Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN

European Doctorate in Law and Economics

Ciclo XXXV

Settore concorsuale di afferenza: 13/A3

Settore scientifico-disciplinare: SECS-P/03

The Thin Boundary Between Custom and Law – a Law and Economics Approach

Presentata da: **Karol Bolko Zdybel**

Coordinatore Dottorato

prof. Matteo Barigozzi

Supervisore

Prof.dr. S. Voigt
Prof.dr. M.G. Faure LLM

Esame finale anno 2026

Doctoral Committee:

Promotors:	Prof.dr. M.G. Faure LLM Prof.dr. S. Voigt
Other members:	Prof.dr. N.J. Philipsen Prof.mr. L.A.J. Senden Prof.mr.dr. H.S. Taekema

This thesis was written as part of the European Doctorate in Law and Economics programme

An international collaboration between the Universities of Bologna, Hamburg and Rotterdam.
As part of this programme, the thesis has been submitted to the Universities of Bologna, Hamburg and Rotterdam to obtain a doctoral degree.

Acknowledgments

The contents of this thesis have been in development from October 2019 until December 2024. In the first half of this period, there was only a vague idea of what would eventually be written. This idea was initially centered on understanding the survival and contemporary significance of "traditional law" – a concept whose meaning soon proved to be challenging in its own right. This research problem was first suggested to me by Prof. Stefan Voigt in Bologna in Autumn 2019, marking one of my first steps on a journey into law and economics and legal theory. Since then, through the invention and elaboration for which only I am to blame, the topic of the thesis has undergone significant refinement. As will become clear, the process stopped at a point at which the topic's original formulation is barely recognizable. Hopefully, it can still be found interesting.

Individual chapters or sections of the thesis were reviewed and commented on by numerous scholars on several occasions. The commentators were economists (including, first and foremost, institutional economists), law and economics scholars, legal scholars from other paradigms, political scientists, and legal philosophers. Feedback was also received during conferences and meetings organized under the European Doctorate in Law and Economics program in Bologna, Hamburg, Rotterdam, Maastricht, and Rennes, as well as during international conferences and through the peer review process of academic journals. I am deeply grateful for all the support and believe that this feedback significantly improved my work.

The supervisors, Prof. Stefan Voigt and Prof. Michel Faure, provided invaluable assistance while also respecting my preference for thinking and working primarily in solitude. Without their help, guidance, and especially patience, the thesis would not be possible. The quality and style of academic works by Prof. Francesco Parisi were another important, although indirect, source of inspiration and influence.

Akram Zaki, Program Coordinator of the European Doctorate in Law and Economics at the Rotterdam Institute of Law and Economics, assisted with the final stages of preparation and editing of this document prior to its submission. My Ph.D. research has been generously supported by Deutscher Akademischer Austauschdienst (German Academic Exchange Service – DAAD). Needless to say, I am thankful for the support.

Special thanks are owed to the administrators, hosts, and all the other staff who bravely operated Library Genesis (defunct as of February 2025) at least since 2008. Their network of websites was an invaluable repository of academic materials. It that helped millions of students and researchers all around the world over many years, including me while writing this thesis. I hope their project will be reactivated someday.

Last but certainly not least, I am grateful to my family for the unwavering understanding and joy they brought me throughout this journey – one that was undoubtedly exhausting not only for me but especially for them.

Foreword

There is a field of research dedicated to social order emerging without the apparatus typical of present-day nation states: without legislatures with undisputed competence to make laws, without enforcement agencies capable of coercion into compliance, or even without courts having the final say in disputes. In fact, there is more than one field of this type, or at least this field has been explored from multiple angles across various academic disciplines.

Libertarian social theorists study it, when they seek "legal systems very different from ours", as the title of a recent book reads (Friedman et al., 2019). Each of these systems outlined in the book differs from a typical legal system in a contemporary Western democracy in its own unique way. They are typically old, sometimes ancient. However, what unites them all is that they do not utilize the state capacity to an extent even remotely close to what people in modern nation-states are used to. They lack various forms of centralization: for example, centralization of enforcement power or the sense of political legitimacy enjoyed by political officials. Replace with: There is no organized police, no lawmakers, or no political leadership – sometimes none of these at the same time.

Political theorists study this field as well. Sometimes they have an abstract interest in the possibility and stability of social order under anarchy, or the fundamental reasons for which states exist in the first place. However, political scientists are more often interested in the possibility, conditions, and scope of a rules-based order in international relations. This motivation is understandable: anarchy is not a hypothetical possibility but a lived reality it the international sphere. There is no superstate among states. Therefore, understanding the enablers of a peaceful coexistence of states, a coexistence in which states' actions are predictable, orderly, and respectful of each other, are crucial for peace and prosperity. Proponents of the so-called "regime theory" (Bradford, 2008) and "liberal institutionalism" (Abbott, 2008) study the rules-based order among sovereign states and how it can be achieved under anarchy.

Legal anthropologists and students of early legal history are also deeply interested in this field. Early law – i.e., law functioning among mostly illiterate men, archaic law, "primitive law", social order of the historically first human civilizations – could not rely on the state capacity as we understand it today because such capacity was often non-existent. The study of the evolution of early law must take these conditions into account. In the last several decades, there was even an emerging trend among some legal anthropologists to emphasize the absence

of Western-style nation-states in the shaping of the social system of indigenous peoples. This trend, known under the name of "legal pluralism," (see, Griffiths, 1986) offered novelty in terminology rather than advancing our understanding of law. But this emphasis is already telling, because it shows the self-awareness of scholars regarding the structural conditions in which social orders emerge and function.

Naturally, institutional economists and law and economics scholars also study this field. There is a thick strand of literature, both theoretical and empirical, on the protection of property and contract enforcement in circumstances where contemporary institutions of a nation-state are absent (see, Powell and Stringham, 2009). This literature on "private ordering" explores how agents may secure the protection of property rights and facilitate non-simultaneous exchange without the involvement of third parties.

However, despite the apparent overlap, the boundaries of this field can sometimes be blurred. Early in the writing of this thesis, I sought to develop a model of "traditional" or "customary" law based on the methodological foundations of economics: individual choice under constraints. When I attempted to incorporate some of the insights from political scientists, institutional economists, and legal theorists, I quickly encountered difficulties. Apparently, there is more than one way to define terms such as "anarchy," "lawlessness," "spontaneous law," "custom," and "private ordering." In building definitions and developing models, some scholars emphasize the distribution of power in society; others focus on the ways social rules and expectations emerge; still others prioritize the presence or absence of legal officials. What one group of researchers emphasizes is not necessarily the same as what another group envisions, even when they happen to use similar analytical tools.

Nevertheless, interest from scholars working on diverse topics, such as international stability on the one hand and social norms on the other, demonstrates that the aforementioned field has substantial explanatory potential. Therefore, the first part of this thesis outlines the basic building blocks of the field of "spontaneous" social orders. It identifies three such building blocks: three meanings commonly associated with "spontaneous" social order or "spontaneous institutions," each highlighting a specific aspect of decentralization in institutions. The first refers to the decentralized creation of rules, where customs serve as the source of rules. The second focuses on the decentralized enforcement of rules. Finally, the third highlights the absence of officials specialized in announcing, explaining, and clarifying rules.

The result is a classification of institutions studied within this field into several ideal types, which are similar in some respects and different in others. It is also shown that, in the

rational choice framework, these similarities and differences are "structural": they can be represented with the corresponding designs of games in the game-theoretical sense.

Initially, I regarded this procedure as a personal exercise – a method to close gaps in my knowledge of law, legal theory, and legal history, as well as to address the general incompetence of a novice. It was intended as a private starting point, a prerequisite for engaging later with more advanced considerations. But I no longer hold this view. Instead, I have become increasingly confident that the issues addressed in the first part of the thesis are significant in their own right.

The second, longer part of the thesis takes a deeper dive into a specific area of the field in question, examining the different outcomes of two rule-making regimes: custom and design. The focus is on institutions with decentralized, or collective, enforcement, which aligns closely with the realities of societies governed by "early" or "primitive" law, as well as the structural anarchy of the international scene.

The novelty of this part lies primarily in incorporating incomplete information into the analysis. It begins with a basic observation: institutions are not just any rules but rules that operate within the social realm. Therefore, they must be accessible to social actors. They are not merely rules of thumb that help individuals navigate the world independently of what others do, think, and believe. Instead, they must be known and potentially communicated – for example, by parents to children or by someone defending their case in court to the entire audience. They are effective, in part, because others know them, and also know that still others know them. In game-theoretical terms, institutions constitute public knowledge.

At the same time, rules are sustained and enforced only to the extent that relevant social actors – those who sustain and enforce the rules in a given context – find it in their interest to do so. However, whether this condition holds is *not* public knowledge. The reason is simple: the existence of such interest depends on individual preferences, which are inherently private information. Traditional game-theoretical models often start from the opposite premise: that players know the game they are playing, including how much others value specific outcomes. Preferences are known, and incomplete information about preferences is assumed away. Yet, this rarely holds in practice. In fact, the emergence of social rules, and the shape of these rules, often reflects the subjective importance of the interactions governed by them and the value actors assign to the outcomes of these interactions.

Against this backdrop of incomplete information, the thesis analyzes, on the one hand, customary rules – those that emerge from the interactions of agents – and, on the other hand, "designed" rules. The latter term refers to rules that can be accessed in the form of centralized, unified, and explicit normative logic. Such logic aims to fully characterize parties' rights, liberties, powers, and immunities within a given social setting, even though it often falls short of accomplishing this ambition for practical reasons. The most important feature of designed rules is that they are subject to deliberate control, unlike customary rules, which result from "blind" processes of historical evolution. Legislation, of course, is the most natural example of rules that are designed.

Table of contents

Introduction

Some academic works have clear and straightforward topics, with titles that almost perfectly represent the essence of their content. Unfortunately, this is not one of these works, even though it began with such intent. Its initial goal was to explore traditional and customary law using the tools of law and economics, with a particular focus on their persistence in contemporary realities. Why do some societies continue to adhere to legal rules or entire legal systems rooted in oral, religious, or indigenous traditions, claimed to have existed since time immemorial? Why do we observe qualitatively different models of social control across the globe, with informal, tradition-based norms playing a more prominent role in some places than in other places? These were the questions that motivated this thesis in its early stages.

However, this ambition quickly proved overly broad. The fields of institutional economics and law and economics are replete with theories and models of custom, customary law, social norms, and related concepts. These theories form a dense, interconnected network of ideas and involve equally many interrelated modeling techniques. In fact, this field proved to be so intricate and rich that defining a specific research topic and structuring it into a coherent research question quickly became a challenge in itself.

On the one hand, legal scholars use several types of models to explain the emergence of customary law. On the other hand, similar models are often employed by other social scientists to explain the development and stability of social norms. Institutional economists studying primitive law, or the law of pre-literate societies, repurpose these models, as do scholars of international law and institutional economists examining "private ordering." The same is true not only of models but also of other elements of broadly conceived theoretical frameworks. At some point, the overlaps in the way various phenomena are conceptualized became apparent, as did the differences in what these conceptualizations tried to capture. Are customary law, social norms, and primitive law truly the same? A superficial reading of the law and economics literature may suggest so, but scholars who rely less on formal tools, such as economic models or game theory, would very likely disagree.

Over time, the idea of the "spontaneous" development of socially shared rules – i.e., the emergence of rules in the absence of what we conventionally associate with the modern state – began to diverge from the thesis's original motivation and eventually overtook it entirely. Of

course, the idea itself is far from original. Within the broad category of institutions – or socially shared rules that structure human interactions – social scientists have been particularly interested in those characterized by minimal degree of design and centralization. This kind of institutions can be included under the umbrella term of "spontaneous" institutions (see, Luban, 2020).

It seems that there are two primary sources of interest in spontaneous institutions. Some scholars advocate for them on efficiency grounds. They argue that spontaneous institutions may be effective in regulating behavior, in contrast to the commonly held belief that centralized and bureaucratized legal systems, characteristic of the present-day nation state, are essential for social order. At the very least, spontaneous institutions should be considered an alternative (e.g., Kinsella, 1995; Benson, 1990; Friedman, 1989 [1973]). Other scholars keep more distance from the problem. Instead of engaging in advocacy, their contributions identify fields where spontaneous institutions are more economically efficient than state-enacted and state-enforced social orders (e.g., Druzin, 2016; Williamson, 2005; Taylor, 1982).[1]

0.1. Key subjects raised and research questions

Starting from these considerations, the eventual topic and structure of the thesis took shape. First, the thesis lays the groundwork by developing a typology of what is referred to as "spontaneous" institutions – in most general and vague terms, a typology of economic institutions formed or sustained through decentralized collective behavior within a community. This part is conceptual and definitional. It seeks to answer the following research questions:

1) In what sense can institutions be called "spontaneous"?
2) If there are multiple ways to answer this question, what classification of spontaneous institutions emerges?
3) How can it enhance our understanding of spontaneous institutions in the real world?

We proceed in several steps in this part of the thesis. First, we distinguish between three criteria for identifying spontaneous institutions that can be found in the literature – i.e., custom

[1] There is also a non-economic perspective, insisting that spontaneous institutions are ethically desirable (even when economically inferior) because other methods of institution-making lack moral legitimacy. This perspective is left out because it is largely irrelevant to the thesis.

as a source of rules, decentralized structure of enforcement, and the absence of third-party coordination – and elaborate on why these three criteria are in use. In the next step, the criteria are combined into a single classification. This allows us to find "ideal types" of spontaneous institutions – "ideal types" standing for abstract representations of real-world phenomena that emphasize their key structural aspects while disregarding contingencies. The classification into ideal types is subsequently illustrated with numerous examples from legal anthropology, legal history, sociology, and international law. It takes phenomena such as primitive law, social norms, customary international law, or commercial usages, and highlights their structural similarities and differences.

Incidentally, the usefulness of the concept of spontaneous institutions for conceptualizing various contemporary and historical normative orders is also demonstrated. The typology organizes the scattered field of spontaneous institutions into structurally similar clusters, allowing for more systematic thinking based on transparent criteria. A natural follow-up, and the last step, is to propose ideas for representing the ideal types game-theoretically, summarizing the survey of formal representations of spontaneous institutions. This topic closes the typological part of the dissertation.

The remainder of the thesis is more detailed. It starts with the basic observation taken from the so-called focal point theory of law. This theory states that the law makes one of the possible collective responses to a social problem "salient," and therefore consistently chosen. In game-theoretical terms, it sees law as an equilibrium selection mechanism. Should farmers expect ranchers to pay compensation when cattle destroy crops? Should ranchers expect to be required to pay and, preventively, build fences? Or, rather, should farmers build fences themselves because no compensation from ranchers is to be expected? Read the letter of the law, inspect the past decisions of judges, and the answer to these questions will be clear. Law "works" by alleviating uncertainty. It aligns the expectations of multiple individuals, making their actions predictable and the rules reliable. This is especially important in situations permeated by uncertainty about the objectives, goals, and desires of others. Laws provide a shared understanding of behaviors that should be expected and of behaviors that constitute a transgression or violation of norms – this is, in essence, the coordination account of law.

It is easy to realize that this idea is so powerful that it should not be confined to law, at least not in its narrow sense. Law, as we understand it today and also in forms in which it has historically existed, is just one type of institution. Beyond laws, there are rules of different origins or structures that effectively constrain behavior and regulate social interactions. In one

way or another, all institutions play the role of equilibrium selection mechanisms; they guide actors toward one of many possible equilibria. In particular, if law addresses the challenges associated with uncertainty, then all other institutions – spontaneous and non-spontaneous alike – must necessarily address them as well. The difference lies in *how* they do it.

One possible source of uncertainty is the fact that the "structure of the game" in social interactions – most importantly, the outcomes desirable, undesirable, or neutral to each party – is not fully known to anyone. In more formal terms, information is incomplete. For the sake of simplification, incomplete information is often absent from institutional theories. However, this absence can be reasonably contested. In realistic scenarios, individuals have diverse desires and long-term plans, which are often unknown to others or only partially understood. All social rules develop and thrive in such conditions, which is a well-known fact to students of international relations, where the assumption of uncertainty about other actors' intentions is bedrock (e.g., Mearsheimer, 2001).

As will be seen, taking private, idiosyncratic objectives and preferences into account is consequential. If social interactions are to be governed by rules – spontaneous or not – these rules must somehow break the veil of private information and be manifested as public. "Publicity" means that they need to be understood similarly by all parties involved; they need to provide a common understanding of expected, desired, and prohibited behaviors. The challenge associated with incomplete information is to create such a common, public notion while starting from a state of uncertainty and ambiguity about the structure of the social interaction in which agents are engaged.

