On the validity of the division into formal and informal institutions

IZABELA OSTOJ

University of Economics in Katowice, Faculty of Economics, Department of Economics, ul. 1 Maja 50, 40-287 Katowice, Poland

izabela.ostoj@ue.katowice.pl

orcid.org/0000-0003-2997-280X

Abstract

Motivation: Definitions and classifications constitute the most important elements of each theory. The division of institutions into formal and informal is commonly applied within new institutional economics. In his papers devoted to institution measurement S. Voigt developed the division of institutions into internal and external, argued its functionality and, to a certain extent, its superiority.

Aim: The aim of this paper is to investigate and assess the validity and applicability of the division of institutions into formal and informal from the perspective of the analysis of social and economic phenomena.

Results: Based on literature, as well as drawing on the origins of the institution classification and referenced examples, it was proved that the division into formal and informal institutions is not only at least equally precise as other classifications, but it is also consistent and logical, which determines its high usefulness in scientific research.

Keywords: institutions; formal institutions; informal institutions; legal origin

JEL: B5; K4

1. Introduction

Definitions and classifications constitute the most important elements of each theory, determining its cognitive and application merits. The organization and classification of knowledge based on the principles of explanation are the main goal of science (Madden, 1963, pp. 64–70; Mormann, 2008; Nagel, 1961, p. 4). Their role is particularly important in the ‘young’ fields, whose
notions and classifications may not be sufficiently mature or structured. One of such approaches is new institutional economics. Research conducted within this field commonly applies the division of institutions into formal and informal (e.g. Marcesse, 2018, pp. 284–296; Meon & Sekkat, 2015, pp. 754–771; Pejovich, 2008, p. 11; Redmond, 2008, pp. 569–576; Seidler, 2018, pp. 289–312; Williamson, 2009, pp. 371–387). This division, adopted and disseminated by North (1992, p. 9), is one of the most frequently used and relatively rarely criticized classifications of institutions. In the articles published in the Journal of Institutional Economics, Voigt (2013a, pp. 1–26; 2018, pp. 1–22) undermines, to a certain extent, the accuracy and relevance of the division of institutions into formal and informal ones. At the same time, he argues for the validity and superiority of the division of institutions into internal and external ones, adopting an entity enforcing compliance with the rules as a criterion. The article triggered a discussion (Robinson, 2013, pp. 27–29; Shirley, 2013, pp. 31–33; Voigt, 2013b, pp. 35–37) in which the issue of the classification of institutions was mentioned as part of the problems involved in their measurement¹. However, those publications did not discuss the issue in greater depth.

The aim of this article is to assess the validity and applicability of the division of institutions into formal and informal. It proposes the arguments in confirmation of the thesis that the division of institutions, proposed by Voigt, into external and internal according to the entity enforcing the compliance with the rules is imprecise and difficult to apply, while the division of institutions into formal and informal, despite certain shortcomings, enables the complete and logically coherent analysis of socio-economic phenomena. The article presents the core assumptions and sources of the division of institutions into formal and informal as well as its critique. Additionally, it provides arguments revealing the weaknesses of the proposed new division of institutions. An attempt is made to prove the validity and universal character of the division of institutions into formal and informal ones. This was achieved based on the literature on the subject, in particular journals and case studies.

2. Literature review

2.1. The division into formal and informal institutions

In his research on institutional change processes, North (1992, p. 9) assumed that institutions include formal rules, informal constraints (norms of behavior, conventions, and self-imposed codes of conduct), and the enforcement characteristics of both. Thus, he adopted the division into formal and informal institutions, according to the rule component as the criterion. This division has become the most widely used and recognizable element of studies conducted

