual cannot make herself better off by unilaterally breaking the rule. Type 2 institutions rely on individual self-commitment and hence, are not the consequence of individuals who interact. The examples named in Table 1 help to clarify the difference between type 1 and type 2 institutions: when I aim to be understood by others, I cannot make myself better off by renegeging on the grammatical rules of a language. These rules are, hence, self-enforcing. If I am the last person to leave after a long day on a lonely beach, I might be tempted to leave my garbage right where I put it. If my ethics have been influenced by Kant, I might, however, ask if my behaviour is universalizable and take the garbage to the next trash can. If I were to renege on the rule that those who produce garbage are responsible for putting it away, I would probably feel bad, which is why I prefer to comply with it. In type 4 institutions, an organization is assigned the task to sanction rule breakers. This could, for instance, be private courts, as offered by many chambers of commerce.

This is why type 3 institutions promise to be of most interest here: rules that members of a certain group are expected to comply with. If there is non-compliance, sanctions will be implemented by other private individuals. If there is a generally accepted norm of queuing and a single person jumps the queue while the others have been patiently waiting, these people face a free-rider problem:

³ This taxonomy was first proposed in Kiwit and Voigt (1995). An abridged version in English is contained in Voigt and Kiwit (1998).

all of them want to see the queue jumper sanctioned. Yet sanctioning is costly (risking an evil eye or even being hit) so everybody hopes that someone else might volunteer to sanction the behaviour. Compared with the other three institutions, sanctions seem to be most unlikely because of the free-rider problem involved – and measures for such institutions are most interesting. This is why the main focus of this paper is on how to measure type 3 institutions.

Logically, one can distinguish four kinds of relationships between internal and external institutions: they can be (1) neutral, (2) complementary, (3) substitutive and (4) conflicting. The last category is of central interest here as it could hinder economic development. If institutions are conflicting, compliance with an external (internal) institution implies renegeing upon an internal (external) one.⁴ Viewed from a functional perspective, institutions reduce uncertainty in an otherwise very uncertain world (North, 1990: 6). By reducing uncertainty, institutions enable individuals to reduce the contingencies they need to consider when making plans for the future and acting today. But if internal and external institutions are in conflict, this can lead to increased uncertainty. As a consequence, transaction costs are higher than in situations where external and internal institutions are aligned, which, in turn, implies that there will be fewer transactions, less division of labour and – as a consequence – a lower standard of living.

Looked at more closely, the increased transaction costs due to conflicting institutions can be unbundled into a number of components. We could consider the first component to be the cost of familiarizing oneself with more than one set of institutions. These costs might not be tangible for individuals ‘born’ into the respective societies but might be very tangible for outsiders such as foreign companies. A possible consequence is that investments promise lower returns, and less foreign direct investment might be a consequence.

The coexistence of conflicting institutions implies that if an individual behaves in accordance with one institution, he simultaneously reneges on the other. To escape sanctions, it might be necessary to execute transactions in a clandestine way that will lead to sub-optimal firm size, the omission of advertisement and marketing campaigns, the impossibility of financing certain projects externally and so on. The same holds, of course, true for individuals. If I want to drink alcohol in a country in which its acquisition is prohibited by external institutions, I have to search for outlets very carefully, which will drive up transaction costs.

This is a second component of the overall transaction costs.

Suppose that the representatives of the state are keen on monopolizing the supply of institutions. This assumption seems well justified considering the

⁴ It is, of course, also possible that internal institutions conflict with each other. In the Arab world, for example, there is the widespread use of *wasta*, implying that relatives are expected to be privileged in hiring and promotion decisions. This is likely to be in conflict with institutions favouring meritocracy and the rule of law more generally (one of the few papers on *wasta* is Barnett et al. 2013).

Table 2. Combining types of rules with types of sanctions

	Internal	External
Informal	Societal norms	Vague legal concepts
Formal	Private arbitration	Most legislation

attempts of colonizers and revolutionaries alike to change entire societies in a short time frame. Some people who are supposed to implement the newly passed legislation might perceive the legitimacy of newly established external institutions as low and might continue to behave in accordance with time-honoured internal ones. Think, for instance, of policemen in a newly colonized country. If institutions exist that enable individuals to structure their interactions without reliance on the state, representatives of the state will have to spend more resources on implementing their system. In part, these costs will be made up of additional monitoring costs that are incurred to ensure that alternative institutions are not used. If there were no conflicting institutions around, these resources could be put to a more productive use.

Combining the distinction of informal *versus* formal institutions (i.e. referring to the rule component) with those of internal *versus* external institutions (i.e. referring to the sanctioning component) leads to a 2x2 matrix, depicted in [Table 2](#).

All four cells are more than theoretical possibilities and can be filled with well-known real-world examples as in [Table 2](#). Here, we are primarily concerned with institutions that rely on informal rules and whose non-compliance is sanctioned by members internal to society. I will focus, in other words, on the cell in the northwest corner of [Table 2](#), societal norms. These can range from rules on when and how to queue, to rules on how to cooperate in situations in which a public good is to be provided privately and collective action is necessary. Private arbitration organizations frequently refer to very formalized rules but do not rely on the state to enforce their sanctions. The respective institutions were called type 4 in [Table 1](#).

Many laws that are enforced by the state and are, hence, external according to the taxonomy here proposed consist of substantial amounts of informality. ‘Fair price’, ‘best interests of the child’ or simply ‘public interest’ are examples for such vague legal terms. On the other hand, most legislation is based on highly formal rules.

If we look at the taxonomy proposed in [Table 1](#), informal internal institutions correspond to type 3 internal institutions. Because the distinction regarding who is to do the sanctioning in case of non-compliance with the rule seems to be more readily ascertainable than whether a rule is formal or not, I propose to continue referring to the institutions being discussed in this paper as internal institutions, rather than informal ones.