Mimicking Hayek's (2002 [1968]) phraseology, we call it a discovery challenge. The discovery challenge is related to coordination but different. In technical terms, coordination means that parties choose strategies that, when combined, form an equilibrium. Everybody does what others expect, and everybody's expectations are confirmed by what others do: farmers expect ranchers to pay compensation for damage caused by cattle; ranchers do pay compensation if damage is done; ranchers take precautions by building fences to prevent damage. In turn, the discovery challenge involves determining which strategies *can be combined* into an equilibrium in the first place. If farmers expect ranchers to pay compensation, will they do it? And if they do, will they preventively make fences? Discovery is a task unique to scenarios of incomplete information. Under complete information, this problem is assumed away: the range of possible equilibria is known from the outset, and the difficulty lies only in ensuring that agents produce one of them.

In this context, we study the differences between customary institutions, where rules that are to be followed in equilibrium are discovered by observing past similar situations, and designed institutions, where such rules are deliberately constructed beforehand. In sum, this part of the thesis addresses the following research questions:

4) If information is incomplete (in the above sense), how do customary and designed institutions compare with respect to:

    a. Specificity: which should be expected to be more case-specific, and which more general, and why?

    b. Susceptibility to manipulation: which are more easily manipulated by vested interest, and why?

    c. Flexibility in response to external shocks: which adapt more flexibly to changing circumstances?

    d. Influence of group cohesion: how do both perform, depending on how cohesive and internally connected the groups are in which these institutions operate?

0.2. Making sense of rules, institutions, and social equilibria

To provide the thesis with a proper social-scientific basis, it is essential to begin with definitions before moving to the later parts. Since "institutions" and "rules" will be, in one way or another, in the spotlight throughout the entire text, these concepts should be first defined on the grounds of institutional economics. As is perhaps already apparent, we take "rules", or "social rules" to be equivalent to "economic institutions" – a well-used but still somewhat elusive concept.

Defining institutions poses a challenge in its own right. In institutional economics, institutions are typically understood as shared rules governing social behavior (e.g., Streeck and Thelen, 2005; Knight, 1992; North, 1991). Importantly, the definitional characteristics of institutions include their incentivizing role. Institutions are more than ethical postulates, guidelines stipulating desirable actions, or personal rules of thumb. In the institutional economics framework, rules are considered institutions insofar as they are widely expected to

be implemented in the social realm, thus effectively constraining the behavior of agents. In other words, what qualifies as institutions are "rules in actual or potential use in a community, and not merely rules in form" (Hodgson, 2015:7).

The incentivizing role of institutions makes it convenient to distinguish their two major components: an abstract normative rule and an enforcement mechanism by which compliance with the rule is achieved in social realities (Voigt, 2019). The normative rule specifies an action that is permitted, required, or forbidden in given circumstances (Hodgson, 2006). To again use the landmark ranchers-and-farmers example in which ranchers' cattle may damage farmers' crops, the normative part may read "If you are a rancher and your cattle destroyed crops on farmer's land, pay liability damages amounting to X to the farmer." In turn, the enforcement mechanism represents how agents are motivated to comply with the normative part. Enforcement mechanisms may take various forms such as reputational sanctions, shaming, outcasting, or state-enforced repercussions against ranchers who refuse to abide by the liability rules for cattle trespass; in other contexts, other mechanisms are conceivable. All in all, enforcement mechanisms may generally be interpreted as actions and expectations about actions taken by others that jointly produce compliance with a given rule (Aoki, 2001).

A more fine-grained account of institutions that emphasizes the incentives to comply can be sought in the "governance structure" concept by Kornhauser (2022; 2004). In this view, a governance structure is a broad term encompassing how the following four functions are performed in a society: (i) rule creation, (ii) detection of noncompliance with existing rules, (iii) the application of rules to the assessment of conduct, and (iv) sanctioning.[2] The first function corresponds to the normative part of economic institutions: it pertains to the problem of how rules are devised or discovered. The other three functions specify various aspects of rule implementation: while two governance structures (or two institutions) may carry the same rule into action, they may differ in how (and by whom) compliance is monitored; in how (and by whom) compliance is assessed; and in how (and by whom) non-compliance is sanctioned or compliance rewarded. A political system in which a parliament makes laws, specialized government agencies monitor citizens' behavior, appointed judges adjudicate, and security agencies enforce the rules represents a governance structure in the said sense: a stylized modern legal system of parliamentary democracy.

[2] This use of the term "governance structure" differs from the one employed by Williamson, who defines governance structures as methods of organizing transactions to economize on transaction costs; for example, business firms are such governance structures (e.g., Williamson, 2002).

Several remarks need to be made at this point. First, "rules" in this thesis are differentiated from "norms." Although "norms" is a term sometimes used to denote self-enforcing patterns of group behavior (e.g., Young, 2015), it is used more often to refer to the "emotional and behavioral propensities of individuals" (Elster, 1989). This notion of norms, which identifies them with predispositions toward certain behaviors in specific social interactions, has extensive literature in behavioral economics (e.g., Cardenas, 2011), institutional economics (see Voigt, 2024a, 2024b, for a survey), and evolutionary biology (e.g., Gavrilets and Richerson, 2017; Boyd and Richerson, 2009). Institutions may or may not depend on norms understood as agents' predispositions, i.e., may or may not be agent-sensitive (Hodgson, 2006). Although the thesis is not concerned with norms in this sense, the term "norms" may occasionally appear when the context requires.

In the provisional definition of institutions, we have left aside the controversy of whether it is more helpful to characterize institutions as equilibria in the games social actors play rather than simply as social rules. The theory of institutions and game theory have a complicated relationship. Institutional economists have long discussed how to appropriately define institutions: as rules constraining agents (e.g., Hodgson, 2015; 2006; Ostrom, 2005; North, 1991) or as game-theoretical equilibria (Hindriks and Guala, 2015; Greif and Kingston, 2011; Calvert, 1995). Some scholars take a middle-ground approach, identifying institutions with a set of conditions necessary for equilibrium to materialize, particularly beliefs and expectations regarding other agents' objectives and behavior (Aoki, 2001; Knight, 1992). This perspective is also preferred in this thesis. Importantly, although the definitional differences remain important, all sides of the discussion tend to agree that game theory is a useful (Hodgson, 2015) or perhaps even the optimal (Greif and Kingston, 2011) method of analyzing and understanding institutions.

In fact, one of the leading contemporary approaches to defining institutions explicitly uses the language of game theory. It characterizes institutions as "rules-in-equilibrium," (Hindriks and Guala, 2015; Guala, 2016) which is an idea that unites the old rules-based approach with the game-theoretic view of institutions. Although seemingly complicated, this idea at its core is surprisingly simple. Institutions are defined as an intersection (in the mathematical sense) of rules and equilibria. The intersection does not properly include neither the set of rules or the set of social equilibria: some rules are not institutions, and not all social equilibria are induced or shaped by institutions.

But the delineation of rules that are indeed institutions requires a precise, game-theoretical language. Terms like "effectively guiding and constraining" human interaction, or "rules in actual or potential use in a community" must be given an unambiguous meaning. This is where game theory steps in. According to the rules-in-equilibrium approach, a rule – a statement in the form "in situation X do Y" – is an institution if it simultaneously satisfies two conditions:

i) It is strategic;

ii) If adopted by sufficiently many agents, it can coordinate them onto an equilibrium in the game-theoretical sense.

Being strategic means that it makes a difference to the actions others are willing to take whether someone follows the rule or not. For example, a rule of leaving the house always 15 minutes before the bus arrives, adopted by a precautious passenger, is generally *not* strategic, and therefore cannot be an institution. The bus driver will arrive at and leave the stop irrespective of how any given passenger behaves; likewise, other passengers do not rely on the presence or absence of any single passenger in their voyage. Therefore, a rule of leaving the house 15 minutes before the bus arrival, although probably a useful rule of thumb, does not qualify as an institution. On the other hand, a rule of reprisals (limited response in kind to a violation of international law) in international relations is strategic. Before committing a potentially offensive act toward another state in the international system, rational state leadership will take into account how likely this act would be considered a violation and, if it is, whether reprisals would follow. Thus, reprisals can potentially be institutions on the ground of the condition of being strategic.

The second condition lies at the heart of the rules-in-equilibrium approach. This condition requires that rules are equilibrium-inducing. If the rule is followed with sufficient consistency, then there is little incentive for agents to drift away from following the rule, and therefore it becomes a game-theoretic equilibrium. It satisfies the handbook description of an equilibrium: given the expected course of action of other agents, each agent has no incentive to change his course of action. Because the actions in question are driven by rules, this statement can be rephrased: given the expectation that others will generally follow the rule, each agent has no incentive to drift away from following the rule themselves.

For example, if a bus driver believes that other drivers will abide by the rule of driving on the right, he will also drive on the right. If a newcomer to an elevator believes that those

already inside expect a short greeting from those who enter, this person will likely say "good morning" in order to not be judged impolite. However, if the driver does not believe the others to drive in any particular way, it is unlikely that he would still drive on the right-hand side, even if this is what he was taught his entire career. Likewise, if the newcomer lives in a country where elevator greeting is not expected or even seen as weird, "good morning" is unlikely.[3]

It is noteworthy that there are strategic rules that are not equilibrium-inducing. Take the following, very polite, rule that addresses the problem of who should come first through a door: always come through doors after the other person. If followed consistently, adopting it would block all traffic through any doors, as no one would ever find themselves on the other side, always waiting for the other person to go. However, the rule "ladies first" can work fine when people of two sexes are caught in such a situation, and an even superior rule "outgoing first, ingoing next" also works well if two people are on the opposite sides of a door. Such rules can induce an equilibrium and thus count as institutions.

All in all, game theory seems crucial for institutional theory. Even the definition of institutions, although it can be formulated in a natural language, seems best suited for being expressed in game-theoretical terms. This is why game-theoretical modeling will recur as an underlying, background analytical context throughout all chapters and many of the subchapters in this thesis.

0.3. Plan of the thesis

The thesis is divided into four substantive chapters. Chapter 1 classifies institutions that are, for various reasons, labeled as "spontaneous" in academic literature. As previously mentioned, there are many ways to theorize what it means for institutions to be "spontaneous," generally understood as being void of centralization, deliberation, or design. This chapter develops a typology of such institutions. Based on this typology, the chapter proceeds to discuss numerous real-world examples that demonstrate its utility. Additionally, a game-theoretical interpretation of the typology is provided. This interpretation points out that various types of spontaneous institutions can be, and often are, represented through corresponding game designs.

[3] The author can certify to the fact that both country types exist.

Chapter 2 discusses incomplete information in the context of institutional analysis. According to most conventional accounts, institutions need to overcome two fundamental problems: the incentive problem and the coordination problem. The first refers to ensuring that agents face sufficient incentives to make the institution an equilibrium. The second refers to agents understanding, which one of the many possible equilibria they should play. In Chapter 2, we introduce an additional factor into this framework: the discovery challenge. When information is incomplete, knowledge of possible equilibria is not instantaneous. Instead, they must be discovered. We propose a simple model that simultaneously incorporates the incentive problem, the coordination problem, and the discovery challenge.

Chapter 3 applies the model to analyze customary rules and designed rules. It begins by introducing this distinction from both philosophical and historical perspectives. The chapter then represents both institutional types as (rule-induced) equilibria within the model. We argue that customary rules can emerge when the discovery challenge is addressed in a specific way: a common, objective, and public notion of desirable and undesirable behavior gradually develops as information is revealed over time. Agents contribute to this notion by uncovering bits of information that are essential for the long-term stability of a customary rule. Next, we focus on designed rules. We suggest that the provision of designed rules should be broadly understood as the creation of focal points before the game starts. Whether a given set of rules indeed facilitates playing an equilibrium can be observed relatively quickly.

Chapter 4 offers a comparative analysis of the findings from the previous chapter. It observes that the development of customary rules is better facilitated by simpler strategies. Consequently, we suggest that customary rules tend to be less internally differentiated (or less case-specific) because practices that are easier observed and emulated are more likely to become normatively expected than those that are highly case-specific. When efficiency demands high complexity, designed rules may have an advantage. However, deliberately designed rules effectively serve as third-party coordination devices, making them systemic: their specific components cannot be vetoed or rejected. This indivisibility of incentives to follow designed rules is a double-edged sword. While it allows for potentially higher levels of efficiency, it also makes rule design more susceptible to manipulation by vested interests and errors that may result in unintended adverse consequences. The chapter further argues that customary rules are more responsive to underlying socio-economic conditions and more resilient when these conditions change. Finally, we emphasize the personal nature of customary rules and their dependence on the social distance between the parties involved.

1. Spontaneous institutions – a typology

Summary

The chapter offers a concise typology of spontaneous institutions – i.e., institutions formed or sustained through decentralized collective behavior in a community. It combines three commonly encountered criteria for determining what counts as a spontaneous institution: (i) implicit formation of (customary) rules, as opposed to deliberate design of rules; (ii) lack of third-party enforcement; (iii) lack of third-party assessment of compliance with rules or third-party rule validation, as opposed to the presence of a public interpreter of compliance. The typology is subsequently illustrated with examples derived from legal history, legal anthropology, and international law. Supposedly dissimilar normative systems (e.g., customary international law and primitive law; historically emergent rules of warfare and domestic social norms) are shown to exhibit structural resemblance. Finally, the paper discusses how various types of spontaneous institutions can be represented game-theoretically. In sum, the chapter conceptually organizes the dissipated field of research of spontaneous institutions from the law and economics perspective.

1.1. Introduction

At least for several decades, institutional economists and law and economics scholars have been systematically interested in complex rules of behavior that emerge spontaneously. In most expositions, such rules are placed in contradistinction to those with a clear element of deliberation, design, or involvement of a third-party authority. Spontaneous institutions present themselves as a valuable empirical research topic in various historical, legal, and anthropological contexts. Moreover, spontaneous institutions are often perceived as an alternative to those normative depictions of social organization that emphasize the organizational role of the state. The field is rich: by employing methods from the economics toolkit, many scholars have strived to demonstrate how social actors may produce and sustain rules through self-organization, without relying on state capacity to impose social order (e.g., Mahoney and Sanchirico, 2001; Young, 2001 [1998]; Parisi, 1995; Knight, 1992; Milgrom et al. 1990; Taylor, 1987; 1982; Sugden, 1986; Ullmann-Margalit; see also, Powell and Stringham, 2009, for a survey). Over the years, the relevant literature has grown vast, naturally making navigation within the field increasingly challenging. Many leading concepts have become blurry, having been used, borrowed, and cross-cited by numerous authors over time.

In light of this motivation, this chapter attempts to systematically classify spontaneous institutions: i.e., those formed or sustained through decentralized collective behavior in a community. This will be done by providing a scheme of Weberian "ideal types", i.e., idealized and abstract representations of real-world phenomena (Weber, 1949). The classification into ideal types is based on criteria derived from legal theory and institutional economics; it emphasizes structural similarities and differences between the identified types. Thus, the chapter contributes to law and economics research by presenting a framework that categorizes spontaneous institutions into groups of structurally similar entities. It suggests a structural isomorphism among various types of spontaneous institutions. As a result, the chapter highlights that seemingly dissimilar institutional regimes (e.g., primitive law, customary international law, and domestic social norms) share structural similarities, which enables cross-fertilization of research from a rational choice perspective.

In more concrete terms, the chapter builds on and takes inspiration from the broad and interdisciplinary research on social norms, private ordering, customary law, and extralegal governance, broadly conceived. This literature encompasses an array of related themes from

different epochs and geographies. Spontaneous institutions, as understood in the current chapter, are often evoked in the research of legal history (e.g., Kim, 2021; Kadens, 2012a; Bellomo, 1995; Maine, 1883); colonial legal systems (e.g., Moore, 1986; Hooker, 1975); legal anthropology (e.g., Gutmann and Voigt, 2020; Ndulo, 2011; Igbokwe, 1998; Worby, 1997; Benda-Beckmann, 1981; Pospíšil, 1958; Malinowski, 2017 [1926]); extralegal governance of close-knit communities (e.g., Ellickson, 1986); business and industry self-governance (e.g., Bernstein, 2001; 1992; Ellickson, 1989); anarchic or stateless social environments (e.g., Leeson, 2009; Lesaffer, 2007; Anderson and Hill, 1978); international law and cooperation on the international stage (e.g., Guzman, 2008; Fon and Parisi, 2006; Posner and Goldsmith, 1999), and numerous others. The chapter conceptually organizes this diverse and dissipated field.

The chapter is divided into seven sections. Section 1.2 surveys the literature to identify three approaches to determining what makes institutions deserving the label "spontaneous." On this basis, Section 1.3 proceeds to classify spontaneous institutions into Weberian ideal types. Section 1.4 illustrates the typology with selected examples from legal history, legal anthropology, and international law, emphasizing structural similarities and differences between various real-world institutions. Section 1.5 provides a short game-theoretical interpretation of the typology. Section 1.6 discusses and summarizes the entire chapter. Finally, Section 1.7 concludes.