¹ The publication of the articles engaging in the discussing with S. Voigt’s views was possible in the same issue, since the article was also published online several months earlier.
in new institutional economics. It is, in fact, the further development and concretization of the approach proposed by Veblen, the founder of institutionalism (which was labeled with the adjective ‘original’ in later years), to define social institutions. Veblen described them as a complex of habits of thought and conventional behavior (Blaug, 1985, p. 709). Veblen needed this definition to critique social and economic life, of which he was an insightful observer and proposed a specific and timeless interpretation of people’s beliefs and attitudes in business relations. Institutions as a set of formal and informal rules, including their enforcement arrangements, had already been defined by Schmoller in his publication in 1900 (Furubotn & Richter, 1998, p. 6). A similar division was also developed by Hayek (1998, p. 37), who referred to the origins of institutions and identified spontaneous institutions, emerging in an evolutionary manner, thus endogenous in nature, and exogenous institutions, created by the rulers in an intentional manner, often based on earlier informal principles. Voigt (2013a, pp. 1–26; 2018, pp. 1–22) refers to this division, but he distorts it, changing the criterion of classifying institutions as internal and external, which will be discussed further in this article.

North used this division of institutions primarily to explain the mechanisms of economic change. In North’s (2005, pp. 1–2) approach, the structure people impose on their lives to reduce uncertainty is an accumulation of prescriptions and proscriptions together with the artifacts that have evolved as part of this accumulation. The result is a complex mix of formal and informal constraints. As he emphasizes, these are imperfect constructions, they bring about uncertain results due to the inadequate understanding of the environment they concern, as well as the shortcomings of formal and informal mechanisms of their enforcement. According to North, this division seems the most appropriate and sufficient for the analysis that he conducts, despite the weaknesses of which he is aware.

If we view the problem in historical terms, institutions were initially only of informal character. Formalization was influenced by the development of the interests of groups belonging to a given community and striving to assign to them the appropriate meaning, interpretation and enforcement rules. Formalization is usually in the interest of individuals or groups that possess certain power or have gained a dominant position based on the existing informal rules (Redmond, 2008, pp. 569–576). Looking at the historical sources of the division of institutions into formal and informal, it seems natural that, by its very nature, such division will be imprecise due to the possible transitional nature of some institutions, from informal to formal, as well as diverse forms and degrees of their formalization.

An important contribution to the analysis and interpretation of the division of institutions into formal and informal was made by Hodgson (2006, pp. 1–25; 2015, pp. 497–505). He posits that the commonly used division of institutions into formal and informal is not sufficiently precise, but he does not criticize it. He argues that definitions in the social sciences are likely to have fuzzy bounda-
ries. This is also the case of institutions and, consequently, it will affect the precision of their classification. Hodgson proposes the definition of institutions as the integrated systems of rules that structure social interactions and gives us a rough but useful demarcation criterion to distinguish institutions from other social phenomena. He points out that some authors perceive formal institutions as codified, others emphasize that they are designed, while still others use the term formal in relation to laws. Another interpretation proposition involves understanding formal as legal, whereas informal as non-legal (even if they were written down), but then it is uncertain whether informal norms should be understood as only non-legal, or as illegal as well. He also identifies the possibility of understanding formal institutions as explicit (clear, precise) rules and informal institutions as tacit rules. Hodgson quotes the approach proposed by Menger, who wrote about pragmatic and organic institutions. Hodgson does not decide which meaning is the best, although in 2006 he proposes using more precise notions to distinguish between institutions, such as legal and non-legal. He suggests, however, that we should always specify how we understand them in a given case. Therefore, he allows some freedom or ‘blur’ in the interpretation, admittedly though he uses categories that are semantically related and the reader can intuitively understand them in the way intended by the author. Notably, he does not use the distinction of institutions based on the enforcement criterion.

### 2.2. S. Voigt’s critique of the division of institutions into formal and informal

Voigt (2013a; 2013b) formulates his criticism regarding the division of institutions into formal and informal, maintaining that the rules emerging spontaneously tend to become formalized over time. He argues that it is not known to what extent an institution must be formalized so that it can be regarded as formal, pointing out that the division lacks precision. He poses the questions: does an institution have to be recorded? Does it have to pass the legislative procedure? Therefore, he proposes the division of institutions into external and internal according to the criterion of the entity enforcing compliance with the adopted norms. According to Voigt, if the state enforces compliance with the rules, institutions can be described as external, and if society members enforce compliance with the rules, institutions can be regarded as internal; while discussing this division, he also explains external standards as publicly sanctioned, while internal ones as privately sanctioned. Interestingly, Voigt’s (2011, pp. 319–330) earlier publication, which analyzes the issues of constitution adoption, ‘constitutional equilibria’ and their changes in various countries, raises the problem of endogenizing constitutions. The endogenization of constitutional rules helps to understand why different societies adopt different principles in their con-