3. Why should internal institutions matter?

This is not a theory paper. Its main focus is on examining why it is difficult to measure internal institutions, and explore possible ways to measure them. Categorizing certain interaction situations in which internal institutions are relevant might help to structure our endeavour. I propose to distinguish four stylized interaction situations in which internal institutions are expected to affect the concrete outcome. The first stylized situation refers to bilateral interactions between private actors. These can be individuals, but can also be firms. The second interaction situation refers to multilateral interactions, also between private actors. The third situation examines coordinating devices that allow a multitude of private actors to control the representatives of the state. The last situation refers to institutions that enable members of government to coordinate their behaviour.

- (1) In the first stylized interaction situation, private individuals are interested in coordinating their behaviour, and internal institutions might help them to be successful in their endeavour. Think of a one-time business transaction that takes place sequentially. Party 1 first delivers the good, party 2 promises to pay a couple of weeks down the road. Absent any institutions, no matter whether external or internal, economists predict that party 2 will not pay because it already enjoys the good and would prefer not to pay over paying for the good. If we observe such transactions taking place in the absence of any external enforcement in a one-shot game, we have reason to believe that some internal institution must be present. The situation just described is not a pure coordination game, and type 1 institutions are, hence, not relevant here. But any of the remaining three types of internal institutions might be relevant. A possible example for type 3 institutions, i.e. those we are most interested in here, is that information regarding the behaviour of party 2 is widespread, which will make it more difficult for her to find transaction partners in the future.⁵
- (2) In the second stylized interaction situation, individuals are interested in some public good, and internal institutions can help them in providing themselves with public goods. As nobody can be excluded from their consumption, standard economics teaches that those goods will be underprovided. This is a standard case of a market failure that is usually interpreted as sufficient reason for government intervention. But the provision of public goods without government intervention does occur. Much of Elinor Ostrom's work (most prominently 1990) describes not only cases in which groups of private individuals have been successful in managing common pool resources (and cases in which they have been unsuccessful), but also the conditions that need to be met for non-state provision to work. The conjecture here is that adequate internal institutions can increase the likelihood of the provision of public goods by allocating to certain individuals the responsibility to do certain things

⁵ As shown by Milgrom et al. (1990) with regard to information exchange among medieval merchants.

(e.g. participate in the provision every so often, participate in – costly – monitoring, etc.) and by sanctioning those who do not comply with their obligations.⁶

- (3) In many situations, politicians would be better off if they were not constrained by any constitutional rules, which are a specific kind of institution. If, however, rule-breaking is connected with a high probability of being sanctioned *ex post*, they might prefer to comply with the relevant institutions. Checks and balances can be thought of as an attempt to make politicians monitor each other. But this might not always be sufficient. In such instances, the capacity of individual citizens to coordinate their behaviour so as to put an effective check on government can be crucial in keeping the government within its constitutional confines (e.g. Weingast, 1993). This is the third stylized interaction situation. In a sense, of course, it is a specific case of the second situation; monitoring the government can be viewed as the provision of a public good.
- (4) Finally, there might also be internal norms that help politicians monitor each other's behaviour. These might be important for the actual implementation of some constitutionally mandated separation of powers. We do not dig any deeper into the analysis of these internal norms here, as the number of people whose behaviour is channelled by them is minuscule in comparison to the three other interaction situations just described.⁷

If one is interested in the effects of internal institutions on economic development, it might help to think of the four interaction situations just sketched as possibly having both a direct and an indirect effect on economic development. The direct effect on development is determined by whether a particular institution is conducive (or not) to development. The indirect effect is determined by the particular relationship of the respective internal institution with external institutions. Those internal institutions helping individuals to monitor their government would primarily have an indirect effect as they would impact the quality of government which, in turn, has an effect on economic development.

This distinction between a direct and an indirect effect also points at different functions of internal institutions. The direct effect refers to the capacity of society to coordinate its behaviour without resorting to the state. The indirect effect indicates to what degree the political institutions of a country are in line with the norms and customs of its people.

Having identified four interaction situations in which internal institutions could possibly have a direct or an indirect effect on economic development, it is important to emphasize that internal institutions are not necessarily conducive to development. As one example, if internal institutions restrict the

⁶ Sugden (1986) contains a more abstract description of the conditions that need to be met to increase the probability of private provision of such goods.

⁷ Shepsle and Weingast (1987) showed that the power of congressional committees that cannot be explained by formal institutions can be attributed to informal ones. For an analysis of the relevance of such institutions in Latin American politics, see Helmke and Levitsky (2006).

voluntary exchange of many goods, on religious grounds for example, economic development could be slowed down.

4. Why is measuring internal institutions so difficult?

A good starting point for discussing the difficulties in measuring internal institutions is an observation made by Elinor Ostrom (1996: 208): ‘These rules may be almost invisible to outsiders, especially when they are well accepted by participants who do not even see them as noteworthy’. This section serves to spell out some of these difficulties.