1.2. Three approaches to delineating spontaneous institutions

As was emphasized, the current chapter focuses only on a specific subclass of economic institutions: those that are spontaneous. Within the broad category of institutions, social scientists have been particularly interested in those where deliberation, design, or centralization is minimal or none. There are many reasons why such institutions become objects of scholarly interest. For example, several authors are enthusiastic about spontaneous institutions because of their alleged efficiency (e.g., Kinsella, 1995; Benson, 1990; Friedman, 1989 [1973]). Others focus on spontaneous institutions to countervail what they consider to be the "centrist" view of domestic law, which claims that the organized modern nation state is the primary, if not only, source of social order, and diminish other (e.g., indigenous) sources (e.g., Griffiths, 1986). These scholars suggest that non-state modes of social order are often omitted from contemporary social sciences, despite the fact that they have historically played and still

continue to play an important role in human societies. Still others study normative regimes of the international scene, where spontaneous institutions play the primary role. Some of them believe that spontaneous institutions can safeguard international peace and stability (e.g., Keohane, 1984); others challenge this view (e.g., Grieco, 1988).

In any case, before any of such claims can be assessed, it is necessary to answer a prerequisite question: how to systematically distinguish spontaneous institutions within the wider class of institutions as such? In other words: which institutions count as spontaneous and why? Because the scholarship on spontaneous institutions has become immense, this question has lived to see many answers. As a result, leading terms and concepts started to overlap. Scholars have begun to notice that the concept of spontaneous institutions and related concepts are used to characterize such supposedly diverse phenomena as "rancher/farmer relations, (…) extralegal contractual relations among wholesale diamond traders, (…) aboriginal customs in Papua New Guinea, (…) [and] the rulemaking procedures of the American Law Institute" (Katz, 1995:1745). It is far from self-evident, then, what the proper topic of such research is, or even whether it is just one of many topics.

Attempts to pinpoint spontaneous institutions touch upon some fundamental philosophical problems, such as the nature of law and its relation to other methods of social control and social organization. For this reason, they depend on the choice of legal philosophy and involve, at least implicitly, a philosophical discussion. Aware of this, the current chapter takes a step back to make a more moderate claim. It suggests that various established criteria for determining what counts as a spontaneous institution can be combined into a compelling typology. This typology can serve as a useful analytical tool for navigating the broadly conceived field of institutions that possess a spontaneous element. An additional advantage of this typology is its "structural" nature. It classifies institutions based on their underlying incentive structures. For this reason, our approach aligns with all, or at least most, legal philosophies. In fact, it aligns even with those – or especially those – that reject the notion of law altogether (so-called eliminativist approaches) in favor of focusing on the incentives and strategic constraints that agents face (Kornhauser, 2015; Nye, 2022). This structural nature of the typology also allows for mapping the ideal types of spontaneous institutions to classes of game-theoretical models that can represent their key qualities.

Specifically, the subsequent parts of this chapter focus on three structural criteria for delineating spontaneous institutions: (i) custom as a source of rules; (ii) decentralized

enforcement or sanctioning; (iii) private interpretation of compliance, or the absence of third-party coordination in interpreting and validating rules.

The first criterion emphasizes the rule creation function of governance structures, or institutions. It considers the customary character of rules as the distinguishing feature of spontaneous institutions. In these institutions, normative standards of behavior are derived from the actual behavior of community members. The second criterion emphasizes the sanctioning function. According to this criterion, spontaneous institutions should be identified with those characterized by a decentralized enforcement structure (e.g., those that rely on reputational sanctions or ostracism as an enforcement mechanism). Finally, the third criterion emphasizes the conduct assessment function. According to this criterion, we should speak of spontaneous institutions when normative standards of behavior are validated and interpreted privately. In other words, there is no public agent that provides a single interpretation of rules, or classifies behavior into compliant and non-compliant with rules. All in all, the three criteria emphasize three layers of decentralization in institutions: decentralized rule creation, decentralized enforcement, and decentralized validation and interpretation.

1.2.1. Implicit rulemaking: Custom as a source of rules

The first approach to delineating spontaneous institutions emphasizes the customary character of rules as its distinctive feature. Custom is depicted as an organic and spontaneous method through which rules emerge rather than being deliberately constructed (see, e.g., Parisi, 2001; Hayek, 2013 [1973]). This is why scholars skeptical of the "state-centric" view of social order, which gives primacy to command-and-control legislation as a source of rules, often look to bottom-up rulemaking mechanisms as possible alternatives (e.g., Cooter, 1994; Griffiths, 1986).

The commonplace meaning of custom comprises long-standing and well-established practices that withstood the test of time – a portrayal with which respected dictionaries agree. The Cambridge Dictionary defines custom as a "way of behaving (…) that has been established for a long time". The Macmillan Dictionary defines it as "something that people do that is traditional or usual". With time, such established ways of behaving may become normatively expected. They are not only behavioral regularities within the group ("People drive on the right-hand side of the road"; "Men let women go first through doors") but also rules ("People should drive on the right-hand side of the road"; "Men should let women go first through doors"). This

basic concept of custom as a source of rules is well known to legal scholarship (see, e.g., Bederman, 2010; Parisi, 2001).

In legal theory, the distinction between custom and other forms of rulemaking (like legislation and precedent) receives special attention. It emphasizes the difference in the method by which rules are formed. Unlike legislation made at a particular moment through an explicit and deliberate act, customs develop within a community "over a period of time by performing certain actions repeatedly in such a way as implicitly to indicate that the members had accepted that they must perform such actions." (Kadens, 2012b:1163). Thus, the conventional legal theory states that custom as a source of law requires two constituents: the existence of practice in a community and the accompanying belief that this practice is obligatory – i.e., *opinio juris* (Bederman, 2010).

Given the significant presence of custom in general legal scholarship, it comes as no surprise that the juxtaposition of custom and legislation is also popular among law and economics scholars. Because the law and economics paradigm combines legal background with economic analysis, it often considers rules as "prices" that agents pay for various behaviors (Cooter and Ulen, 2016). Rule design by a legislator is thus likened to fixing prices in a planned economy; analogously, customary mechanisms for rule-making are pictured as a flexible discovery procedure in which interested actors reveal their preferences for or against a specific rule, thus collectively shaping rules-prices (Cooter, 1994). Although most scholars are aware that the details of the custom formation process are more nuanced and claims about its efficiency in setting right "prices" need to be qualified (see, e.g., Bertolini, 2016), custom as a mechanism for devising rules remains widely used as one of the primary indicators of spontaneity in institutions.

1.2.2. Lack of third-party enforcement

A different approach to delineating spontaneous institutions associates them with institutions functioning under a dissipated distribution of power, i.e., in social settings where sanctions and rewards for various behaviors are administered in a decentralized manner. Under decentralized enforcement, each agent has limited capability to incentivize others and may separately decide on the application of punishment. In other words, the problem of spontaneous institutions is "posed as a pre-state process: How does social cooperation emerge without external enforcement?" (Knight, 1992:173) Institutions whose implementation relies on

decentralized enforcement are contrasted with those enforced by a specialized agent who enjoys a superior power position within the group.

Philosophically, the juxtaposition of decentralized and centralized enforcement is founded on the Austinian positivist notion of law. This notion envisions law as a set of coercive orders. It boils down to a command of a sovereign backed by a threat of punishment or other adverse consequences. What differentiates law from non-legal forms of governance, such as morality or social mores, is its enforcement by a distinct agent (the "sovereign") who applies force in a consistent manner (Schauer, 2015; Austin, 1832). In contrast, non-legal institutions are not backed by centralized sanctions but are sustained through different incentive structures, such as moral conviction, peer pressures, or group mobilization. Echoes of this view can also be noticed in predominant law and economics contributions. In the previously mentioned rules-as-prices approach, legal rules are frequently seen as prices that are *exogenously* imposed on economic agents for taking specific actions (Cooter and Ulen, 2016). The usual problem is to assist the planner in designing such prices to achieve predefined policy objectives.

Against this approach, decentralized enforcement attracted the attention of rational choice scholars who attempted to explain compliance with rules of cooperation in the absence of a central political authority, e.g., in preliterate societies (Benson, 1989; Rasmusen and Hirshleifer, 1989) or under the regime of self-help characteristic of the international scene (Norman and Trachtman, 2005). This interest resulted in the development of models that attempt to specify the conditions under which decentralized sanctioning structures may sustain rules-based social order (Acemoglu and Wolitzky, 2020; Powell and Stringham, 2009 for a survey). It also produced many case studies of rules-based cooperation achieved under conditions of low or no state capacity (e.g. Friedman et al., 2019; Stringham, 2015). Among scholars of international relations and international law, decentralized enforcement is considered the cornerstone of rational-choice theories on rules, norms, and cooperation among sovereign states (e.g., Abbott, 2008; Guzman, 2008).

The emphasis on structure of enforcement as a factor differentiating various forms of social control can also be found in sociology (e.g., Durkheim, 1984 [1893]) and political sciences (e.g., Axelrod, 1986; Taylor, 1982). Durkheim (1984 [1893]) juxtaposes "diffuse" and "organized" sanctions as two fundamental methods of disciplining society. The former are applied by individual community members in a decentralized fashion; the latter are administered by designated persons or corporate entities and are seen as a characteristic feature of legal orders. Taylor (1982:7) studies social order under anarchy by contrasting a situation of

"a limited concentration of force but no means of enforcing collective decisions" with one resembling a monopoly of coercion. In the first scenario, power is

> "dispersed amongst the members of the group; or, the greater the proportion of the group's members involved in solving the collective action problem (e.g., applying sanctions to free riders), the more decentralized the solution. Contrariwise, a solution is centralized to the extent that such involvement is concentrated in the hands of only a few members of the group." (Taylor, 1987:23)

Similarly, Axelrod (1986) models social norms, understood as spontaneously emergent and decentralized means of social control, by introducing an option to voluntarily sanction defectors as the only available means of enforcement.

The exact nature of decentralized enforcement is contextual: it depends on the actions that agents can perform within their technological constraints. Depending on who the agents are and what kind of interaction with others they are involved in, such actions may include, e.g., gossip (Ellickson, 1991), shunning (Gruter, 1986), refusal to share or reciprocate benefits (Malinowski, 2017 [1926]), cessation of cooperation (Guzman, 2008; Bernstein, 2001; 1992), symmetric reciprocation of the offensive act (Barsalou, 2010), or feud (Leeson, 2009; Friedman, 1979). All in all, regardless of the precise nature of decentralized sanctioning, its usage as the sole (or at least primary) means of enforcement is often treated as a criterion that delineates the meaning of spontaneous institutions.

1.2.3. Lack of ex post *third-party coordination (Private interpretation and validation)*

Finally, a factor differentiating spontaneous institutions may be sought in the absence of a third-party assessment of compliance and third-party validation of rules. This criterion is inspired by the coordination account of law (Basu, 2018; Marmor, 2001; Postema, 1982) and, more fundamentally, by Hartian legal positivism (Hart, 1994 [1961]).

According to the coordination account of law, the defining feature of legal orders, as opposed to non-legal institutions, is *ex post* third-party coordination. The role of a third-party coordinator is to provide social actors with clear statements about the validity of legal rules (i.e., whether a rule belongs to the legal system or not) and about compliance (i.e., whether a rule has been violated or not). For instance, a coordinator might declare: "Ranchers have a duty to compensate for crops destroyed by their cattle if there is no fence around the rancher's land, and do not have this duty otherwise" or, "Rancher X has built an enclosure that does not qualify as a fence and must therefore compensate for the damage caused by his cattle."

The statements of a designated third party are conventionally recognized as having special status, which differentiates them from mere opinions of private parties. Individuals may differ in their views about what the rules are and whether rules have been breached in particular cases. For example, witnesses of a traffic accident may voice various opinions about who is responsible, and support these opinions with various interpretations of traffic rules that the individuals involved should have obeyed and various recollections of events. However, statements supplied by a special authority (e.g., a relevant court) override those of private parties. They are conventionally treated as definitive. In other words, this authority has "the capacity to articulate, clarify, and adapt the […] system" of rules in the form of public knowledge (Hadfield and Weingast, 2012:491), and to decide how it should be applied in specific circumstances. In the real world, this role is typically played by various legislators, courts, tribunals, councils, and other functionaries of the legal bureaucracy.

The idea of third-party coordination is closely related to the focal point theory of law. It claims that law serves as a collection of focal points around which members of society coordinate their expectations and behaviors, particularly in situations where multiple equilibrium outcomes are possible. Influenced by Schelling (1980 [1960]) and further developed by McAdams (2009; 2000), this theory suggests that laws help individuals form shared expectations. This, in turn, helps agents to predict the actions of others and accordingly adjust their own, thereby reducing uncertainty and the amount of conflict. For example, while many mutually exclusive methods of acquiring unowned property are conceivable, the law designates a specific method as valid. Thus, agents know which actions result in successful acquisitions of unowned things and which do not, and can be reasonably confident that others have the same understanding of these actions.

However, the coordination account of law and the focal point theory differ in how they conceptualize coordination. The coordination account of law specifically emphasizes the

importance of ongoing coordination. It asserts that a third party plays a proactive role by resolving disputes or disagreements as they arise. Such disagreements may occur when parties have already failed to coordinate, for example, by interpreting rules or facts differently. All parties may know that ranchers are freed from liability if they build fences, but what qualifies as a fence? When such doubts arise, the third party restores coordination by providing definitive judgments about is "right" or "wrong." If subsequent disagreements emerge, it continues to do so until a final conclusion is reached. For this reason, and unlike the regular (ex ante) coordination provided by focal points, ex post coordination requires agency: the coordinator must actively respond to emerging controversies.

In contrast, the same criterion of *ex post* third-party coordination suggests that non-legal institutions rest upon the private assessment of compliance. The classification of actions into compliant and non-compliant is supplied independently by agents based on their private assessment. To frame the issue in Hartian terms, there are no "secondary rules," i.e., rules stipulating who has the mandate to recognize and change other rules, and how this can be done (Hart, 1994 [1960]). For example, although there are multiple elaborate rules of courtesy functioning in contemporary societies (often written down in savoir-vivre manuals), no specialized tribunal exists to validate the behavior of individuals against those rules, or to decide that some rules are outdated and no longer binding. Such validation is performed privately and independently by each agent.

In other words, the lack of third-party assessment of compliance does not preclude the emergence of common normative concepts and shared expectations. Rather, such expectations are developed differently. Instead of being externally supplied, they are established through the "confluence of individual decision-making exercised in the absence of external coordination" (Hadfield and Weingast, 2012:491). According to the private interpretation criterion, institutions whose normative components emanate from such decentralized confluence are classified as cases of spontaneous institutions.

1.3. Typology

So far, the discussion has revolved around the three criteria for delineating spontaneous institutions within the general category of institutions. In the next step, these criteria will be

combined into a systematic classification of Weberian ideal types: abstract, idealized, and simplified notions constructed to represent crucial features of real-world phenomena (Weber, 1949). Because ideal types are characterized by reduced complexity and model-like simplifications, they may be useful tools for navigating infinitely complex realities.

Since the three classification criteria are binary (custom or explicit source of rules; decentralized or centralized enforcement; *ex post* third-party coordination or private assessment), there are eight possible combinations. However, if logical dependencies exist between two or more classification criteria, the actual number of ideal types can be further limited. Indeed, it can be claimed that one such dependency exists between the enforcement method and the presence of *ex post* third-party coordination: the existence of a specialized enforcement agency entails such coordination. It is safe to assume that the actions of the enforcement agency are salient for other agents, because the agency occupies a special position of power in the society. Thus, it publicly communicates what should be considered "right," "wrong," or "neutral," and therefore can coordinate agents' expectations.

With this exception, other possible combinations of the three criteria seem logically plausible, resulting in a total of six ideal types of institutions summarized in Figure 1. The residual type aside, they will be now concisely discussed.

Figure 1

Typology of spontaneous institutions

Custom as a source of rules		
	No *ex post* third-party coordination	*Ex post* third-party coordination
No third-party enforcement	Ambient rules	Pure customary law
Third-party enforcement		Centrally enforced customary law

Explicit formation of rules		
	No *ex post* third-party coordination	*Ex post* third-party coordination
No third-party enforcement	Relational agreements	Private orderings
Third-party enforcement		(residual type)

1.3.1. Ambient rules, pure customary law, and centrally enforced customary law

The upper panel of Figure 1 includes ideal types of institutions that correspond with the dictionary definition of custom – i.e., institutions in which the rule creation role is played by community practice. Within this broader category, the panel contains three ideal types, called ambient rules, pure customary law, and centrally enforced customary law.