---

2 This division is also imprecise because it does not account for quasilegal norms emerging as a result of collective bargaining.
stitutions. If constitutional change is interpreted as the consequence of some disequilibrium, then a constitution that is respected by society can be regarded as in equilibrium. The endogenization of constitutions can essentially explain the need to change a constitution with seeking a new equilibrium, if a constitution does not meet the needs of society (constitutions are, on average, changed every 20 years). The compliance with a constitution is generally enforced by the state (more or less effectively), but the endogenization of a constitution does not mean that the society will start to enforce its provisions, which would result from the criterion of the division of institutions into internal and external. There is therefore a contradiction in the arguments presented in different periods.

In his critique, Voigt (2013a, 2013b) starts with the definition of an institution as ‘commonly known rules used to structure recurrent interaction situations that are endowed with a sanctioning mechanism’. In the division of institutions into external and internal focusing on mechanisms relating to the enforcement of the adopted norms, he does not clearly identify the sources of external and internal institutions, whereas it is in this area that problems arise. The adoption of the criterion of the enforcing entity causes that the state and society are put in opposition; the state is external in relation to the society and interactions occurring in this society.

3. Methods

The formulation of conclusions about the validity and role of the division of institutions into formal and informal was conducted based on the review of literature, primarily articles published in current journals dedicated to research in the field of institutional economics, mainly in the Journal of Institutional Economics, and to institutional analysis of socio-economic phenomena. The method involved critical analysis of publications relating to the classification of institutions and critique of the arguments relating, in particular, to the division of institutions proposed by S. Voigt. It allowed the author to distinguish three sources of these arguments: the analysis of the views of the main representatives of new institutional economics, legal tradition and case studies. Logical deduction and comparative analysis of various theoretical approaches were also used.

4. Results

4.1. Arguments identifying the shortcomings of the division of institutions into external and internal

This part of the article reviews the division of institutions into internal and external or related ones, the context of which is however different from the division developed by Voigt. They show that such a division is highly ambiguous,
because authors, in fact, adopt different classification criteria, while applying the notions commonly used in many other classifications.

The division of institutions into external and internal proposed by Voigt bears some similarity to Hayek’s (1998, p. 36) approach, because of the terms used by Hayek when he discusses the order created by forces outside the system (exogenously) and equilibrium set up inside the system (endogenously). Voigt, however, chooses a different criterion for classifying institutions — the entity enforcing the compliance with the rules. The overlap of these two approaches is misleading, because research practice does not involve using the same notions for divisions based on different criteria. Another aspect of the division emerges when Voigt (2018) treats external institutions as national, in the sense of imposing and enforcing them by the state. He does not take into account the rules set and enforced at other levels, e.g. by international organizations, which is a natural process in the conditions of economic integration and globalization. This points to a limited scope of the proposed division.

M. Aoki refers to the division of institutions into internal and external in his original concept of institutions-as-equilibria and the definition of institution as reconciliation of internalism and externalism. Yet, he defines this division completely differently, because the internal refers to the state of mind (stance of the philosophy of mind), which treats actions as external phenomena, explained by internal determinants, e.g. beliefs or preferences, while ‘externalism’ is seen more broadly, with the emphasis on interaction between the brain, the body and the external environment (Takizawa, 2017, pp. 523–540).

Voigt’s division manifests some similarities to the idea of institution used by Ostrom & Basurto (2012, pp. 317–343), referring to Commons, who, however, writes about legal formal rules and informal social norms, which, either in separation or in different combinations, function as ‘the rules’, ‘the working rules’ or ‘the rules-in-use’. When Ostrom makes such a distinction between institutions, she does not refer to a specific criterion in the classification of institutions, although these institutions are treated as accepted by society, which can be referred to internal institutions. According to the researchers of her work, Ostrom does not take into account the sources of institutions, but rather their functions (Cole, 2017, pp. 829–847). The division of institutions into internal and external, based on a source of institutions as a criterion, is proposed by Kasper (2002, p. 37). According to his approach, internal institutions arise from spontaneous social interaction, while external institutions are designed and imposed by political agents. This has a certain affinity with Hayek’s concept.