- (1) According to the definition of institutions proposed here, institutions consist of a commonly known rule to structure recurrent interaction situations (and some sort of sanctioning threat). Now, this rule might be commonly known among its users but might be difficult to grasp for outsiders. Problems promise to be severe if the rule has never been formally written down. In a word: the lack of formality is likely to imply difficulties in observability.
- (2) To be able to ‘measure’ an institution, we also need to know how sanctions are designed. Are they effective? Are there varying degrees of non-compliance with the rule? Are there differentiated sanctions depending on the degree of non-compliance? If the contents of the institution are not formally written down, it is difficult to empirically ascertain how specific sanctions are being used.
- (3) How do we delineate one institution from another? There are good arguments in favour of attempting to document institutions one by one.⁸ This attempt may be confounded by the fact that institutions are very likely to display their effects only in conjunction with other institutions. And we are primarily interested in their functions, not in institutions *per se*.⁹
- (4) While many groups who rely on a set of internal institutions are likely to live and act entirely within one nation state, the reach of internal institutions will often extend beyond the borders of a nation state. It is likely that groups of individuals who belong to the same ethnic group or who work in the same sector share some institutions. For example, a group of international merchants might belong to different nation states, but will likely share a set of internal institutions. This is also likely with regard to the adherents of a religious group spread across nation state borders. A further complication is individuals who might belong to a variety of groups and thereby be subject to more than one set of internal institutions. Think of an international merchant who also belongs to some local religious group and comes from an ethnic minority. Identifying the group that applies a specific set of internal institutions can also constitute a challenge.
- (5) As we already noted, Ostrom points out that some norms might be so self-evident to the users that they are unlikely ever to mention them explicitly. This implies that interviews are an incomplete method for producing a list of internal institutions.

⁸ Some of these are spelled out in Voigt (2013).

⁹ This point is made in Robinson (2013).

- (6) Surveys (a close relative of interviews) are also problematic. Surveys can help to elicit the attitudes of those being surveyed but not necessarily their real behaviour. Often, surveys seem to reflect behaviour deemed to be desirable by the respondents but not necessarily the way they actually behave.
- (7) It is frequently asserted that internal institutions are very slow to change (e.g. Williamson, 2000). It is also argued that, once a certain threshold is passed, they can change very quickly (e.g. Bicchieri, 2006). Even if we assume that we are able to measure currently valid institutions, such a measure would be insufficient to check the competing conjectures regarding the pace of change of internal institutions. In addition, we would need to be able to ascertain the content of internal institutions as they were hundreds of years ago. This constitutes an additional difficulty.

In sum, measuring internal institutions constitutes a major challenge. In the next section of this paper, I critically discuss some attempts that have been made to do so.

5. A critical look at some variables used as proxies for internal institutions

There are a number of papers that try to show empirically the importance of internal institutions. In this section, I take a critical look at the indicators most frequently used in these studies and conclude that they are not well-suited for the intended purpose. I make my point by discussing two concepts in an exemplary fashion, namely ‘trust’ and ‘social capital’, and a specific tool to elicit information on internal institutions, namely the World Values Survey. Trust and social capital seem to be the most frequently chosen proxies, and the World Values Survey is one of the most widely used surveys.¹⁰

Some authors who propose to show the relevance of internal institutions spend little or no time thinking about appropriate ways to measure them, but choose a seemingly innocuous shortcut by claiming that internal institutions and culture are more or less identical. They then continue with their analyses by relying solely on frequently used measures of culture. Tabellini (2010), who is very careful in not confounding internal institutions and culture, considers trust to be a central component of culture. Others are less careful and use trust as an indicator for internal institutions (e.g. C. Williamson, 2009).

It is argued that individuals who trust each other are able to exchange goods at lower cost than individuals who do not trust each other. But trust itself is not

¹⁰ I have picked trust and social capital because they are often used in conjunction with internal institutions. A different approach could have been to start with an area in which internal institutions are assumed to be highly relevant and ask for available measures. Referring to an example given in the introduction, one could ask for measures on traditional land rights in Africa and beyond. Such measures are a precondition for answering questions such as what factors determine the continued use of these institutions, what factors determine their quality, their compatibility with the rule of law and so forth. In a recent attempt to analyse these questions, I did not find any convincing measures (Gutmann and Voigt, 2016).

an institution but rather a propensity to believe the statements of others that refer to possible future behaviour. It can certainly be argued that the propensity to believe others is also a consequence of internal institutions. Trust could then be interpreted as a consequence of a great many institutions, and this is exactly the problem with this measure. Well-functioning external institutions might contribute to this propensity just as much as internal institutions. This implies that if trust is used as a measure of internal institutions, it is at best very imprecise.¹¹

Many lab experiments teasing out the degree to which participants trust other participants and behave in a trustworthy fashion have been run. In the lab external institutions are absent, which makes it possible to identify the effects of internal institutions on trust. Yet the observed behaviour can be the consequence of a great many institutions. Future experiments should, hence, shift their focus away from trust and toward institutions possibly inducing it.

A very similar argument can be made with regard to social capital, which is often declared to be an internal institution (see, for instance, various contributions to the special journal issue edited by Casson et al., 2010). Social capital is frequently equated with the capacity of a social group to solve collective action problems (e.g. Putnam, 1993). If a social group manages to coordinate without relying on the state, then its members must have been able to bring this coordination about by relying on internal institutions. Social capital, then, is not itself an internal institution, but the consequence of any possible number of internal institutions. The use of concepts such as trust and social capital as aggregate measures of institutions constitutes the central problem. In addition, we cannot be sure that the creation and conservation of social capital is brought about by relying exclusively on internal institutions. Given these issues, social capital, too, is a very imprecise measure for internal institutions at best.¹²

These two examples illustrate how researchers often prefer to deal not with single institutions but with amalgamations of them. It is not convincing to assume that both trust and social capital are determined exclusively by internal

11 Guiso et al. (2008) offered a well-documented example of how external institutions can determine trust levels. Putnam (1993) conjectured that the experience of being a citizen of an independent city state implies experiences of self-administration and of being responsible for one's own *res publica*. Putnam went on to suggest that successfully mastering the administration of their own affairs was likely to lead people to trust each other. Guiso et al. (ibid.) identified the 400 largest Italian cities in 1871 (when the first census after Italy became a nation state was taken) and traced their history back to the beginning of the second millennium. They found that cities that were independent then were very different from cities that had not experienced a long history of independence. There were more not-for-profit organizations in the former, participation in referendums was higher, and the likelihood that an organ donor organization existed was also higher. The authors claim that at least half of the differences in social capital can be explained by referring to the history of the specific city.