Ambient rules refer to institutions that exhibit all three features associated with spontaneous institutions; they would be considered spontaneous according to any of the three criteria. The normative component of ambient rules is derived from the usual pattern of action in the wider social environment. They are enforced through decentralized sanctioning (or rewarding) in a situation of dispersed agents' power. Finally, ambient rules are interpreted and articulated privately. These three properties jointly make ambient rules fully "emergent," in the sense that deliberate rulemaking, centralized enforcement, or third-party assessment of compliance are absent.

Further, the ideal types of pure customary law and centrally enforced customary law encompass institutions that can be aligned with the legalistic notion of customary law. In building definitions of customary law, legal scholars routinely stress the conjunction of the sociological fact of custom-driven behavior and the corresponding elevation of this fact to the status of binding law by legal authorities. Conformity with customary rules is assessed by a specialized official (e.g., a court, tribunal, council of elders, etc.), and thus customary rules are given a single public interpretation.

For example, Parisi describes customary law as "a spontaneous norm which is recognized by the legal system (…) as a proper legal rule" (Parisi, 1998:672).[4] Likewise, Kadens (2013) systematically distinguishes social rules that emerge from spontaneous activity and legally binding customary rules proclaimed by an authorized official. According to this approach, customary law consists of two counterparts: factual behavior-custom and legalistic rule-custom. Behavior-custom is a recurrent behavior in the community. Rule-custom is the variant of said behavior endorsed and authoritatively communicated as a normative standard. Historical and empirical research on the strategies employed by courts to identify customary rules suggests a significant role played by courts and tribunals in shaping the binding variants of customary law (Petersen, 2017; Kadens, 2013).

The difference between the ideal types of pure customary law and centrally enforced customary law can be specified in terms of Hohfeldian analysis (Hohfeld, 1917). In a social environment lacking a monopoly of coercion, a breach of a rule typically gives a subgroup of agents (e.g., the offended party, the cognatic group, the clan, or all interested agents) the freedom to undertake steps toward redress. The role of legal officials comes down to authorizing and legitimizing the subsequent use of decentralized means of enforcement. For example, a court may authorize the kin group of the injured party to enter the injurer's estate and recover damage payments; the rest of the enforcement process is left to the kin group itself. Such enforcement would be characteristic of the ideal type of pure customary law. In contrast, in the case of centrally enforced customary law, a breach confirmed by an authorized third party gives the offended party a claim right to the actions of the enforcement agency. The enforcer, such as the police or specialized agents of the court, becomes burdened with a duty to act on the offended party's behalf (Hoebel, 1967).

[4] Kim (2009, 2007) argues that this concept of customary law is specific to the broadly understood Western legal tradition and thus absent from other legal cultures, such as Far Eastern ones, that have developed well-organized legal systems with strong top-down enforcement.

1.3.2. Relational agreements and private orderings

The bottom panel of Figure 1 includes ideal types of institutions that do not fit the commonplace definition of custom. Nonetheless, two of them exhibit at least one trait associated with spontaneous institutions. These traits are decentralized enforcement and the absence of *ex post* third-party coordination. Institutions with at least one of these qualities are often discussed, among others institutions, in the literature on private and non-legal ordering (e.g., McAdams and Rasmusen, 2007; Dixit, 2004). We refer to the ideal type of an institution where rules are formed explicitly but lack external enforcement and ex-post third-party coordination as a "relational agreement." The second type, characterized by the presence of ex-post coordination, will be referred to as a "private ordering."

The ideal type of a relational agreement corresponds to a situation where explicitly adopted rules (e.g., through pacts, exchanged promises, or authoritative religious texts) are sustained entirely via in-group enforcement mechanisms. Consequently, adherence to the rules depends on the value each group member attaches to the future relationship between the member and the other group members. For this reason, relational agreements structurally resemble relational contracts studied by legal scholars and institutional economists. Relational contracts are contracts in which ongoing contractual performance depends on a mutual belief that the beneficial relationship between the parties will continue (e.g., Dixit, 2004; Macneil, 1980; Tesler, 1980). However, relational contracts typically involve relatively few (typically two) parties. In contrast, relational agreements generate a broader expectation of compliance within wider society. Agents who are expected to abide by the rules are not necessarily the same agents who established them.

Finally, when explicitly adopted rules are adjudicated or interpreted by an authorized social agent, the institution will be labeled as a private ordering. In general, private ordering refers to a concept in which a group of actors decides how to govern and police dealings and interactions among themselves. This involves both establishing rules and in-group regulations as well as resolving potential conflicts between members. In a theoretical paper, Sagy (2011:923) defines private ordering as "a development of extralegal forums and forms of dispute processing by nonhierarchical groups," but suggests that the capacity to deliberately make rules also belongs to this notion.

An important feature of private orderings is the absence of legal sanctions for non-compliance. Decisions of dispute-resolution bodies lack the backing of the state, and therefore,

the only sanctions for non-compliance may be administered by group members. The purest form of private ordering is represented by those forms of self-organization where no state authority is present, even as a fallback option (Renner, 2021). This form, in which agents cannot have recourse to a higher-order authority, will be understood as the ideal type within the context of the typology.

1.4. Filling the matrix: real-world examples

The previous section developed a typology of spontaneous institutions, using existing legal and economic scholarship to distill abstract types of spontaneous institutions that can be formally represented and analyzed. The ideal types summarized in Figure 1 are designed as conceptual shells, ignoring contingencies but capturing key structural features of real-world phenomena. In turn, this section of the chapter aims to make the typology practical by assigning real-world examples to the ideal types. While the correspondence between ideal types and actual institutions can never be perfect, the purpose of providing illustrations is to demonstrate that the typology can be a useful classification tool. Equally important, this mapping reveals structural similarities between various (*prima facie* dissimilar) institutions.

Figure 2

Examples corresponding to the ideal types from Figure 1

Custom as a source of rules		
	No *ex post* third-party coordination	*Ex post* third-party coordination
No third-party enforcement	Social norms; Primitive law without public adjudication; Customs of waging war; Customary international law in the pre-international dispute resolution era	Primitive law with public adjudication; Customary international law in the (post-1945) international dispute resolution era, to the extent that compulsory adjudication is recognized
Third-party enforcement		Diligence standards based on custom in tort law; Commercial custom, when incorporated into the law; Colonial "customary law"

Explicit formation of rules		
	No *ex post* third-party coordination	*Ex post* third-party coordination
No third-party enforcement	International agreements in the pre-international dispute resolution era, or without external dispute resolution	Some international agreements in the international dispute resolution era; Religious laws in diasporas (e.g., among Mennonites); social order of medieval Iceland; private commercial law, e.g., in the diamond and cotton industries
Third-party enforcement		(residual type)

1.4.1. Ambient rules: Social norms, primitive law without adjudication, customary international law in the pre-international dispute resolution era

It is worthy of reminding that ambient rules represent customary rules that are both enforced and interpreted privately between agents. Such institutions are best exemplified in social norms. The rationale for conforming to social norms stems from their historical establishment and sufficiently widespread acceptance in society. Moreover, their continued existence depends on decentralized incentivization by group members. They are transmitted and interpreted privately by individuals, families, or other organic social units. Researchers have acknowledged similar characteristics of social norms when modeling them as emergent and evolutionary phenomena in evolutionary game theory (e.g., Young, 2015; 2008; 2001 [1998]; Sugden, 1986).[5]

Social norms have been extensively studied from the rational choice perspective in various contexts. Such studies often consider social norms as substitutes for legal rules when the latter are too costly or infeasible to establish or enforce. Examples include norms of liability for animal-caused property damage among cattlemen and ranchers in Shasta County, California (Ellickson, 1991); norms of property rights in hunted animals among whale fishers in the North Atlantic Ocean (Ellickson, 1989); norms regulating inheritance of real property in rural Catalonia (Assier-Andrieu, 1983); and footbinding norms in Imperial China (Mackie, 1996). In all these cases, the relevant rules are unwritten, derived from practice, privately interpreted, and enforced through decentralized social sanctions by the community.

Beyond contemporary domestic social norms, other examples of ambient rules can be sought in legal anthropology – more precisely, in primitive law. While the term "primitive law" can be interpreted in various ways, we follow the convention established by A. Diamond in the seminal work *Primitive Law, Past and Present* (1971). According to Diamond, the term refers to social rules that predate the development of formalized state structures and, especially,

[5] Naturally, alternative approaches also exist in the law and economics literature. For example, in their extensive survey paper on social norms from the perspective of law and economics, McAdams and Rasmusen (2007) do not require that social norms are implicitly created by repeated actions of individual agents, nor that they are privately interpreted. Such centrally devised or created norms (e.g., by a professional association) are given the name "organization norms". However, the additional requirement stipulated by McAdams and Rasmusen is that social norms are obeyed because they are at least in part supported by "normative attitudes". These attitudes differentiate social norms from "conventions" that are abided by because of the interplay of purely external incentives, e.g., threats of sanctions by others or preexisting equilibrium-supporting beliefs in the society.

codified laws. In other words, the history of primitive law begins with the advent of humanity and concludes around the time when the first legal codes are established.

Legal historians and anthropologists often observe that a common trait of primitive law is the prevalence of rules embedded in flexible oral traditions or commonly followed practices – in other words, "whatever is regularly or generally done is considered rightly done" (A. Diamond, 1971:164). In the absence of written codes, legislation and precedent are unlikely to emerge as socially approved sources of rules, and thus "the remaining source of law, and the one that dominates primitive law, is custom" (Posner, 1980:31). Additionally, the egalitarian social structure prevents any individual or small group from amassing enough power or wealth to dominate the community (see, e.g., Taylor, 1982; Hoebel, 1967). Under such conditions, the provision of incentives to comply with social rules requires collective participation (e.g., through ostracism or ex-communication) or at least widespread acceptance of the legitimate use of force by others.

The existence or non-existence of public adjudication is a well-known source of variety in primitive law. Organized and ritualized dispute resolution is observed in some non-literate societies, typically those that are more economically advanced, while being absent in others (MacCormack, 1973; A. Diamond, 1971).[6] It is precisely those regimes of primitive law lacking public adjudication – systems of sophisticated social norms in non-literate communities – that can be considered real-world exemplifications of the ideal type of ambient rules.

Anthropologists have long studied regimes of primitive law without public dispute resolution mechanisms. For example, in his pioneering work on the "savage society" of the Melanesian people in the Trobriand Archipelago, Malinowski (2017 [1926]) denies the existence of anything resembling a tribal court for settling disputes:

> "The rare quarrels which occur at times take the form of an exchange of public expostulation (*yakala*) in which the two parties assisted by friends and relatives meet, harangue one another, hurl and hurl back recriminations. Such litigation (...) may be of assistance in settling disputes. Sometimes it seems,

[6] A. Diamond (1971) generalizes his extensive case studies of preliterate legal institutions by suggesting that societies of food gatherers and those in "lower grades" of agriculture or pastoralism typically have no recognizable adjudication institutions. Mechanisms of public dispute settlement are present only in more economically developed primitive societies (e.g., among cattle keepers or advanced agriculturalists), yet even in those cases, their emergence is not universally observable.

> however, only to harden the litigants. In no case is there any definite sentence pronounced by a third party, and agreement is but seldom reached then and there." (Malinowski, 2017 [1926]:85)

In the absence of mandatory litigation, tribe members would willingly resort to reciprocal sanctioning to punish perceived transgressions of community norms. They would refuse to share means of subsistence, cooperate, or associate with someone they considered a wrongdoer. Similarly, MacCormack (1973:77) observes that many African peoples like the Nuer, the Dinka, the Tiv, the Amba, the Konkomba, and the Lugbara had "no chiefs and no courts and not even village headmen" capable of deciding disputes. He reports on complex systems of unwritten rules regulating behavior within a single tribe and relationships between different tribes. A. Diamond emphasizes that among the Nuer, "there are no governmental or judicial organs of people, tribe, tribal sections, village or settlement and nowhere developed leadership except on the part of the lineage head" (A. Diamond, 1971:238).[7] All in all, it can be asserted that institutions of many societies governed by primitive law represent the ideal type of ambient rules: they lack implicit rule formation, third-party enforcement, and commonly recognized legal officials.

Another locus of ambient rules may be found in the international realm. Scholars of international law noticed that customary norms governing international relations structurally resemble social norms in the domestic context (e.g., Norman and Trachtman, 2005) or primitive law (e.g., Barkun, 1968; Dinstein, 1986; Lefkowitz, 2017). Therefore, norms of customary international law can be considered ambient rules. Indeed, the fundamental features of customary international law, as recognized in the literature, include the derivation of norms from the observed regularities in the behavior of states and decentralized enforcement, i.e., enforcement through actions taken by individual states in response to perceived infringements (e.g., Shaw, 2017; Guzman, 2008). According to Guzman (2008), in the absence of a hierarchy in the international system, the enforcement mechanism crucial for the operation of international law consists of "the three Rs of compliance": reputation, reciprocity, and retaliation.

[7] Although the Nuer recognized figures of go-betweens (called leopard-skin chiefs) that alleviate inter-group disputes, they are mere assistants of the conflicted parties. Greuel (1971) argues that the authority of leopard-skin chiefs is founded on the ability to build large coalitions on a case-by-case basis and thus channel the threats of social sanctions against those refusing to make peace with other tribe members.

Like in the case of primitive law, the derivation of normative standards from the past behavior of states and the decentralized method of sanctioning suffice to classify customary international law in two out of the three identified dimensions. The presence or absence of third-party interpretation and validation of rules poses a more complicated challenge. Contemporary it may seem that customary international law is supported by a set of broadly recognized dispute resolution authorities (e.g., the International Court of Justice) that play the role of a public third-party interpreter of compliance. Nevertheless, two caveats need to be made.

First, permanent international dispute resolution authorities are a historically novel phenomenon. For instance, the International Court of Justice (ICJ) was established in 1945, and its predecessor, the Permanent Court of International Justice, was established in 1920. Before their creation, no bureaucratic body could reasonably claim to have the authority to interpret rules of customary international law or adjudicate disputes between states.

Secondly, even after the establishment of the ICJ, there is no mandatory jurisdiction in international law that is independent of states' will. This means that states must consent to the ICJ's jurisdiction, either by recognizing it in advance or agreeing to it in specific disputes. By default, states are free to deny jurisdiction. Only when all parties to a dispute have explicitly recognized the authority of the ICJ can it rule on a contentious issue (see, e.g., Alexandrov, 2006). Nevertheless, many states choose to bind their hands by willingly subordinating themselves to the ICJ's jurisdiction. As of late 2024, 74 states have declared, typically with various reservations, that they accept the "compulsory jurisdiction" of the ICJ in prospective, hypothetical disputes involving other states that have made similar declarations (International Court of Justice, 2024). In effect, these states recognize, among other competences, the power of the ICJ to clarify and interpret customary international law, and to adjudicate disputes between them based on customary international law. Thus, they participate in a system with *ex post* third-party coordination.

However, the participation is far from universal. Currently, there are around 200 states in the world, including 193 members of the United Nations. Both figures are much greater than the number of states that have declared their recognition of the ICJ's compulsory jurisdiction. Thus, fewer than 40% of states globally make such declarations. This figure decreases further when weighted with state power. Only one permanent member of the United Nations Security Council, i.e., the United Kingdom, is among the declaring states. Neither the United States,

China, Russia, nor France – of which the first three have nuclear triad capabilities[8] and are universally regarded as the most powerful actors in the international system – recognizes compulsory jurisdiction by the ICJ.

For these reasons, it should be concluded that historically, i.e., before the advent of the ICJ and the "international dispute resolution era," customary international law represents the ideal type of ambient rules: spontaneous institutions with custom as the primary source of rules, decentralized enforcement, and no third-party coordination (see, e.g., Bederman, 2001). Even today, to the extent to which states hesitate to recognize the authority of international dispute resolution bodies, customary international law retains this historical nature as a collection of ambient rules governing the interactions between international actors.[9]

1.4.2. Pure customary law: Primitive law with adjudicative mechanisms, present-day customary international law, to the extent that compulsory adjudication is recognized

As suggested in the previous section, the ideal type of pure customary law represents an ambient rule that is articulated and interpreted through a public process conducted by a specialized social agent. Consequently, it is reasonable to expect that many real-world institutions exemplifying the ambient rules type may also have a corresponding variant falling under the ideal type of pure customary law. Indeed, it has been noted that primitive legal systems differ in whether they have public dispute resolution mechanisms. Therefore, those with public adjudication can be seen as instances of the ideal type of pure customary law.

The legal organization of the Indian Yurok tribe may serve as an example. Their social system combined traditionalist rules with self-help as the primary enforcement method. This was supported by well-established public adjudication procedures, which legitimized acts of self-help within the community (Benson, 1989; Hoebel, 1967). According to Hoebel (1967:52-

[8] I.e., the capability to launch nuclear strikes from land, air, and sea.