Finally, it is worthwhile to quote the correspondence between Hodgson (2006, pp. 1–25) and North, when North, explaining his division of institutions into formal and informal, defines formal as ‘enforced by the courts or things like that’ and informal as enforced by persons who will bear the costs in the event of failure to comply. To a certain extent, he introduces the criterion of the entity enforcing the compliance with the norms to his division of institutions. As a result, he delineates formal institutions in slightly broader terms than in the orig-
inal definitions, whereas at the same time he does not propose a new division based on the criterion of the enforcing entity. He maintains the division into formal and informal institutions.

4.2. Law origin and its significance for the division of institutions

The discussion of the division of institutions into external and internal according to S. Voigt’s criterion may require the reference to law origins and various legal traditions. If we assume that the state is an entity enforcing external norms, we would also have to assume that the norms are imposed in a top-down manner, which mainly refers to law and determines the rules of law (Voigt, 2012, pp. 262–284). However, it should be noted that not every law is equally imposed in a top-down manner. In general, two legal traditions are distinguished: common law (Anglo-Saxon law, precedent law) and civil law (continental law, codified law), although various national modifications have also developed (e.g. Germanic or Nordic civil law). The civil law system is derived from the Roman law and developed in Western Europe in the Middle Ages (Hicks, 1969, pp. 69–71). It was included in the civil codes in France and Germany in the 19th century. The basic form of legislation in this system is a legal act, while lawmaking consists in determining, based on abstract norms, the legal consequences of real actions. Such regulations require rigorous enforcement, so they adopt the external character.

In contrast, the common law system, which was created in England, developed its own tradition, institutions and principles, only indirectly influenced by the Roman law. In this system, the importance of codes is lower, while the law-making role is played by precedents (Johnson et al., 2017, pp. 646–665). Law-making takes place through single verdicts in specific cases, announced by independent judges in compliance with social principles and values, equipped with a judge’s meticulous justification. Such verdicts become the norms for the future. The common law plays a significant role in this system as it should prevent conflicts of interest between people acting on their own initiative, based on their knowledge, and not under the influence of top-down orders. According to Hayek (1989, p. 21), the freedom of the individual is ensured in this way. These norms, according to the division into external and internal ones, should be classified as external, enforced by the state, but their character is completely different from the norms held in the civil law system, because it takes into account the systems of values of a given society and does not have a strictly limiting function. Moreover, even if the common law system is transferred to remote territories, for example from England to Canada, certain states of the USA, Australia, India and many others (Gregory & Stuart, 2004, p. 209), it is essentially open to local common law, its enforcement is not completely external. The degree to which the law will differ from internal, informal systems of values and norms will vary from country to country. In this situation, the division becomes similarly vague as the distinction between formal and infor-
mal institutions, and therefore unnecessary. The problem becomes even more complex when considering the Islamic law (Greif & Mokyr, 2017, pp. 25–52), in which religious and secular laws co-exist, regulating various spheres of life. It would be difficult to recognize them as being equally enforced by the state, but they can be undoubtedly defined as formalized.

4.3. Case studies

Strong arguments for the validity of the division of institutions into formal and informal are also provided by the examples of atypical institutional solutions that are difficult to classify. Many such examples come from the countries at a relatively low level of development, where the institutional order is mainly shaped by tradition, and the law is sometimes included in it in an unorthodox manner. Four cases are cited below: the first case concerns informal institutions that come from the outside, but are enforced internally; the second involves the enforcement of informal institutions supported by the state; in the third case, the state, as the guardian of the formal order, treats the representatives of the informal sector of the economy as a party in negotiations; the fourth case discusses the effectiveness of formal institutions supported by informal institutions, which ensures a satisfactory level of compliance with formal rules.