12 To make things even messier, Guiso et al. (2004) can be read as arguing that trust is a consequence of a society's level of social capital.

institutions. It is conceivable that an identical level of trust or social capital might be brought about by very different combinations of internal institutions.¹³

Moving on from our brief discussion of two frequently used proxies for internal institutions, we turn our attention to a discussion of the World Values Survey as a source of information about internal institutions. A number of researchers interested in the effects of internal institutions have relied on items contained in the World Value Surveys. In addition to the trust variable already criticized above, Tabellini (2010) attempts to measure culture by using questions referring to the respect of others, the value parents attach to obedience in raising their children and the degree to which surveyed persons feel that they are in control of their own lives. As alluded to above, Williamson (2009) does not shy away from using items from the World Values Survey as handy proxies for internal institutions. As a way of measuring institutions defined as we have here, these variables make little sense. Beliefs about the connection between effort and success in life are just that, beliefs. These beliefs do not, however, constitute an institution. The same applies to the set of priorities parents rely on to raise their children.

6. What to do then? Benefits from using experiments

In this section, I propose a way to measure internal institutions that I suggest is more systematic than the approaches described and criticized in [Section 5](#).

The triumphal procession of experiments in economics

Experiments have experienced a veritable boost in economics in recent years. Experiments allow researchers to observe the behaviour of subjects under precisely controlled conditions. They have been used to challenge some of the basic assumptions of economics and have already led to a new branch of economics, namely behavioural economics. These days, no major economics department can afford not to have its own laboratory to run experiments. But experiments are not confined to the laboratory. More and more, researchers have attempted to replicate experimental conditions in the field. I suggest that experiments can be used to identify the internal institutions that people use to structure specific interaction situations. When compared to surveys, experiments have the advantage that we are not surveying attitudes but observing behaviour that is incentivized by real payoffs.

¹³ The necessity to ‘unbundle’ sets of institutions into single institutions is not as evident as with regard to formal institutions, especially if that necessity is justified by pointing to policy concerns. Recommendations to improve ‘democracy’ or the ‘the rule of law’, two examples of broad sets of institutions that comprise possibly hundreds of single institutions, are such vague suggestions that they are not policy advice at all. Provided that informal institutions are largely exempt from intentional change, this argument does not apply to informal institutions. Then again, this is a strong assumption and it might be wrong. To test it empirically, informal institutions need to be unbundled too.

The environment in which experiments are being conducted can be controlled. The researcher is free to design an experiment that will cover the functions she is interested in analysing. Experiments can be carried out with precisely defined groups. Since experiments can be precisely designed, they can be used to mimic at least the first three stylized interaction situations described above, namely bilateral cooperation on a voluntary basis, the voluntary provision of public goods, and the ability to coordinate behaviour to monitor the governing. Literally hundreds of experiments have already been carried out investigating the first two interaction situations (for an overview regarding the first, see Fehr and Schmidt, 2006, for an overview regarding the second, see Chaudhuri, 2011).¹⁴

Running experiments to ascertain internal institutions

The major limitation of using these experiments as proxies for internal institutions is that they are usually played in a single place. My proposal here is that conducting identical experiments in many different societies can serve to establish measures of internal institutions that could be used for cross-country studies. The idea that it might be possible to conduct an ‘identical’ experiment in a variety of settings poses a challenge. Societies differ in mean income, education attainment and so on. Adjusting payoffs according to the average income of a country seems the least of the challenges. Non-familiarity with written texts can be dealt with by reading the experiment out loud, etc. But frequently, concepts are not easily translatable across languages.

Before dealing with some of the challenges of such an approach in a little more detail, I would like to point out that experiments regarding the first three stylized interaction situations described in Section 3 above have already been conducted. The challenge consists in modifying the experiments so that they are amenable to a cross-national experiment. As a first take on internal institutions, it seems to make sense to have the experiment played in a context in which external institutions are assumed to be absent. True, trying to coordinate ‘in the wild’ will be closer to real life in some countries than in others, but it seems to be the most straightforward way to ascertain internal institutions untampered by re-enforcing or mitigating factors. Since external institutions can be ascertained by non-experimental means, the relationship between internal and external institutions can be ascertained once the valid internal institutions have been identified.

The three stylized interaction situations are important in all societies regardless of size, wealth or geographical condition. However, the concretization of the stylized situations might differ substantially across societies. This means that experiments should be sufficiently abstract so that subjects can easily apply them to a situation relevant in their lives.

¹⁴ Regarding the third interaction, fewer experiments have been run. Nevertheless, the situation has also been subject to a number of experiments such as Cason and Mui (2014).

Possible pitfalls in cross-national experiments

If one is interested in the institutions that are being employed by the population at large, relying on pools of university students is unconvincing. This problem has long been acknowledged by experimenters, but most experiments are still run on this basis. To ascertain internal institutions, a pool of participants that is representative of the entire population is clearly desirable. Attaining representative samples might also be a challenge as, for example, women might be less inclined to participate in experiments.¹⁵ Assume a highly homogeneous society and the argument that differences in the way the game is played across countries are due to differences in institutions. Such attribution of differences on the country level will only be credible once within-country heterogeneity has been accounted for; this implies that even in homogeneous societies, the game should be played at least twice. In diverse societies it is conceivable that different parts of the population adhere to different internal institutions, hence, the possibility of institutional heterogeneity within countries exists. It would, thus, not suffice to run the experiment only twice per country.