[9] This claim is particularly applicable to the customs regulating warfare. By definition, belligerents normally reject any possibility of external coordination of actions and are not subject to a single superior force capable of resolving the conflict. However, even ongoing hostilities do not preclude the spontaneous development of rules in war. For example, customs that specified the acceptable ways of conducting siege in the early modern era (i.e., in the 16th – 17th century) were created through a learning-by-doing process as a byproduct of siege operations, enforced by the warring parties through adjustments to their future siege tactics vis-à-vis the opposing party (tit-for-tat strategies), and required individual interpretative abilities of the field commanders to be applied consistently and understandably to others (Lesaffer, 2007).

53), Yurok people did not themselves "(…) arraign the offender or determine the extent of the damages to be assessed. This was done by the informal court of "crossers" who were chosen from among nonrelatives living in different communities than those occupied by parties to the litigation." The Yurok procedure was structured and evidence-based. Importantly, the procedure culminated in a definitive verdict that was binding on the parties. If the crossers "found the defendant guilty, [they] declared an explicit judgment against him. (…) [T]he judgment assessed the customary damages against him, which he had to pay over to the plaintiff."

Other similar examples have been extensively reported in legal anthropology. Among the Lango people of northern Uganda, "the only administrative or judicial bodies were the informal gatherings of the village elders to settle intra-village and intervillage disputes (…). But there was no power to enforce their decisions except public opinion" (A. Diamond, 1971:239). Likewise, in the Vogusu and Logoli Bantu tribes, "the main judicial authority is exercised by the old men of the sub-clan, but there is no organized judicial assembly and no means of enforcing a judgment except public opinion" (A. Diamond, 1971:241). The research on "legal pluralism" in developing countries suggests that community-involving adjudication based on customary law continues in the developing world (Igbokwe, 1998; Worby, 1997).

Importantly, in all the mentioned cases, adjudication procedures typically do not involve action by organized enforcement authorities, as they are either absent or weak. Instead, verdicts grant the offended party and their allies the privilege to seek redress on their own. For example, among the Yuroks,

> "in default, the defendant normally became the plaintiff's debtor-slave; otherwise, his execution by the plaintiff and his kin was warranted, although there was risk of engendering feud in this kind of action even though public opinion supported the plaintiff." (Hoebel, 1967:52-53)

In other words, although imperfect, the procedure serves as a coordination mechanism that grants widespread social legitimacy to private sanctioning efforts. Yurok courts of crossers distinguish illegitimate, wrongful acts of violence from legitimate, rightful ones in the eyes of the wider public.

Another example can be cited by continuing the discussion about customary international law from the previous subsection. Indeed, present-day customary international law adjudicated in international tribunals may represent another example of pure customary law as characterized in Figure 1. The derivation of rules from the regular behavior of states and reliance on decentralized sanctions as the primary enforcement tool remain constant features of the international legal order. However, a 20th-century novelty lies in the prominence of permanent international dispute resolution bodies such as the ICJ or the International Tribunal for the Law of the Sea.

The functions of international dispute resolution organizations, emphasized in the law and economics literature, overlap with the characteristics of a third-party assessor of compliance outlined in the preceding section. International courts interpret, clarify, and articulate norms of customary international law, and thereby "assist states to come to a common understanding regarding relevant (...) law" (Guzman, 2008:51-52). They provide focal points within the spectrum of possible interpretations of rules, potentially reducing the effects of states' private, idiosyncratic understandings of rules of international law (Ginsburg and McAdams, 2004). Specifically concerning custom as a source of international law, courts play a role in identifying relevant customs, assessing their binding nature, and delivering clear rulings on their legal implications.

Moreover, by publicly announcing outcomes, international dispute resolution bodies disseminate information about states' behavior to interested parties. Scholars of international law, such as Guzman (2008) and Norman and Trachtman (2005), highlight the crucial role of courts in spreading information, which enhances the effective administration of responses by individual states within the international system. This dissemination facilitates the coordination of sanctions and reduces the costs associated with developing a state's track record, thereby contributing to more accurate reputation-building. However, it must be stressed once again that the classification of present-day customary international law as a system of pure customary law, as understood in Figure 1, is limited by the extent to which states accept the compulsory jurisdiction of courts like the ICJ and the International Tribunal for the Law of the Sea. At present, this acceptance is far from universal.

Nevertheless, history provides further examples of customary international law that differ from how it operates in the contemporary international system, which are marked by the broader acceptance of third-party interpreters and adjudicators. An early case can be observed in the system of rules known as *Leges Marchiarum*. *Leges Marchiarum* was a set of rules which

regulated cross-border interactions during the prolonged borderland conflict between England and Scotland in the 16th century. Despite the warlike conditions, unwritten rules of cooperation emerged organically from these interactions (Leeson, 2009). They constrained violence and unprovoked everyday hostilities. Enforcement of these customary rules relied on private violence, including raiding and hostage-taking.

Significantly, despite the ongoing low-intensity conflict between England and Scotland, *Leges Marchiarum* saw the development of an organized public adjudication forum. This forum aimed to prevent arbitrary acts of vengeance that could escalate into a full-scale war. The dispute resolution body supervised compliance with meta-rules governing raiding and revenge. It ensured that borderland customs, as brutal and tough as they were, did not descend into a cycle of ever-accelerating, alternate violence. The case of *Leges Marchiarum* suggests that third-party coordination extends beyond providing focal points in simple games that approximate pure coordination problems. It can also support equilibria in games where conflicting motives prevail, yet opportunities to avoid unnecessary waste and destruction through cooperation remain present.

1.4.3. Centrally enforced customary law: Customary standards of diligence, enforceable commercial custom, colonial and post-colonial "customary law"

Centrally enforced customary law refers to customs that are authoritatively recognized as legally binding and enforced by a specialized enforcement agency. As such, it can be represented by rules within contemporary domestic law whose content originates from community practices.

An illustration can be found in the adjudicative philosophy underlying the American Uniform Commercial Code (UCC). The UCC, a document that has served as a blueprint for commercial codes in most American state jurisdictions since the 1950s, enshrines the idea of commercial customary law. It directs that courts should look into "immanent business norms," or trade usages practiced among merchants. In litigation, these usages should be interpreted as state-of-the-art conduct expected from the litigating parties. They "vary or qualify the meaning of either trade rules or explicit contractual provisions" (Bernstein, 1996:1777). Court rulings based on this principle are then granted enforcement.

Further examples of centrally enforced customary law may include customary standards of care in American tort law, e.g., in boating accidents or medical malpractice. The standards of care expected in both cases are often determined by courts based on the typical practices within the relevant group of professionals (Peters, 2000; Epstein, 1992).

Beyond the context of contemporary legal systems, the ideal type of centrally enforced customary law aligns conceptually with the construct of "customary law" developed during the colonial era. Colonial customary law encompassed the practices of indigenous peoples recognized by colonial authorities as substantive legal rules applicable within their respective communities and enforceable in courts (Hooker, 1975). Admission of these customary rules typically depended on their compliance with the colonizer's standards of decent morals (they were subject to "repugnancy clauses") and broader political order. Today, similar approaches to customary law persist in many post-colonial countries, particularly in Sub-Saharan Africa (Zenker and Hoehne, 2018).

However, the translation of indigenous practices into legal frameworks foreign to their originators created incentives to modify, reinvent, or abandon old customs and usages. This often resulted in a significant transformation of their original meanings and functions. Therefore, researchers studying colonial and post-colonial legal systems have frequently observed a notable gap between pre-colonial social arrangements and the technocratic customary law created through colonization (Kim, 2009; Moore, 1986; Snyder, 1981). This finding led them to distinguish between custom "pronounced in court judgments, textbooks, and codifications" on the one hand and "living customary law" that consists of "norms that regulate people's daily lives" on the other (Diala, 2017:143).

Critics of incorporating trade usages into commercial law make a similar argument: the involvement of third-party interpreters and enforcers creates a new set of incentives. It inadvertently changes the motivation to develop and sustain customary practices. Most importantly, it transforms customary practices meant to sustain ongoing relationships into endgame rules enforced by a third party – i.e., rules applicable to situations in which the relationship irrevocably deteriorates. The argument runs as follows: agents often engage in informal, customary business practices, such as providing extracontractual services or routinely disregarding contractual deadlines. They do so precisely because these practices are informal and customary – meaning they are not enforceable by a third party. If the relationship between agents deteriorates or if a member of an informal business community no longer finds participation in the community advantageous, withdrawing from these practices becomes a

desirable option. For example, extracontractual services may not be offered any longer, or strict adherence to deadlines may become expected in all circumstances. However, when these informal practices are subjected to third-party enforcement, their role in sustaining relationships loses its meaning, and the incentive to engage in them diminishes (Ben-Shahar, 1999; Bernstein, 1996).

On the other hand, Kostritsky (2006) suggests that at least some trade usages and business norms should be recognized by courts as valid legal rules and enforced by a third party. She argues that courts should evaluate the purpose of each business practice and, based on the purpose assessment, decide whether to incorporate it. However, this criticism seems to abandon the fundamental premise of the "incorporation strategy." According to this premise, insiders, such as merchants and other members of the business community, have superior knowledge and situational awareness compared to courts. For this reason, they are better suited to judge the purposes and functions of trade usages.

In any case, the fact remains that the possibility of some trade usages being granted third-party enforcement changes the underlying incentives to create and sustain such usages. In general, this criticism aligns with the idea that centrally enforced customary law and other ideal types of spontaneous institutions, such as ambient rules and pure customary law, are structurally different, and institutional transitions of one into another are therefore consequential.

1.4.4. Relational agreements and private orderings: International agreements and other agreements under anarchy, self-governance of religious diasporas

In the previous section, we defined relational agreements and private orderings as spontaneous institutions in which rules are not customary in origin but are adopted explicitly. Nevertheless, relational agreements still retain two other features often associated with spontaneous institutions: decentralized sanctioning and the absence of *ex post* third-party coordination.

As such, the ideal type of relational agreements can be represented by international agreements. Agreements between international actors, such as conventions and multilateral treaties, are explicitly formulated and typically put down in writing. However, the underlying feature of the contemporary international order is structural anarchy, and the resulting security

competition between states naturally countervails the possibility that a distinct enforcement agent emerges in the system (Waltz, 1979). Therefore, they lack third-party enforcement. Moreover, as noticed by Guzman, "most international agreements exist without any form of dispute resolution. Agreements might be entirely silent on the question of dispute resolution or might include the singularly unhelpful command that the parties work together to resolve the dispute" (Guzman, 2008:50). Thus, the performance of such agreements lacks external coordination: the parties do not recognize any third party with the authority to interpret or decide questions of compliance.

Of course, in present-day realities of the international realm, the possibility of using an external dispute resolution mechanism is open to willing states, even if not generally used. The Statute of the ICJ in article 36 stipulates that "[t]he states parties to the present Statute may at any time declare that they recognise as compulsory ipso facto and without special agreement, in relation to any other state accepting the same obligation, the jurisdiction of the Court in all legal disputes concerning: a. the interpretation of a treaty." Moreover, some important contemporary international conventions, such as the Law of the Sea Convention, explicitly establish mandatory third-party dispute resolution mechanisms for their signatories (Noyes, 1989). According to our typology, international agreements that use such mechanisms represent the final category in the classification of spontaneous institutions – private orderings.[10]

Naturally, examples of private orderings extend way beyond contemporary international law. An interesting historical case of a society-wide private ordering can be found in the social system of early medieval Iceland, a historical case study often cited by enthusiasts of spontaneous institutions. The period between the 9th century, when the Vikings arrived on the island, and 1262, when they pledged loyalty to Norwegian King Hakon, is often considered a successful example of a "stateless society" (e.g., Geloso and Leeson, 2020; Miller, 1996;

[10] A side note about relational contracts. Relational contracts exhibit characteristics that, with some qualifications, resemble the types of relational agreements or private orderings, as defined above. These contracts are usually understood as agreements whose ongoing execution depends on the mutual willingness of the parties to sustain cooperation in the future (Macneil, 1980). While the terms are typically set in advance, as opposed to being formed over time from historical behavior, relational contracts are often vague and flexible. In response to changing circumstances, the parties involved may frequently overlook imperfections in each other's performance, prioritizing the anticipated value of future cooperation (Campbell and Harris, 1993). Relational contracting is prevalent in various sectors of the modern economy, including the automotive industry (Helper and Henderson, 2014) and construction (Cheung et al., 2006). Since relational contracts avoid third-party enforcement and rely on a tacit understanding of mutual needs and situational specifics, they could, in principle, be classified as examples of relational agreements in the sense specified in the text. However, relational contracts are agent-specific: they are operational only between the involved parties, with no general knowledge of their content or expectation of validity in the wider society. Due to their limited scope and lack of general applicability, they typically do not establish "rules governing *social* behavior" (in other words, institutions).

Friedman, 1979). During this period, the country functioned as a loose confederation of small chieftaincies that were themselves internally divided. Even by the standards of medieval Europe, Icelanders lacked political authority and organized force, having neither dukes nor any equivalent figure of similar power. Therefore, the legal system of medieval Iceland can be considered international law in microcosm.

Within this system, the laws of Iceland were shaped and proclaimed during annual assemblies of freemen (*Althing*) held at a designated place, the Law Rock. A special public officer called a lawspeaker memorized the laws and recited them upon request, particularly during each Law Rock assembly. The lawspeaker's sole role was to accumulate knowledge of Iceland's laws and ensure they were accurately communicated. Beyond their legislative function, the Law Rock meetings heard disputes between individual Vikings or families, effectively imposing obligations on the parties involved. The outcomes of disputes settled during the Law Rock meetings were enforced through a sophisticated system of private vigilantism authorized and regulated by the *Althing*. The Icelandic system attempted "to limit the permissible range of self-help, but it did not try to prohibit it altogether. It sought to limit the class of expiators and the time and place where self-help could be legitimately taken." (Miller, 1996:232) Compliance with a decision made during the *Althing* in one year was verified during the following assembly. Persistent refusals to comply would result in declaring the transgressor an outlaw, to whom no one was allowed to offer help or shelter (Friedman, 1979; Stein-Wilkeshuis, 1986).

In sum, the political system of medieval Iceland combined the three characteristics of the ideal type of private orderings: explicit rulemaking through the *Althing*, with a repository of legal knowledge embodied in the lawspeaker; *ex post* third-party coordination provided by dispute resolution meetings during the *Althing*; the absence of a third-party enforcer and thus the resulting exclusive reliance on decentralized sanctions for enforcement.

Outside the realm of broadly conceived international law, other real-world examples of private orderings can be sought within religious diasporas subject to secular laws of modern nation-states that do not recognize their religious rights and duties. For instance, Mennonites, a branch of Protestants, are obligated by their faith to live in strict accordance with the Dordrecht Confession – a foundational document specifying their religiously permissible lifestyle. The Dordrecht Confession grants Mennonite religious leaders supreme authority in interpreting the document and identifying infringements. Although these leaders lack the power to forcibly subordinate fellow believers, their elevated position within Mennonite communities allows

them to effectively coordinate social sanctions against transgressors (Gruter, 1986). A similar phenomenon has been observed within contemporary Muslim diasporas in England (The British Home Office – Siddiqui et al., 2018) and Germany (Jaraba, 2020).

Finally, well-known case studies of private orderings can be found in the self-organization of various industries, such as New York diamond retailers (Bernstein, 1992) or the American and international cotton trade (Bernstein, 2001). The latter reportedly functions so effectively that it has earned the name "private commercial law." Its distinctive feature is the complete absence of centralized (state-backed) sanctions. Noncompliance with cotton trade rules simply results in expulsion from trader organizations, which usually carries heavy reputational consequences. Aside from this, the system exhibits centralization in rule-making and *ex post* coordination through arbitration tribunals. Rules for the American cotton trade are shaped and proclaimed by several national and regional business associations. Importantly, these rules are not based on commercial customs; in fact, cotton trade rules explicitly prohibit custom from trumping or varying trade rules or explicit contractual provisions. Disputes between traders are resolved by a panel of arbitrators. According to Bernstein (2001:1725)

> "The trade rules are periodically revised to respond to technological advancements, market changes, and ambiguities revealed during disputes. Their content is known and understood by most market participants. The arbitration tribunals that resolve disputes do so expeditiously and inexpensively."

1.5. Game-theoretical interpretation of the typology

Game theory has long been established as an elementary tool of theoretical institutional analysis. In particular, scholars studying economic institutions have developed multiple game-theoretical models of social norms, customary law, relation-based governance, and other phenomena relevant to this chapter, in short: of spontaneous institutions. With some exceptions (notably Dixit, 2004), the modeling techniques typically involve the use of repeated games to

represent various spontaneous institutions.[11] The repeated game design reflects the recurring nature of a social interaction, e.g., a commercial transaction, accidentally caused damage, or a conflict over the ownership of a resource.