The first example refers to the processes observed in African countries, consisting in displacing some informal norms, e.g. traditional beliefs, by other informal norms (although with a greater degree of formalization than the former ones), such as Christian values transferred from Europe by missionaries (Osei-Tutu, 2017, pp. 114–123). These situations show that informal norms can also come from the outside of the system, but they can be enforced by a society that is gradually adopting them. Some communal forms of organizations are replaced by others, still informal (Karacuka, 2018, pp. 57–79). As a result, changes take place within informal institutions, but it would be difficult to see them as internal in this case.

The next example involves one-day oral contracts of employment, popular in countries with an extensive informal sector of the economy, e.g. India (Harris-White, 2017, pp. 1727–1755), especially in the construction industry. They consist in the verbal conclusion of a contract for a given day with the setting of the remuneration that should be paid on the same day. They are completely informal, based on trust. Apart from negative social response, it is not possible to enforce the terms of such a contract in the event of non-compliance. Such a norm would, therefore, be considered internal. Studies conducted in India showed that employers often do not keep the terms of the contract, reducing the amount of remuneration or refusing to pay. This concerned 40% of the surveyed employees. However, over 90% of them did not undertake any action (Naraparaju, 2016, pp. 651–676). The ineffectiveness of informal institutions led to the proposals that the interests of employees on a one-day, completely unregulated, labor market should be protected by formal solutions, such as going
to court or appointing quasi-formal institutions such as labor commissioners. This would require the creation of relevant state regulations, so the compliance with informal agreements, if in breach, would be supported by state intervention. In this case, the internal institutions would have to become external depending on whether the parties involved would comply with the agreement, which makes it impossible to classify such agreements unambiguously. In consequence, the division of institutions into internal and external becomes unstable and does not classify institutions effectively. Also, it does not provide for an intermediate level between internal and external institutions.

Another example of intermediate level institutions in India are the associations representing enterprises in the informal sector of the economy (Harris-White, 2017), more or less registered. They play a key role in regulating urban economy. Business association presidents represent particular industries in liaising with authorities or banks. In this case, we can talk about the hybrid of formal and informal institutions and the informalized manifestation of the state. It would be difficult to define these norms as internal, because they are not only socially enforced, but similarly they are not entirely external (because the state clearly does not require them to be respected, but only honors them).

The final case also comes from India and concerns the national program of public works in rural areas, which has been implemented in stages since the 1990s, aiming at the reduction in unemployment and poverty in villages. It was introduced in the form of a legal act, so it was unambiguously formal and its implementation was controlled by the state, but Marcesse’s (2018, pp. 284–296) research revealed that its implementation would not be possible without the simultaneous observance and enforcement of informal rules such as: reciprocity, non-interference and equity, i.e. external institutions were in fact partially enforced by local communities. Three of the four cited cases concern the labor market, as it is often here that indirect institutions that are difficult to classify emerge.

5. Conclusion

Classifications should serve a better description of the nature of different phenomena. Therefore, some may seem more useful than others, depending on research goals. It seems, however, that some divisions become more widely applicable and are attributed with universality, while the use of others is significantly limited and in practice they tend to complicate rather than facilitate the research process. The division into internal and external institutions proposed by S. Voigt does not offer a clear and unambiguous insight due to the different meaning of the same notions used, for example, by Hayek or Aoki, as well as their intuitive or colloquial sense. Moreover, it does not cover complex institutional solutions developed in some countries. In his division of institutions into formal and informal, North assigns to institutions a different method of enforcement as an additional distinguishing criterion, though not very precise.
This criterion would therefore appear in two different classifications, which constitutes methodological inconsistency.

It is uncertain whether, with the current state of new institutional economics, it will be possible to work out a division of institutions different from the one that is currently dominant, namely the distinction of institutions into formal and informal ones. The shortcomings of this division do not seem to be an obstacle to its application, on the contrary, they show openness to institutional processes that have not yet been researched and cannot be classified in an unambiguous manner.

References


**Acknowledgements**

Author contributions: author has given an approval to the final version of the article.

Funding: this research was funded by the University of Economics in Katowice, Faculty of Economics, Department of Economics statutory sources.

Note: the results of this study were presented at *Second Scientific Conference Institutions: theory and practice* (June, 19–20 2018, Toruń, Poland).