I assume the most frequent argument against this proposal will be to doubt the external validity of experiments, i.e. the validity of the results beyond the laboratory. Experimental economists are, of course, confronted with this doubt on a daily basis. Some have dealt with it in innovative ways. Rustagi et al. (2010), for example, ask whether the degree to which groups cooperate conditionally in the lab is reflected in the degree to which they manage their common pool resources in real life. *In concreto*, they experimented with 679 individuals who are members of 49 different groups that use the forest as a commons in Ethiopia. They found that a higher proportion of conditional cooperators (those who contribute more to the commons, the more others contribute) is, indeed, a significant predictor for the quality of (real-life) commons. In addition, they elicited information on the willingness to spend resources on costly monitoring via a survey instrument and found that the declared willingness to invest is, again, highly significant for the quality of the (real-life) commons. In a sense, experimental and survey results can be used to predict real-life outcomes. This is an important study on the external validity of both experiments and surveys.

Bringing surveys back in?

But running experiments in 100 or more countries is extremely costly, if it is even possible. Costs refer not just to the monetary costs, but also to the need to find reliable academic partners in all of these countries, or travel

¹⁵ This constitutes a challenge but it can be risen to. In running experiments with refugees from the Arab world, we were confronted with exactly this problem. We increased our share of female refugees by contacting groups that were exclusively frequented by women, such as sewing clubs. Ensminger and Henrich (2014) contains the description of how these and similar issues have been dealt with by very experienced experimenters.

there oneself. There have been various attempts to draw on surveys to mimic experiments. Fehr et al. (2002) proposed to integrate behavioural experiments into representative surveys. They believed that by such integration, the shortcomings of both approaches could be overcome. More recently, Falk et al. (2015) ran comprehensive pre-tests eliciting information on preferences via both surveys and experiments. They then chose survey questions that were most aligned with how subjects behaved in the experiments. Ultimately, they had the survey implemented in 76 countries interviewing 80,000 respondents.

Because the entire population of a country might not share some internal institutions, it would be helpful to have a measure for the degree of heterogeneity among a specific population. If a survey is administered in such a way that it purports to be representative of a country's entire population, then the measured standard deviation from the mean could be used to determine the heterogeneity of that group. The same procedure can be used with experiments.

Some forerunners

A very limited number of experiments have been carried out among different groups. Henrich et al. (2001) is probably not only the first such study but also the most prominent one. In this study, Henrich and his co-authors cover 15 very small societies in 12 countries on five continents. Their results uncover a greater variation in behaviour than found in previous studies, which relied on student populations. Although some of their results are very interesting, they are not useful for our purposes as none of the societies is in any way representative of an entire country.¹⁶ Over the last couple of years, experiments have been carried out in a higher number of countries. Vieider et al. (2012), who identify risk attitudes in 30 countries and enquire into their growth effects, is one such example. Compared to more standard growth models, 30 countries is still a small sample (see also Vieider et al. 2015).

The study closest to what we have in mind for making internal institutions measurable was conducted by Herrmann et al. (2008). The study involves playing a particular public good game. The participants were drawn from 16 groups ranging from Boston and Melbourne to Samara, Riyadh, Athens and Muscat. The game was played in groups of four. After having contributed (or not) to the public good in the first round, all participants were given the opportunity to punish anybody in the second round. Intuitively, it would seem that those who tried to free ride, i.e. those who contributed nothing or below average, would be prime targets for being punished. If this is the case, and everybody in society knows this, such punishment can be instrumental in making the voluntary provision of a public good sustainable over time. Herrmann et al. (ibid.) call this

¹⁶ One of the determinants for differences in how subjects played their game was the degree to which the societies were integrated into market transactions. Regarding this criterion, the overwhelming majority of today's countries do not display much variation as most of their members are integrated via markets.

behaviour ‘pro-social punishment’. But they also observed that participants who had contributed above average were also being punished and call this ‘anti-social punishment’. If anti-social punishment is widespread, the likelihood that many individuals will be ready to contribute voluntarily to the provision of public goods is expected to be low. In other words: if anti-social punishment is widespread, we expect few public goods to be provided on a voluntary basis.

Interestingly, the prevalence of anti-social punishment seems to follow a clear pattern: it is lowest in highly developed places belonging to the common law legal family (Boston, Melbourne, Nottingham), followed by high-income places on the European continent (St. Gallen, Zurich, Bonn, Copenhagen). Groups with the highest proportion of anti-social punishment are located in the Middle East (Muscat and Riyadh). Interestingly, at least by this standard, Athens seems to belong to the Middle East as its proportion of anti-social punishers is between those of Riyadh and Muscat.

Measuring the prevalence of anti-social punishment is not measuring a particular internal institution. Yet the prevalence of anti-social punishment could be relevant to a wide range of institutions: if (high, above average) contributors face a high probability of being sanctioned for their behaviour, the size of contributions is likely to diminish over time. In other words: societies in which anti-social punishment is widespread are unlikely to be able to provide themselves with public goods in the absence of government intervention. If I am interested in a public place being clean and contribute to its cleanliness but am sanctioned due to my voluntary contribution, the public place will be dirty fairly soon. Herrmann et al. (2008) find that contributions to their game negatively depend on the amount of anti-social punishment.¹⁷

It is fascinating to speculate about further possible implications. It was mentioned above that the capacity to monitor one’s own government is a special case of the voluntary provision of a public good. If those contributing to the monitoring of the government are sanctioned and they contribute less as a consequence, a society’s ability to monitor its own government will be low, implying that the government is likely to get away with sub-standard provision of public goods, high degrees of corruption and so forth. Speculating even more, one could expect individuals in such societies to value the private over the public.