This section offers a game-theoretic interpretation of the typology developed in Section 1.3. Its objectives are limited. It does not attempt to construct an all-encompassing model. Instead, this section suggests several design principles that can be useful in modeling various types of spontaneous institutions. The discussion proceeds as follows. First, we outline how the customary formation of rules, decentralized enforcement, and the absence of *ex post* third-party coordination can be (and typically are) represented in game theory models. Second, building on these considerations, we propose principles for designing models that represent the ideal types of spontaneous institutions identified in Section 1.3.

1.5.1. Representing implicit rule formation, decentralized sanctions, and ex post *third-party coordination*

In game-theoretical terms, explicit rule formation can be interpreted as a provision of focal points by a third party before the game unfolds. Such focal points are common

[11] The precise techniques of representing various types of spontaneous institutions differ from subject to subject. Within this variety, two aspects stand out as the key model design features: (i) informational assumptions and (ii) solution concepts.
Depending on the subject and context, researchers of spontaneous institutions use games with complete information (in which players know the full structure of the game) or with incomplete information (in which players lack some knowledge about the game's structure). The assumption of incomplete information corresponds to the observation that the preferences or characteristics of interaction partners in most social settings are not explicitly given to others but need to be discovered in the interaction itself. Likewise, researchers use games with perfect information (in which players have full knowledge of the past events in the game) or imperfect information (in which the knowledge of the past events is limited). The assumption of imperfect information may reflect the idea that episodes from the distant past are unknown or doubtful and only recent events can be recalled with certainty.
Repeated games with perfect and complete information have been used to study customary international law (e.g., Norman and Trachtman, 2005); with imperfect but complete information have been used to study the rules of medieval trade backed by reputational sanctions (Milgrom et al., 1990); with perfect but incomplete information, to study the possibility of a legal order without centralized enforcement (Hadfield and Weingast, 2012).
In turn, the solution concepts employed in the research of spontaneous institutions can be grouped into two categories: perfectly rational and boundedly rational. Under perfect rationality, players make optimal choices given the available information. In the most consistent variant of perfect rationality, players' ability to carry out a strategy is constrained by the requirement of subgame perfection: at every node of the game, a player chooses a locally optimal action. In contrast, such stringent requirements do not apply under bounded rationality. Players do not adapt their choices to information and beliefs but instead use simple decision rules. For instance, they may imitate strategies that have proven successful in preceding periods (e.g., Axelrod, 1986; Skyrms, 1996; Mahoney and Sanchirico, 2001; Young, 2001 [1998], 2015).

knowledge[12] and outline the "plan for the game" for all players, specifying how the game should be played.

Focal points play two roles. For one, they provide agents with guidance about how they should behave when making decisions within the game. For example, in a repeated prisoner's dilemma, preannounced focal points may specify the rules of playing the game by indicating when an agent should cooperate and when to defect. In this way, focal points inform agents' actions. Second, and equally important, focal points shape agents' expectations about how others will behave – i.e., they set expectations for when others will cooperate or defect. However, even if actions and expectations can be preconfigured in this way, it does not guarantee that the plan for the game will be carried out; it can be carried out only if it aligns with the incentives of all agents, meaning it constitutes an equilibrium of the game. Agents must not only know what to do and correctly anticipate others' actions but also need to have the incentives to act accordingly.

Designing focal points may involve adding a pre-game "constitutional" stage (Buchanan and Tullock, 1962) at which agents determine the rules of the subsequent repeated game with a predefined selection method. Conversely, the absence of the constitutional stage corresponds to the formation of customary rules. Customary rules are not designed within any structured procedure but are fully emergent from the players' actions in the game. Expectations are set throughout the game by observing how players act and hypothesizing what they are likely to do in the future.

The easiest way to represent the absence of third-party enforcement is to equip all agents with the same range of possible strategies. This ensures that (many) agents have equal opportunities to punish or reward behaviors they deem undesirable or desirable.

As in most game-theoretic models of decentralized enforcement, punishment should have two key characteristics. First, it must be costly not only for the punished party but also for the punishing party. The disutility experienced by one side of the punishment interaction must be coupled with a disutility experienced by the other side as well. For example, in the case of ostracism, both the ostracizing and the ostracized parties suffer from the termination of their relationship (see, e.g., Rasmusen and Hirshleifer, 1989). Similarly, in a commercial boycott, both the boycotting and boycotted parties lose a trading opportunity (see, e.g., Hadfield and

[12] X is common knowledge if all players know X, know that others know X, know that others know that others know X, etc. *ad infinitum*.

Weingast, 2012). Second, punishment can only influence behavior if it is collective, meaning that a critical mass of punishers is reached.[13] In other words, decentralized sanctions face a collective action problem.

In contradistinction, third-party enforcement requires an introduction of an additional player, i.e., the enforcer, who alters agents' payoffs in ways others cannot. When legal rules are modeled as being enforced by a third party, this player is often considered non-strategic, meaning that they lack a utility function of their own and follow a pre-programmed strategy of rewarding or punishing other players' actions. Such a non-strategic player can be omitted from the game entirely, and their actions can be incorporated directly into other players' payoffs (see, e.g., Picker, 1994). For example, the enforcer may punish defectors in a prisoner's dilemma, which increases the cost of defection regardless of the actions taken by other players. If the enforcer is non-strategic, their activity can be represented simply as lower payoffs earned by defectors. Because third-party enforcement is salient, changes to payoffs made by the enforcer are common knowledge, even if other payoff components may be private information.

Finally, we turn to the absence of *ex post* third-party coordination. As suggested in Section 1.2, third-party coordination refers to the provision of coordination in real time. Even if rules have been firmly established either at a "constitutional stage" or via custom, playing the game may lead to various conflicts. One possible conflict is a dispute over facts. Players may have imperfect knowledge of past events in the game or may be unwilling to share their information with others. For instance, in a prisoner's dilemma, players may not know whether others fully cooperated in the past, and those who failed to cooperate may choose to conceal this fact. Another possible conflict is a conflict over the rules themselves. Even if all parties observed past events, they might interpret them differently. Does my performance count as cooperation or not? What exactly does the rule require in a given situation? This issue becomes especially important when circumstances change. If new situations arise that were not anticipated, which rule should apply? Or how should an existing rule be adapted? (see, McAdams, 2009)

Solving such conflicts as they arise during the game requires an additional player. Unlike the enforcer, this player's actions are payoff-irrelevant. Their role is limited to making announcements that are common knowledge. These announcements can influence the equilibrium of the game solely because they are information-relevant, i.e., players know them

[13] This critical mass may involve only one agent in special cases (e.g., in the case of rivarly among major geopolitical powers, where one power has sufficient capabilities to punish another).

and know that others know them as well. Conversely, in the absence of third-party coordination, conflicts over facts and rules cannot be resolved by appealing to such announcements. Whatever institutions emerge must form an equilibrium in a game without *ex post* third-party coordination.

1.5.2. The structure of spontaneous institutions as the design of stage games

We can now illustrate the structural differences between the ideal types of spontaneous institutions with a simple structure of a repeated game. Imagine a one-off game where:

(i) $N = (1, \dots, n)$ is a set of players,

(ii) A_i (where $i \in N$) is a set of all technologically feasible actions a_i of player i,

(iii) $A = A_1 \times \dots \times A_n$ is a set of all technologically feasible action profiles,

(iv) $U = (u_1, \dots, u_n)$ is a vector of utility functions $u_i: A \to R$ attributing player's i utility value to a given action profile.

Points (i) to (iii) describe the objective structure of a one-off social interaction. They specify the participants of the interaction (e.g., traders, firms, states, etc.) and the combinations of actions that can be taken by the participants.[14] The interaction can be anything that involves a group (or a randomly selected subset) of agents making interdependent choices. For example, a simple two-person game of chicken is often used to represent a conflict between two actors over an ownerless resource (e.g., Sugden, 1986; Ginsburg and McAdams, 2004); two-person stag hunt games are used to illustrate the dilemma between cooperation and security faced by international actors (e.g., Weiss and Agassi, 2020).

Point (iv) characterizes the relationship between players' choices and the utility enjoyed by players. It describes how players value various outcomes: e.g., being a "chicken" while others are also "chickens" or hunting stag while others hunt hare. Many game theory models assume that all utility functions in the vector U are known to all players. In general, this may

[14] This description can be further complicated to account for a broader range of possible scenarios. First, this can be achieved by introducing a temporal dimension to the one-off game, which transforms it into a sequential (but still finite) game. In this setup, agents move in sequence: for example, sellers may act first by making an offer, followed by buyers, who can either accept or reject the deal. Second, some outcomes may be unobservable. Players may observe the outcomes without knowing how other players contributed to them; in other words, actions selected by players can remain private information. In our description, we focus on the simplest case: a simultaneous game (i.e., a game without a temporal structure) where action profiles are observable.

not be the case. The real shape of u_i may depend on player's i private type. In this case, other players hold beliefs about others' types without knowing them with certainty.

The game outlined in points (i) to (iv) will be treated as a basic stage game. Together with the initial beliefs about players' private types (if applicable), this basic stage game can be used to build repeated games corresponding to the ideal types of spontaneous institutions identified in Section 1.3. Beside the basic stage game, there are three optional building blocks of the repeated games. First, there are focal points that can be provided before the repeated game starts. Second, there is Enforcer who can affect the payoffs of other players. Third, there is Coordinator who voices payoff-irrelevant announcements that are public knowledge. The repeated games are depicted in Figure 3.

Figure 3

Game designs corresponding to the items in the typology of spontaneous institutions

	Players	**Pre-game focal points**	**Stage game sequence**
Ambient rules	N	No	(1) Basic stage game
Pure customary law	$N \cup$ Coordinator	No	(1) Basic stage game (2) Coordinator makes announcements
Centrally enforced customary law	$N \cup$ Enforcer	No	(1) Basic stage game (2) Enforcer affects payoffs
Relational agreements	N	Yes	(1) Basic stage game
Private orderings	$N \cup$ Coordinator	Yes	(1) Basic stage game (2) Coordinator makes announcements

Ambient rules are the easiest type to represent in the game-theoretical framework. They can be envisioned as an infinite string of basic stage games. The absence of the "constitutional" stage accentuates their unplanned emergence.

Formal analyses of spontaneous institutions that utilize model designs similar to the model suggested for ambient rules are plenty (e.g., Axelrod, 1986; Sugden, 1986; Young, 2001 [1998]; 2015). The relative simplicity of such models can be counterbalanced by realistic assumptions about the process through which players select their actions or, especially, information available to players. For example, players' discount factors may be treated as private information. Differences in discount factors reflect varying degrees of cooperativeness among agents. The higher the player's discount factor, the more valuable the cooperation in the future becomes relative to the immediate gains from opportunistic behavior. In this setup, adhering to spontaneously emergent rules that we classify as ambient rules can be an indication of a player's future orientation (Posner, 1998).

Moreover, several models of ambient rules assume imperfect information by limiting players' awareness of the past events in the game (e.g., Young 2001 [1998]). This assumption emphasizes the limited cognitive capacity of agents and thus underlies the importance of relatively recent history in shaping expectations about the future behavior of others. For example, Acemoglu and Jackson (2014) study the importance of historically prominent figures in shaping social norms by assuming that, unlike ordinary events, their actions are visible forever. In their overlapping generations model, such figures can divert the pattern of development of social norms driven by the recent past.

In turn, pure customary law by definition embodies a spontaneous rule (i.e., a rule that emerges without a "constitutional" stage) which is explicitly recognized as a legal rule by an authorized agent but enforced in a decentralized manner. In a game-theoretical setting, the event of recognition can be represented through the introduction of an additional player – i.e., Coordinator – that moves alternately with all other players.

The actions available to Coordinator are limited to (payoff-irrelevant) opinionating on the compliance of agents in the preceding iterations of the repeated game. The decision rule used to formulate opinions corresponds to a method of "finding" customary norms among the previous iterations. While the decision rule itself may remain private information, the public character of the assessment might enable a rational reconstruction, making future decisions, at least to a certain extent, predictable in the eyes of other players (see, Hadfield and Weingast, 2012; 2013).

The public assessment by Coordinator plays two important roles: interpretative and informative. Its common knowledge characteristic allows for the coordination of responses to other players' actions and, if future announcements are predicted with certain accuracy,

coordination of expectations. Especially in cases when preferences are private, the announcements can support a coordinated, single view of what the rules are and who has breached them.

Game-theoretical models that emphasize the interpretative role and resemble the design proposed in this subsection can be found, e.g., in papers by Aldashev, Chaara, Platteau, and Wahhaj (2012), and Hadfield and Weingast (2012). The argument in the second paper is that, if payoff-irrelevant announcements by Coordinator possess certain desirable characteristics (i.e., universality, generality, stability, prospectivity, congruence, and uniqueness) and do not diverge significantly from the private preferences of regular players, they can support a perfect Bayesian equilibrium in which players obey the rules because violation would be followed by coordinated social sanctions.

In turn, the informative role is key for the operation of reputation-based enforcement mechanisms. The informative role is theorized in a model by Milgrom, North, and Weingast (1990), in which the only function of Coordinator lies in informing players about others' past behavior. According to this approach, the dissemination of information about the past performance of players is a substitute for long-term bonds between parties. The awareness of the track record allows for the formation of correct beliefs about the intended behavior of trading partners. In consequence, it allows for cooperating only with those who are likely to refrain from cheating. Although careful historical examination suggests that the assumptions of the model do not correspond with the historical realities of medieval trade (e.g., Volckart and Mangels, 1999; Kadens, 2012b), it still can be treated as a generic formal representation of a system of rules characterized by decentralized enforcement supported by a centralized dissemination of information.

We now turn to the ideal type of centrally enforced customary law. In the setting of game theory, centrally enforced customary law can be treated as a version of the previous model of pure customary law. However, there is one important difference: the additional player not only assesses compliance but also affects payoffs accordingly. Such a design reflects the idea that compliance is periodically assessed in courts whose decisions are, in consequence, enforced by third-party agencies or some other specialized entity. Of course, the presence of Enforcer may affect the possible equilibria in the repeated game.

Although not based on formal models, the arguments against the enforcement of business norms developed by Bernstein (1996), Ben-Shahar (1999), and Kostritsky (2006) go along the same lines. Enforcement by a third party strategically influences the rule-creation

process, which makes the ambition to enforce such norms without affecting their substantive content implausible. For example, imagine a contract that specifies a certain price. Further, assume that the payer customarily pays less than this price, and the other party regularly accepts such payments. If payments below the contractual price could become enforceable as a new contractual price, the receiving party would have no incentive to accept them. The receiving party would fear that by consistently accepting lower payments, it would lose the option to demand the full contractual price in the future. Rather, it would risk being penalized by Enforcer for refusing to accept the lower amount. However, if lower payments do not alter the enforceable terms of the written contract, contractual performance would remain more flexible (Ben-Shahar 1999).

Finally, relational agreements and private orderings can be represented in the same way as ambient rules and pure customary law, respectively. The only difference is an addition of a "constitutional" stage at the start of the game. The constitutional stage corresponds to the time in which the rules are established or otherwise presented to the players. Since neither of the two institutional types does not entail centralized enforcement, the further course of the game does not differ from the ambient rules or pure customary law types.

1.6. Discussion and closing remarks

This closing section suggests two important research avenues that can be accurately addressed with the help of the theoretical framework developed in this chapter. The first touches upon the phenomenon of the declining role of spontaneous institutions in legal history and their supersedure with more "rigid" or "formal" modes of social organization. Second, we suggest that the framework can be used to address the problem of the viability of governance by spontaneous institutions, as opposed to state-made and state-administered social order.

Many legal historians have emphasized a transition from spontaneous, unwritten, and (by contemporary standards) informal forms of social control toward more formalized, exact, and professionalized legal mechanisms. Researchers of archaic law also pay careful attention to the evidence of social customs available in documents like the Bible (e.g., Parisi et al., 2020). They assume that subsequent forms of early law bear a structural resemblance to the social

norms from which they supposedly developed. Custom is also recognized as the most important source of law at the beginning of European legal development.

However, the role of custom declined over time, giving way to other sources of rules and social order (Van Caenegem, 1988). This general tendency has been documented in many case studies (e.g., Kadens, 2019; Masferrer, 2019; Thompson, 2015 [1992]; Assier-Andrieu, 1983), and the process seems to continue today. As Epstein (1998:579) notes, "It seems clear that the dominant trend of the past century has been towards the demotion of customary norms and to the rise of what has been called legal centrism." In the European context, the gradual displacement of spontaneous institutions was a multi-step process. Its steps include, but are not limited to, black-lettering (i.e., putting orally transmitting rules into a written form), "homologation" (i.e., standardization of customary norms in a single official document), doctrinal marginalization, and monopolization of lawmaking powers (see, e.g., Glenn, 1997; Dalhuisen, 2008).