One way to use the findings by Herrmann et al. (2008) could be to work with the differences between pro-social and anti-social punishment: the greater

¹⁷ Oosterbeek et al. (2004) is a meta-study drawing on 75 ultimatum game results played in 25 different countries. The authors are interested in ascertaining whether differences both in (mean) offers as well as (mean) rejection rates can be explained by the cultural background of the country in which the game was being played. They are hence interested in the determinants of behaviour in a particular game. This is beyond the core question of this paper, namely how to measure internal institutions in the first place (and they do not find that cultural background is significant for explaining differences in observed behaviour).

the difference, the easier it should be for societies to provide themselves with public goods without depending on the state. As interesting as these results (and speculations) might be, they still constitute only a very partial measure of type 3 institutions. We do learn something about the willingness to punish others, i.e. about the sanctioning component, but this is not connected to a particular rule. Experiments designed to identify internal institutions and allow for comparisons across societies should aim to discover information about a relevant rule as well as the sanctioning component of institutions.

The recent paper by Falk et al. (2015) is interesting not only because of its attempt to use experiments to validate surveys but also because it contains a dimension, ‘negative reciprocity’, that covers such items as the willingness to take revenge, the willingness to punish unfair behaviour towards oneself, and the willingness to punish unfair behaviour towards others. In particular the last item seems a precondition for the sustainability of type 3 institutions, i.e. those internal institutions whose non-compliance is sanctioned by others, but in an unorganized manner.

Yet like the paper by Herrmann et al. (2008) discussed above, Falk et al. fall short of measuring concrete institutions as they only deal with the sanctioning component and not the rule part of an institution. The next step would be first to systematically identify the rules used to structure common and recurring interaction situations, and then to elicit information on sanctioning behaviour.

A first pilot study for the cross-country experimental approach to ascertaining internal institutions might rely on Amazon’s Mechanical Turk. This is an internet based labour market for hiring people for minor tasks and has been used to run a number of experiments (both Mason and Suri, 2012, and Paolacci et al., 2010, describe how M-Turk can be used to run experiments). Although it is clearly not ideal because participants are extremely unlikely to be representative of any society, it also has advantages: experiments would not need to be translated into many languages – with all the known problems – and overall costs would remain justifiable.

7. Beyond type 3 institutions

The focus of this paper has been on type 3 institutions, i.e. rules whose non-compliance is sanctioned spontaneously by third parties. I chose this focus because this type of institution seems to be the most difficult to measure. This paper cannot deal with the other internal institutions as deeply as type 3 institutions. Nevertheless, a very brief discussion of possible ways to measure the other types might still be useful.

Pure coordination games have at least two Nash equilibria. Essentially, a type 1 institution is that equilibrium that is actually chosen by those participating in the game. Schelling (1960) was probably the first to point out that the salience of particular equilibria makes them more likely to be chosen by persons wishing

to coordinate their behaviour with others. It seems likely that the salience of a particular equilibrium is determined by various aspects of the context in which the game is played. Framed differently: formally identical games are likely to be solved differently depending on all sorts of context factors such as history, religion, etc. Mehta et al. (1994) experimented with pure coordination games to determine the salience of particular equilibria. Running such experiments in different countries could help identify differences in type 1 institutions. More recently, Jackson and Xing (2014) had students from India and the United States play a variant of the battle-of-the-sexes game and found that the likelihood of coordinating successfully is also determined by nationality. The same holds true with regard to the stag-hunt game, which was played among Japanese, Korean and Chinese students (Stoddard and Leibbrandt, 2014).

We have defined type 2 institutions as the (ethical) self-commitment of actors to a particular behaviour. An actor not complying with her own self-commitment expects to experience a lower level of utility. Since the sanctioning in case of non-compliance is not done by any third party, ‘observing’ or ‘measuring’ type 2 institutions is particularly difficult. There is an important overlap between culture and type 2 institutions. We propose to define it as follows: Culture is the sum of rules, beliefs, preferences and values shared with many other members of a particular group that are neither self-enforcing nor enforced by third parties. According to this proposal, culture does not cover self-enforcing institutions (hence type 1 institutions) nor rules whose non-compliance is sanctioned by others (hence institutions of type 3 and 4). But institutions of type 2 can be thought of as part of culture. This proposal implies that there is an overlap between institutions and culture, but also that there are institutions that are not culture and there is culture that does not have the quality of an institution.

Experiments enquiring into particular institutions of type 2 should ideally contain no sanctioning mechanism that could be attributed to the structure of the game (type 1) or a third party (types 3 or 4). Studies on ‘comparative culture’ frequently rely on Hofstede (1991), who proposes four dimensions (namely power distance, collectivism *versus* individualism, femininity *versus* masculinity, and uncertainty avoidance) or on Inglehart (e.g. Inglehart and Welzel, 2005), who relies on two dimensions (namely traditional *versus* secular values, and survival *versus* self-expression). These dimensions have been derived from numerous surveys such as the World Values Survey. They are, hence, not theory-driven. In enquiring into valid type 2 institutions in our context, at least two modifications seem in order. First, the starting point of the enquiry should be the particular function that one is interested in. What are the possible functions that some of these institutions might have for economically relevant behaviour of actors? Second, the primary focus in previous studies seems to have been in the identification of different cultures. These cultures typically span numerous nation states and can, hence, be found in different nation states. In this

paper, the focus is on the (in-) compatibility of internal and external institutions. External institutions are typically defined as national ones, implying that our focus would be on identifying particular national and possibly subnational institutions.

We propose to define type 4 institutions where non-compliance with rules is sanctioned via organized private enforcement. Private arbitration courts are perhaps the most relevant example. In [Table 2](#), we placed ‘private courts’ at the intersection of formal rules and internal enforcement. To the degree that the rules are formalized, measuring and comparing them should not be extraordinarily difficult. The terms of business used by formally organized providers of private court services will be found regularly on their web pages for example. But there are ‘private courts’ that rely on informal rules. Consider countries in which traditional or indigenous law plays a significant role in sanctioning non-compliance. I am not aware of any attempts to compare systematically the traditional rules underlying the working of private courts that rely on indigenous law in such countries.¹⁸

8. Outlook: lessons for future research

This paper addresses a number of issues and offers several proposals. After presenting a taxonomy for internal institutions, the paper develops a brief critique of some of the variables that have been used as proxies for internal institutions. We propose that there is a need to rely more on experiments (possibly combined with surveys) to make internal institutions measurable, making it possible to compare institutions across countries. The paper by Falk et al. (2015) contains an important first step in that direction, but many questions remain unanswered. Two of the most pressing questions are: What survey questions are best able to catch how people really behave? How do experiments need to be designed to ensure the highest possible amount of external validity?

This paper has focused on type 3 institutions and has argued that the degree of anti-social punishment would be a good starting point for a general measure of type 3 institutions. Suppose this is correct, where can this measure be applied? We see at least three such applications that could add value. First, it can be used to ascertain the direct effect of type 3 institutions on development by plugging it into, for instance, a standard growth model. Second, such a measurement can be used to ascertain the relationship type 3 institutions have with the respective external institutions of a particular country. It might well be that type 3 institutions are conducive to development *per se* but that they are combined

¹⁸ Voigt and Park (2013) is an attempt to deal with the relationship between state and non-state adjudication. It is based on some surveys run by the World Bank to enquire into the frequency with which non-state arbitration is used, but refrains from coding the underlying rules in any way.

with competing or conflicting external institutions. If measures for both external and internal institutions are available, they could both be inserted into a growth model, possibly including their interaction effect. Finally, as soon as measures for the other internal institutions are available, one can look at their relationship, possibly even putting measures for different types of internal institutions into a growth model.

At least two important follow-up questions remain to be dealt with. North (1990: 37) conjectures that internal institutions ('informal constraints') are important sources of continuity in a world in which external institutions can be changed at very short notice.¹⁹ But internal institutions are also man-made and not entirely exogenous. This implies that looking for their determinants or roots is an interesting research question.

Finally, in new institutional economics it might be possible to identify at least two different camps: those who attribute primacy to internal institutions on the one hand and those who attribute primacy to external institutions on the other. Those belonging to the first camp would point out that at the end of the day, implementation of fundamental institutions – such as the constitution – cannot be secured by external institutions but depends on conforming to firmly held preferences and norms, i.e. internal institutions. According to this view, if external and internal institutions do not complement each other, then external institutions need to be modified, since internal ones are largely exempt from intentional modification. On the other side, there are those who also claim that external institutions might inhibit growth and development, but for entirely different reasons. According to them, more inclusive external institutions can be introduced and implemented as long as the elite is willing to support them. Of course, there might be a third camp consisting of those who find a bit of truth in both the first and the second camp. Being able to measure institutions will help to identify the relative merit of these competing views and advance our knowledge regarding the functioning of institutions.

Acknowledgments

The author thanks Gregory Schaffer and Michael Witt for inviting him to the panel “Measuring Institutions: Issues and New Perspectives” at the 2015 SASE conference in London and Afef Ayouni, Lukas Böker, Jerg Gutmann, Amel Hammami, Monia Idoudi, Julia Lemke, Stephan Michel, Nisrine Ouni, Katharina Pfaff, Konstantinos Pilpilidis, and Agnes Strauss for constructive suggestions.

¹⁹ Similar arguments have been made by others, among them Oliver Williamson (2000) who claims that informal institutions are largely exempt from intentional modification and that they are very stable over time. Roland (2009) makes a similar argument introducing the distinction between ‘slow moving’ and ‘fast moving’ institutions.

References

- Allina-Pisano, J. (2007), *The Post-Soviet Potemkin Village: Politics and Property Rights in the Black Earth*, Cambridge University Press.
- Barnett, A., B. Yandle and G. Naufal (2013), Regulation, trust, and cronyism in Middle Eastern societies: the simple economics of ‘wasta’, *Journal of Socio-Economics*, **44**: 41–6.
- Bicchieri, C. (2006), *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge: Cambridge University Press.
- Cason, T. and V. L. Mui (2014), Coordinating resistance through communication and repeated interaction, *Economic Journal*, **124**(574): F226–F256.
- Casson, M., M. della Giusta and U. Kambhampati (2010), Formal and informal institutions and development (Special issue), *World Development*, **38**(2): 137–41.
- Chaudhuri, A. (2011), Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature, *Experimental Economics*, **14**(1): 47–83.
- Ensminger, J. (1997), Changing property rights: reconciling formal and informal rights to land in Africa. In: J. Drobak and J. Nye (eds): *The Frontiers of the New Institutional Economics*, San Diego, Academic Press, pp. 165–96.
- Ensminger, J. and J. Henrich (eds) (2014). *Experimenting with Social Norms – Fairness and Punishment in Cross-Cultural Perspective*. New York et al.: Russell Sage Foundation.
- Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman and U. Sunde (2015), *The Nature and Predictive Power of Preferences: Global Evidence*, typescript, November (Revise & Resubmit at *Quarterly Journal of Economics*).
- Fehr, E., U. Fischbacher, B. v. Rosenblatt, J. Schupp and G. Wagner (2002), A nationwide laboratory examining trust and trustworthiness by integrating experiments into representative surveys, *Schmollers Jahrbuch*, **122**(4): 519–42.
- Fehr, E. and K. Schmidt (2006), The economics of fairness, reciprocity and altruism – experimental evidence and new theories. In: S.-C. Kolm and J. Ythier (eds): *Handbook of the Economics of Giving, Altruism and Reciprocity*, Volume 1. North Holland: Elsevier, pp. 616–91.
- Guiso, L., P. Sapienza and L. Zingales (2004), The role of social capital in financial development, *American Economic Review*, **94**(3): 526–56.
- Gutmann, J. and S. Voigt (2016), Why is traditional law so prevalent? Identifying reasons for its widespread use, mimeo, available from: Institute of Law & Economics, University of Hamburg, Germany.
- Guiso, L., P. Sapienza and L. Zingales (2008), Long-term persistence, NBER Working Paper 14278.
- Helmke, G. and S. Levitsky (2006), *Informal Institutions and Democracy: Lessons from Latin America*, Baltimore: Johns Hopkins University Press.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis and R. McElreath (2001), In search of homo economicus: behavioral experiments in 15 small-scale societies, *American Economic Review*, **91**(2): 73–8.
- Herbst, J. (2014), *States and Power in Africa: Comparative Lessons in Authority and Control*, Second Edition Princeton: Princeton University Press.
- Herrmann, B., C. Thöni and S. Gächter (2008), Anti-social punishment across societies, *Science*, **319**(5868): 1362–7.
- Hofstede, G. (1991), *Cultures and Organizations – Intercultural Cooperation and its Importance for Survival*, McGraw-Hill International.