The gradual qualitative transformation of law has not been limited to sources of rules (custom *versus* explicit provision of rules). More broadly, the process comprises the departure from "spontaneous" institutions in all three dimensions identified in the preceding sections. For example, Schauer notices that the past two centuries have been marked by a conspicuous centralization of law enforcement capabilities:

> "[t]oday, individuals, businesses, and associations operate within the constraints of the administrative state to a much greater extent than Bentham and Austin could ever have imagined in the nineteenth century. And, importantly, the modern administrative state is an environment of pervasive regulation, with a mass of detailed regulations being enforced by the threat of criminal fines, civil liability, loss of privileges, and a panoply of other sanctions." (Schauer, 2015:44-45)

Likewise, Hadfield and Weingast (2013), Carugati, Hadfield, and Weingast (2015), and, in the international context, Ginsburg and McAdams (2004) describe historical case studies of the emergence of centralized rule-interpreting bodies in environments of loose and ambiguous *de facto* rules.

The suggested law and economics interpretation of this problem suggested in this chapter is relatively simple. The journey from spontaneous institutions to the relative centralization of social control can be separated into three parts: (i) the decline of "implicit" rules, (ii) the centralization of enforcement, and (iii) the solidification of third-party coordination. Any explanation for the qualitative changes in dominant forms of social control should seek to elucidate these three distinct phenomena.

The second topic concerns the viability of governance by spontaneous institutions. Many classically liberal or libertarian political philosophers consider such governance a worthy alternative to state-centric legal order. They maintain that a social system with a minimum or outright nonexistent role of government in making, administering, and enforcing the law is viable and desirable. The fascination with "stateless" law is often visible already in the titles of books, papers, or chapters, such as "Private Creation and Enforcement of Law: A Historical Case" (Friedman, 1979), "The Enterprise of Law: Justice without the State" (Benson, 1990), "Law without the State" (Hadfield and Weingast, 2013), or "The Laws of Lawlessness." (Leeson, 2009)

While opposing the coercive and one-size-fits-all nature of contemporary state-run legal systems, these authors believe that privately created or spontaneously emergent rules may be at least equally efficacious in ordering human interactions (e.g., Friedman, 1989 [1973]; Benson, 1990; Kinsella, 1995). When justifying this position from the law and economics perspective, researchers rely on theoretical models and invoke numerous case studies, both historical and contemporary (see, Powell and Stringham, 2009 for an extensive survey).

However, the attempts to undermine the state-centric view of the law often lack due clarity. It is not always certain what kind of alternative the critics envision and where the emphasis of their arguments lies. This ambiguity is also reflected in the fact that umbrella terms used to denote such alternatives, such as "private ordering" or "private legal orders," cover, in fact, a broad array of various institutions (Katz, 1995). It is far from obvious how they are related to each other and why they should be considered elements of a single category. Yet they all used by the critics of state-centric legal systems to illustrate their argument.

Against this unstructured picture, this chapter suggests a refinement of the normative claim in question. It proposes a fine-grained theoretical framework that decomposes the notion of "legal centrism" into three elements: rulemaking, enforcement, and interpretation. By rephrasing the argument in this framework, the discussion about the possibility of private provision of rules-based governance can gain the clarity that it is currently missing. In other

words, the typology developed in the chapter can be a useful tool for future normative debates about the law from the classical liberal or libertarian standpoint.

All in all, the primary ambition of this chapter was to conceptually organize the mosaic of law and economics research on spontaneous institutions. It is difficult to deny that notions like "spontaneous ordering," "private ordering," "non-legal institutions," "informal institutions," "customary rules," "laws of lawlessness," and many similar ones have spread throughout the academic literature. While they tend to be properly defined in every piece of research where they happen to appear, the sheer amount of accumulated scholarship makes the connections between them unclear. Do they constitute a single body of research? To what extent are they related, and how?

The typology presented in this chapter sketches an answer to these questions. Given the vast volume of existing literature, delivering an all-encompassing picture of the field is implausible. However, the classifications outlined in this chapter may hopefully improve its understanding from the law and economics perspective. It transforms the variety of approaches and perspectives – the very source of internal fuzziness and opaqueness within the field of spontaneous institutions taken as a whole – into a classification tool. This tool is shown to be helpful in seeking structural similarities and differences between various institutions considered, for one reason or another, to be spontaneous.

2. Incomplete information and the discovery challenge in institutions

Summary

It is well established that all institutions must address incentive and coordination problems. To overcome the incentive problem, an institution must constitute a social equilibrium. When rules are in place, agents must find it individually rational to comply, punish those who do not comply, and so on. In turn, solving the coordination problem requires agents to converge on the same social equilibrium from among multiple possible ones. These two challenges are universal to all institutions, whether formal or informal, spontaneous or designed deliberately. This chapter introduces a third, equally critical challenge: discovery. Discovery refers to the process of identifying which social equilibria are viable. It encapsulates the idea that the set of possible social equilibria is not simply "given." Often, it is uncertain whether a particular combination of actions and beliefs can create a stable, rule-governed pattern of behavior within a group. Because agents typically operate with varying degrees of incomplete information, they must first discover potential equilibria, or stable social rules, before they can coordinate on a specific one. The chapter elaborates on the discovery challenge and presents a model of cooperative rules in which this challenge is explicitly incorporated.

2.1. Introduction

How can social rules, or institutions, be developed? Despite the broad nature of this question, it has consistently drawn the attention of scholars across various disciplines – evidence of its longevity and ongoing importance. This chapter revisits this familiar issue in a new way. It takes into account a significant and often omitted factor: incomplete information about the objectives, preferences, and desires of other actors. It begins with a commonsense premise: in realistic scenarios, agents often lack full knowledge of the circumstances relevant to their decisions. Institutions, therefore, effectively develop against the backdrop of such partial knowledge.

The idea behind the chapter can be easily illustrated with simple examples. Imagine a tradesman repeatedly facing a dilemma: should he improve product quality, which may delay delivery, or meet the deadline at the expense of quality? The tradesman's decision depends on how he expects his customers to react, which in turn depends on their private preferences – whether they prioritize quality or punctuality. Similarly, customers may have expectations about whether the tradesman will prioritize timely delivery or focus on quality. How can these expectations be met? Likewise, a similar dynamic arises during a major epidemic. Individuals may limit social activities to reduce transmission, and expect others to do the same. Since people value activities differently and these valuations are often private, the question arises: what rules are likely to emerge? When will avoiding personal contact be anticipated, and when not?

Most formal contributions to institutional analysis focus on the stability and efficiency of rules that are fully developed and complete, often overlooking the *process* of rule creation. Typical models depict institutions as game-theoretical equilibria, sustained by expectations about the behavior of other parties. For example, if travelers expect others to keep to the right side of the road, this expectation reinforces a behavioral pattern where everyone consistently keeps to the right. Importantly, agents are assumed to form expectations with full knowledge of each other's preferences, objectives, and the desirable outcomes for all parties. In other words, they have complete knowledge of the game form – i.e., what others can do and what they want to achieve. This allows agents to identify stable rules, such as traveling on the right, from the outset. However, because multiple equilibria often exist (e.g., traveling on the right or left), institutional analysis must address the question of equilibrium selection: which predefined

equilibrium is most likely to materialize in reality? (Mahoney and Sanchirico, 2001; Knight, 1992).[15]

The logic of selecting between predefined, known equilibria can be summarized in the following statement:

> "If all the actors can focus on something that distinguishes one of the two equilibria, then over time they will be able to establish a regularity in behavior that, when confronted with similar interactions in the future, can serve as a guide to coordinated action. (…) [A]ll of this depends on the actors' initial ability to focus on a salient difference between the two coordination equilibria." (Knight, 1992:100)

In other words, the possible ways in which a social interaction can be structured are assumed to be revealed by default; the problem of establishing institutions is reduced to selecting one of these ways. While this approach is popular and undoubtedly useful in many contexts, it overlooks an important point: agents not only need to select an equilibrium to coordinate on, but also must first discover which equilibria are even possible to play.

More generally, typical applications of game theory to the institutional analysis suffer from the abovementioned weakness. They often assume that all relevant equilibria are known and can be compared, which downplays the challenge of recognizing (or discovering) an equilibrium in the first place. The task of discovering the equilibrium is deliberately simplified by displaying the entire game structure, including player payoffs, from the outset. This approach removes considerations like this: "If I behave in a certain way, would others understand and

[15] Aoki (2001) suggests that there are two principal methods for analyzing social rules or economic institutions using game theory. One approach uses variants and refinements of the subgame-perfect solution concept, according to which agents choose what they believe to be the optimal course of action at each decision point. The other approach represents institutions as fixed strategies employed by different segments of the agent population. These strategies may differ in efficiency and in how well they perform when clashed with other strategies – two factors co-determining the rate at which each strategy is reproduced in the population. Over time, certain strategies may come to dominate the agent ecology. The subgame-perfect equilibrium and related concepts assume that agents are fully rational, but they typically rely on the assumption of complete information. On the other hand, evolutionary models allow for some flexibility regarding information but treat agents as following a predetermined program, making them less suitable for representing rational agents similar to human beings. In contrast, this thesis retains subgame perfection (and thus strict rationality) while introducing the more realistic assumption of incomplete information.

appreciate it, or might they misinterpret or disapprove?" Answers to such questions are normally assumed to be already known by both the analyst and the agents involved.[16]

The remainder of this chapter builds on this basic observation. In contrast to the logic founded on the assumption of complete information, we consider a realistic scenario in which agents have limited knowledge about outcomes others consider desirable or undesirable. This approach aligns with the observation by Hayek (1945) that economically useful knowledge is typically dispersed in society among heterogeneous individuals. Consequently, effectively utilizing this knowledge poses a challenge in its own right. Disclosure and transmission of economically relevant knowledge in society are not automatic and must be incentivized. Hayek argued that the system of market prices serves as a mechanism through which this knowledge is shared and communicated; without it, such knowledge cannot be effectively utilized. Similarly, we claim that the processes by which institutions are established face the same challenge. How this discovery challenge is addressed may influence the content and properties of the resulting institutions. In other words, the process of discovering social equilibria may be consequential for the shape of social rules.

The chapter is structured as follows. Section 2.2 discusses, from the rational-choice perspective, three interconnected obstacles that all institutions must overcome: the incentive coordination, the incentive problem, and what we call "the discovery challenge." In Section 2.3, we employ a game-theoretical framework to illustrate our argument. We use a model of a repeated game with incomplete information to account for less-than-perfect knowledge that agents have about each other. Section 2.4 briefly showcases the special case of the model: complete information. This allows for a better understanding of the consequences of incomplete information. Finally, Section 2.5 bridges this chapter with the following chapters. It introduces two stylized processes of institution-making – custom and deliberate design – and elaborates on the philosophical and historical significance of distinguishing between these two. The next chapters will then use the game-theoretical model to discuss custom and deliberate design in more detail.

[16] There are some exceptions to this approach. To emphasize the subjective and perceptional origin of any equilibrium, Aoki (2001) introduces the concept of subjective game models – i.e., agents' perceptions of the game at any given time. This idea comes as close as possible to the one expressed in this part of the thesis.

2.2. Coordination problem, incentive problem, and discovery challenge

We start with the following observation: it is now conventional wisdom among scholars that a mixture of coordination and incentive problems is inherent in all attempts to establish social cooperation based on predictable rules (e.g., Hadfield and Weingast, 2012; Hindriks and Guala, 2015; Bertolini, 2016).

The coordination problem arises from the fact that the same social interaction may have multiple equilibrium solutions. Rules can successfully govern the interaction only when agents expect others to follow the same unique rule – that is, play the same equilibrium. How can it be ensured that all agents select a unique equilibrium? This, in essence, is the coordination problem.

We can simplify things by focusing only on social dilemmas. Social dilemmas are situations of conflict between agents' personal interests and the collective utility of the group. They are best exemplified by the so-called Prisoner's Dilemma, but the argument can be naturally extended to other scenarios as well. Social dilemmas may have cooperative solutions, in which individuals resist the temptation of pursuing immediate personal gains and instead cooperate in the creation of group utility. However, such cooperation requires a logically antecedent common notion of what cooperation means. When is cooperation expected? What kind of behavior qualifies as cooperation? What should be the response if someone fails to cooperate? These and similar questions need a commonly known answer before agents can confidently engage in cooperative endeavors. According to McAdams,

> "[T]he PD (Prisoner's Dilemma – KBZ) is usually complicated by the fact that there is more than one plausible way for the parties to go about cooperating. If the parties attempt to cooperate using different understandings of cooperation, then it is likely that at some point a party will engage in behavior she believes is cooperative but that the other side views as non-cooperative. One side punishes what it views as defection, the other side views the punishment as unjustified defection requiring retaliatory defection." (McAdams, 2009:22)

Likewise, consider the famous analysis by Coase (1960). Coase suggests that cattle trespass can be addressed through various liability regimes – some prompting ranchers to fence in their cattle, and others prompting farmers to fence out their crops. However, any such regime can only be effective if ranchers and farmers expect the same liability rule to be applied. While it can be argued that both ranchers and farmers prefer some liability rule over the chaos that would result from the absence of any rule, there must be a shared understanding and expectation of what that rule is. Otherwise, no rule could guide the behavior of both groups, and ranchers and farmers cannot rationally expect each other to comply, leading to ongoing conflict over who owes what and to whom.

Coordination represents such a vital topic in institutional analysis that it attracted the attention of philosophers. Postema breaks down the coordination problems by claiming that it consists of four major elements. These are:

1) "Strategic Interaction: The outcomes of the parties are jointly determined by the actions of all; so the outcome of the action of any agent depends on the actions of all the others, and the best choice for each depends on what he expects the others to do, knowing that each of the others is trying to guess what he is likely to do (…)
2) Rough Coincidence of Interests: Each party is likely to benefit more by cooperation than by noncooperation
3) Mutually Conditional Preferences: Certain actions are preferred to others if, but only if, other parties also prefer them (or appropriately corresponding actions).
4) Ambiguity: There are at least two combinations of the actions of all the agents which each agent would count as <<successful>> coordination." Postema (1982:173-174)

In turn, the incentive problem pertains to a different challenge: even when a common notion of cooperation exists, individuals may have insufficient incentives to cooperate. For instance, ranchers who are required to compensate for crops destroyed by cattle may be unwilling to do so because they expect no (or only mild) penalty. In this case, cooperation requires punishments against non-cooperators, e.g., in the form of commercial boycotts of

hesitant ranchers. Still, the incentive problem does not end here. Boycotting farmers might introduce higher-order incentive problems: those who are supposed to boycott may lack sufficient incentives themselves, which calls for punishing those who fail to punish, and so on (see, e.g., Axelrod, 1986).

Much of the early game-theoretical scholarship on institutions has specifically focused on how the incentive problem can be addressed. Even today, a large part of research on institutions focuses on the conditions and incentives that foster compliance (see, e.g., McAdams and Rasmusen, 2007). This strand of scholarship recognizes that the presence of opportunistic incentives, such as incentives to cheat or to avoid punishing cheaters when such punishing is costly, is a factor that may preclude cooperative behavior. For example, in the textbook repeated prisoner's dilemma scenarios, the analysis revolves around identifying conditions under which players consistently choose cooperation over defection.

Many of the formal modeling techniques in institutional analysis have specifically focused on the incentive problem. They assumed that agents face incentives to defect while possessing ideal knowledge of the game structure (e.g., Ullman-Margalit, 1977; Sugden, 1989; Cooter, 1996; Young, 2001 [1998]). For instance, agents may be imperfectly monitored, in which case undetected cheating becomes tempting, or they may face the last-period problem, in which case there are no future gains to justify current cooperation. However, it is notable that concepts like "cooperation," "defection," "benefit," "harm," etc., are already predefined within the game structure, which, by assumption, is known to all. Consequently, the challenge of establishing a shared understanding of desirable and undesirable behavior – i.e., of establishing what terms like "cooperation" and "defection" mean – is treated as a separate problem.

While such separation of the incentive and coordination problems may be useful for analytical purposes, they are, in fact, intertwined. Consider a scenario where a buyer and a seller agree on the purchase of a commodity. The buyer makes the payment, and the seller delivers the goods. However, the outcome may deviate from the parties' original expectations: the buyer may find the quality unsatisfactory, the delivery may take longer than expected, the granularity may be nonstandard, and so on. The buyer might try to pressure the seller into improving what they consider a flaw, for example, by organizing a boycott with other buyers. However, buyers may be tempted not to participate in the boycott. Why stop purchasing from the seller, especially when someone else was harmed? This dilemma represents the incentive problem in our scenario: buyers may lack the incentive to punish an underperforming seller, which may lead the seller to consistently underperform. Yet, there is also a clear coordination problem. Which

events should trigger a collective boycott? How should buyers identify seller's performances that warrant costly collective punishment?