- Inglehart, R. and C. Welzel (2005), *Modernization, Cultural Change, and Democracy: The human Development Sequence*, Cambridge University Press, 2005.
- Jackson, M. O. and Y. Xing (2014), Culture-dependent strategies in coordination games, *Proceedings of the National Academy of Sciences*, **111**(3): 10,889–96.
- Kiwit, D. and S. Voigt (1995), Überlegungen zum institutionellen Wandel unter Berücksichtigung des Verhältnisses interner und externer Institutionen, *ORDO – Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*, **46**: 117–48.
- Mason, W. and S. Suri (2012), Conducting behavioral research on Amazon’s Mechanical Turk, *Behavior Research Methods*, **44**(1), 1–23.
- Mehta, J., C. Starmer and R. Sugden (1994), The nature of salience: an experimental investigation of pure coordination games, *American Economic Review*, **84**(3): 658–73.
- Milgrom, P., D. North and B. Weingast (1990), The role of institutions in the revival of trade: the law merchant, private judges, and the champagne fairs, *Economics and Politics*, **2**(1): 23.
- North, D. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge, MA: Cambridge University Press.
- Oosterbeek, H., R. Sloof and G. van de Kuilen (2004), Cultural differences in ultimatum game experiments: evidence from a meta-analysis, *Experimental Economics*, **7**(2): 171–88.
- Ostrom, E. (1990), *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge: Cambridge University Press.
- Ostrom, E. (1996), Incentives, rules of the game, and development, M. Bruno (ed.): *Annual World Bank Conference on Development Economics*, Washington, DC: World Bank, 207–34.
- Paolacci, G., J. Chandler and P. Ipeirotis (2010), Running experiments on Amazon Mechanical Turk, judgment and decision making, **5**(5): 411–19.
- Peters, P. E. (2009), Challenges in land tenure and land reform in Africa: anthropological contributions, *World Development*, **37**(8), 1317–25.
- Putnam, R. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.
- Robinson, J. (2013), Measuring institutions in the Trobriand Islands: a comment on Voigt’s paper, *Journal of Institutional Economics*, **9**(1): 27–9.
- Rodrik, D., A. Subramanian and F. Trebbi (2004), Institutions rule: the primacy of institutions over geography and integration in economic development, *Journal of Economic Growth*, **9**(2): 131–65.
- Roland, G. (2009), Fast-moving and slow-moving institutions, *Institutional Change and Economic Behaviour*, Basingstoke: Palgrave Macmillan UK, **2008**, 134–59.
- Rustagi, D., S. Engel and M. Kosfeld (2010), Conditional cooperation and costly monitoring explain success in forest commons management, *Science*, **330**(6006): 961–5.
- Schelling, T. (1960), *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.
- Shepsle, K. and B. Weingast (1987), The institutional foundations of committee power, *American Political Science Review*, **81**(1): 85–104.
- Sikor, T. and D. Müller (2009), The limits of state-led land reform: an introduction, *World Development* (Special issue on the limits of state-led land reform), **37**(8): 1307–16.
- Stoddard, O. and A. Leibbrandt (2014), An experimental study on the relevance and scope of nationality as a coordination device, *Economic Inquiry*, **52**(4): 1392–407.
- Sugden, R. (1986). *The Economics of Rights, Co-operation and Welfare*, Volume 22. Oxford: Basil Blackwell.

- Tabellini, G. (2010), Culture and institutions: economic development in the regions of Europe, *Journal of the European Economic Association*, 8(4): 677–716.
- Vieider, F. M., T. Chmura and P. Martinsson (2012), Risk attitudes, development, and growth: Macroeconomic evidence from experiments in 30 countries (No. SP II 2012–401), WZB Discussion Paper.
- Vieider, F. M., M. Lefebvre, R. Bouchouicha, T. Chmura, R. Hakimov, M. Krawczyk and P. Martinsson (2015), Common components of risk and uncertainty attitudes across contexts and domains: evidence from 30 countries, *Journal of the European Economic Association*, 13(3): 421–52.
- Voigt, S. (2013), How (not) to measure institutions, *Journal of Institutional Economics*, 9(1): 1–26.
- Voigt, S. and D. Kiwit (1998), The role and evolution of beliefs, habits, moral norms and institutions. In: H. Giersch (ed.): *The Merits of Markets: – Critical Issues of the Open Society*, Berlin et al.: Springer, pp. 83–108.
- Voigt, S. and S.-M. Park (2013), Arbitration is no substitute for state courts, *Journal of Development Studies*, 49(11): 1514–31.
- Weingast, B. (1993), Constitutions as governance structures: the political foundations of secure markets, *Journal of Institutional and Theoretical Economics*, 149(1): 286–311.
- Williamson, C. (2009), Informal institutions rule: institutional arrangements and economic performance, *Public Choice*, 139(3/4): 371–87.
- Williamson, O. (2000), The new institutional economics: taking stock, looking ahead, *Journal of Economic Literature*, 38(3): 595–613.