In other words, establishing institutions involves a dual problem of coordination and incentives. It combines the challenge of coordinating the beliefs and actions of multiple agents (i.e., the provision of a shared understanding of what it means to cooperate) with the challenge of preventing defection (i.e., the provision of incentives to cooperate). In the last two decades, the connection between coordination and incentive problems has been emphasized by many law and economics scholars as well as other social scientists (e.g., Platteau, 2008; McAdams, 2009; Hadfield and Weingast, 2012; Bertolini, 2016; Bicchieri et al., 2023). For instance, Hadfield and Weingast state that when agents are heterogeneous,

> "[a]chieving deterrence (…) requires coordinating collective punishment in response to particular actions. This presents two essential problems. First, because each potential punisher has an idiosyncratic logic for assessing wrongfulness, none are able to determine unilaterally when to punish in response to possible rule violations. (…) Second, because punishment is individually costly, punishers need an incentive to punish." (Hadfield and Weingast, 2013:9)

Likewise, Bertolini characterizes the difference between coordination and incentive problems by stating that:

> "[c]oordination problems arise from the necessity of coordinating individual decisions to punish (…). Assuming that people have incentives to bear the costs of enforcing the norm, coordination problems involve determining how multiple, simultaneous individual decisions to punish can be coordinated to generate a coherent and predictable enforcement process. In comparison, incentive problems arise when self-interested individuals are unwilling (because they have no incentive) to bear the costs of punishing the norm violators." (Bertolini, 2016:16)

The interconnection between the incentive and coordination problems brings both elegance and analytical clarity. When information is complete, demonstrating how the incentive and coordination problems are solved is sufficient for a comprehensive analytical description of an institution. No further rationale is needed to explain why a given rule is operational within a group of rational agents. Agents' expectations and actions are aligned: each party expects others to act as intended, and each has sufficient incentive to carry out their intended actions. This logic is illustrated by a simple assurance game (also known as a trust dilemma or a stag hunt game), as shown in Figure 4 below.

Figure 4

Example: Incentive and coordination problems in the assurance game

		Agent B: Trust	Agent B: Safety
Agent A: Trust		2, 2	0, 1
Agent A: Safety		1, 0	1, 1

Imagine that Agent A and Agent B can trust each other and collaborate. In this case, both agents will specialize in what each of them does best. If they work together and specialize, each will achieve the payoff of 2. However, their specialized abilities are valuable only if they indeed choose to work together; without the other party's input, these abilities are worthless, and the time spent developing them is wasted. On the other hand, if an agent does not specialize and

decides to work alone instead, a payoff of 1 is guaranteed. This choice represents safety: regardless of the other party's decision, refusing to specialize ensures this payoff.

As is well known from elementary game theory, the assurance game has two equilibria, highlighted in grey in Figure 4. The first equilibrium represents a scenario of mutual trust: if Agent A and Agent B are convinced that the other will choose trust, they both do so. Importantly, these convictions are interdependent: Agent A can believe that Agent B will choose trust only if he is convinced that Agent B already believes Agent A will choose trust, and so on. In other words, this equilibrium requires a coordinated set of beliefs: a shared understanding of how the situation should be resolved. However, such a common notion alone is insufficient. For example, if Agent's A payoff for collaboration were only 1 instead of 2, this agent would have no incentive to trust, and mutual trust would not occur. Thus, the existence of incentives to trust, provided the other party already does the same, is equally crucial.

The second equilibrium occurs when both agents choose safety. The reasoning here is similar to the previous case: if Agent A and Agent B believe the other will choose safety, they will each choose safety as well. Again, this outcome arises from two factors: first, the belief that the other party will choose safety instead of trust; second, the incentive each agent has to choose safety if they expect the other to do the same.

Until now, we have assumed that Agent A and Agent B have a full picture of their interaction: they know each other's payoffs resulting from various outcomes in the game. Under this assumption, the simple analysis can be naturally extended to much more complicated cases. However, what if the agents do not have a complete view of the interaction – technically speaking, what if they have incomplete information? Does solving the coordination problem and the incentive problem suffice to analyze an institution? We suggest that the answer to this question is no. Under incomplete information, a new challenge emerges beyond the incentive and coordination problems. When payoffs are private, neither party knows the full structure of the game. This uncertainty means that they cannot determine what constitutes a social equilibrium. Figure 5 below illustrates this situation from Agent's A perspective:

Figure 5

Example: Incomplete information about the game structure

		Agent B: Trust	Agent B: Safety
Agent A: Trust		2 a	0 b
Agent A: Safety		1 c	1 d

In the scenario shown in Figure 5, the values of a, b, c, and d are unknown to Agent A. These values represent the payoffs Agent B receives from various outcomes in the trust-safety interaction. They reflect Agent's B private valuation of the game's outcomes. These valuations can be shaped by factors unknown to Agent A, such as Agent's B value system, long-term objectives, hidden desires, needs of the current moment, and many others. Because a, b, c, and d are private information, Agent A cannot determine which combinations of actions by both agents constitute an equilibrium. And when no equilibrium is known, none can be selected, making coordination impossible. Even if a common notion of how the game should be played was provided to both agents, Agent A lacks sufficient information to assess whether Agent B will rationally abide by this notion.

Thus, incomplete information introduces a new dimension to institutional analysis: the need to identify equilibria before agents can act upon them. This process of determining whether a pattern of collective behavior constitutes an equilibrium is referred to as the "discovery challenge."

Naturally, the significance of the discovery challenge extends beyond simple textbook examples like assurance games. The same reasoning can also be applied to more complex, real-world institutions. For instance, imagine that enforcing any liability rule for cattle trespass depends on a collective boycott within the rancher and farmer community. In such cases, certain liability rules may prove infeasible. Community members might find it too costly or impractical to implement a rule protecting an activity (ranching or farming) they do not particularly value. Before implementing any liability rule, ranchers and farmers must first determine whether it is feasible. More broadly, when agents have private information relevant to their decision-making, establishing cooperative rules involves a challenge beyond incentives and coordination: the discovery challenge.

Thus, we have come full circle. First, we reiterated the received wisdom: institutions must function as game-theoretical equilibria. They need to align with individuals' incentives and beliefs; otherwise, they cannot be implemented in the real world as mechanisms to structure, constrain, and organize human interactions. This is the incentive problem. However, since multiple equilibria often exist, a way to coordinate on a unique equilibrium – i.e., a solution to the coordination problem – must also be provided. Yet before solving the coordination problem, an equilibrium must first be available to choose from. The availability of equilibria depends on their discovery; they must be identified as rules compatible with incentives and beliefs or, in other words, as combinations of actions and beliefs that resolve the incentive problem.

Another way to view the connection between incentives, coordination, and discovery is to consider it as a tension between "private" and "public" rules of behavior. We suggest that institutions must be "public" in the sense of being accessible to everyone, as opposed to "rules that only I follow" or "rules only for me," such as personal rules of thumb. Private rules cannot create a broader, social expectation of compliance. In contrast, public rules provide an objective framework for how the game should be played, independent of private, idiosyncratic opinions or preferences. That being said, private preferences are far from irrelevant. On the contrary, they play a crucial role in shaping institutions: only institutions consistent with the incentive structures of enforcers can be sustained. Discovering whether such consistency exists is a vital step in shaping institutions. The next section introduces a more formal way of expressing this argument.

2.3. Model

This section introduces a formal framework for analyzing institutions under incomplete information. We use a simple game-theoretical model to examine the three problems discussed earlier: the incentive problem, the coordination problem, and the discovery challenge. The model is inspired by a model by Hadfield and Weingast (2012); it depicts a repeated social dilemma in which difficulties posed by heterogeneous agents having private preferences are explicitly incorporated.

The model begins with a relatively straightforward social dilemma. Agents can produce cooperative outcomes if all exert costly effort. However, only one agent benefits from collective effort at any given time, which creates a temptation for others to refrain from making the effort – hence the incentive problem. Moreover, agents face different versions, called *variants*, of the same underlying social dilemma. Each agent can benefit from collective effort exerted in only some of these variants. To achieve benefits from cooperation, agents need a common notion of the variants that should be collectively addressed: a shared notion of desirable and undesirable behavior. In other words, they need to understand when effort is expected. Because many possibilities may exist, this creates a coordination problem. Finally, to account for agent-specific factors that influence decision-making, the model assumes that agents' preferences regarding the variants of the social dilemma that each wishes to address collectively are private information. Therefore, to cooperate based on any shared notion of desirable and undesirable behavior, agents need to determine which cooperation possibilities are feasible – i.e., they must overcome the discovery challenge.

Our model assumes heterogeneous agents with incomplete knowledge of others' private motivations, goals, and plans. For such agents to cooperate, they must share a common, public understanding of what cooperation entails – i.e., must have institutions, or public, interpersonal rules. As suggested above, we particularly strongly highlight the distinction between private motivations, such as personal values, objectives, or hidden agendas, and public, interpersonal rules that are clear and understandable to others. Private motivations cannot provide a foundation for stable cooperation because they are hidden from external parties, making it impossible for others to predict behavior. In contrast, rules for social use must be publicly accessible. Only then can they be understood and followed by those without access to private information about individual participants.

Consider state behavior on the international stage as an example of the model's logic. International relations scholars recognize that states' objectives are an amalgam of goals, agendas, and ambitions shaped by factors such as security concerns, ideologies, religions, internal power struggles, and more (Wendt, 1999). Some of these objectives are communicated to other international actors as official political programs; others are deliberately concealed; still others – like those stemming from ideologies – are only remotely announced, often remain vague, and are prone to misinterpretation. Moreover, states vary in dominant ideologies, prevailing religions, domestic political systems, understanding of security by their leaders, and other ways. As a result, it can be difficult to formulate a general expectation about how states will act in a given situation, such as in war, a diplomatic conflict, or a major natural disaster, based solely on their individual characteristics. This creates a world in which state actors are irrevocably distrustful of each other. Therefore, consistent with the logic of realism in international relations theory (e.g., Waltz, 1979; Mearsheimer, 2001), states would take every immediate opportunity to advance their interests as they perceive them. War, diplomacy, or natural disaster management know no rules, only interests.

However, norms of international law can be understood as a means of overcoming this particularism. According to the so-called English school of international relations theory, when international law or rules of an "international society" (Bull, 1977) emerge (if they do, which is not certain), states acquire a collective understanding of norms and rules. These rules and norms are conceptualized as existing independently from the preferences or agendas of each individual state. In other words, they are binding regardless of state-specific factors, such as its ideology, religion, political system, or persons in power.

Thus, the previous two paragraphs can be summarized as follows. Whatever rules or agendas agents privately follow is a different matter than a social system based on public rules. What matters for the latter is a shared, public notion of cooperation – a notion that can be comprehended, elucidated, communicated to others, and replicated in action. This aspect of institutional analysis is often undertheorized in game-theoretical models of institutions. Unlike our model, most other game-theoretical accounts of institutions assume complete information, and therefore do not address the gap between private preferences and public rules.

However, there is a secondary, yet significant, reason to assume that agents are heterogeneous and have limited knowledge about each other's private preferences. A well-known critique of rational-choice analysis is that it supposedly overlooks non-egoistic motivations in human behavior. Economists are accused of focusing on the analysis of self-

interested agents who are concerned exclusively with satisfying their own desires (see, e.g., Sen, 1977; Sunstein et al., 1998). However, other factors clearly play a major role in decision-making. Values, internalized standards of rightness or decency, honor, moral judgments, philosophical reflection, and even aesthetical tastes can all be relevant to human choices. Some of these factors are so significant that they can override egoistic motivations. For instance, a company owner might choose to pay higher-than-market wages because they prioritize fairness toward workers over profit maximization. In such cases, personal gain in the form of profit is willingly sacrificed for the well-being of others.

The inclusion of various variants of underlying social dilemmas – i.e., various situations where individuals may or may not wish to act collectively – reflects the potential diversity of agent-specific, idiosyncratic factors that influence decisions. Incomplete knowledge about others' preferences corresponds to the fact that these factors are often personal, shaped by lived experience, derived through complex reasoning, or otherwise hidden or opaque to others. Thus, whatever the merits of the critique against rational-choice analysis, the assumptions of our model can withstand this objection.

2.3.1. Basic design

For simplicity, we assume the group consists of only two agents: Agent 1 and Agent 2, though the logic can be extended to larger groups. The social interaction is structured as follows: agents have a choice of exerting effort or not in every period. One agent, referred to as the "affected party," may gain utility benefits if both parties exert effort. The other agent, called the "contributing party," does not gain utility benefits in the current period regardless of agents' actions. Making effort is costly: whoever makes effort incurs a utility cost c. Moreover, agents' roles change randomly. In one period, an agent may be an affected party and become a contributing party in another. In each period, the roles are reassigned, with each agent having a ½ probability of taking on the affected party role or the contributing party role.

Up to this point, the setup combines key features of social dilemmas and assurance games. On one hand, a lack of effort is a risk-free action. When any agent refuses to exert effort, they avoid the cost of contribution, and their outcome does not depend on the other party's choice. This risk-free option links our formal framework to assurance games. On the other hand, the cooperative option involves an incentive problem: because the effort does not benefit all agents simultaneously, one party will only contribute if they hope to benefit from this choice in

the future. This party faces a temptation to defect and avoid the cost of effort in the current period. This temptation turns our scenario into a social dilemma.

The social interaction can be interpreted in various ways. For example, the affected party may be a buyer wronged by a seller's malperformance, and the wrongdoing will be rectified by the seller if two agents pool their resources (e.g., by making a joint threat of boycott) and risk a costly confrontation with the seller. Alternatively, the affected party may seek to benefit from using a resource exclusively over a long period, but the benefits only materialize when both agents refrain from consuming or exploiting it immediately. Because the resource has some apparent immediate value, refraining from consumption means that both agents incur a private opportunity cost. In either scenario, it is assumed that a single agent is too weak, incompetent, or otherwise unable to secure the benefit alone; they need a complementary effort from the other party.

As was explained, the social interaction is not identical each time it occurs but has multiple possible variants. For instance, the seller's wrongdoing experienced by the affected party may involve delivering a product of defective quality, insufficient weight, or non-standard granularity; exploiting the resource immediately may take several forms, each inhibiting different long-term uses while having minimal or no impact on others, etc. Formally, let $\Omega = \{1, \dots, \mathrm{N}\}$ denote the set of all conceivable variants of the social interaction. It is simply assumed that the specific variant that the agents face in any given period is randomly drawn from a uniform distribution over Ω at the beginning of that period.

Variants matter because they affect the payoffs. Each of the agents can benefit from mutual effort being exerted in some variants of the social interaction but is indifferent to the others. They may be wronged by some malperformances by sellers and indifferent to other malperformances; they may value some long-term uses while disregarding others, and so on. Whether an agent benefits from mutual effort in a given variant is a characteristic specific to that agent. We say that an agent who benefits from mutual effort in some variant of the underlying social interaction *cares* about that variant. The agent-specific preference scheme that determines whether Agent i ($i = 1, 2$) cares about variant $k \in \Omega$ will be called this agent's classification scheme. Formally, a classification scheme is the set of all variants about which Agent i cares and will be denoted as Ω_i.

Put simply, a classification scheme describes Agent's i preferences. It captures this agent's subjective valuation of real-world outcomes. Each variant in the set Ω represents a real-

world outcome achievable through the collective efforts of two agents. In our previous example, these outcomes included the absence of qualitative flaws in the delivered goods, specific weights, granularity, and similar characteristics. This framework, however, can be easily generalized: in abstract terms, it simply includes the states of the world that Agent i subjectively considers desirable.

Figure 6

Stage-game payoff matrix – variant k

		Contributing party: Effort	Contributing party: No effort
Affected party: Effort		$u_k - c$ / $-c$	$-c$ / 0
Affected party: No effort		0 / $-c$	0 / 0

Figure 6 shows the payoff matrix for an exemplary variant k. As is conventional, the top line indicates payoffs of the row player, while the bottom line shows those of the column player. The variable u_k represents the affected party's valuation of mutual effort exerted in variant k of the social interaction. u_k can take the value g when Agent i cares about k, or 0 otherwise. We assume that $g > 2c$, so that mutual effort in cases in which agents care about k is socially beneficial.

Classification schemes have an important feature: they are not only subjective (i.e., agent-specific) but also private. Agents do not know each other's classification schemes with

certainty; instead, they hold beliefs. $\beta_i^k(t)$ will denote the belief of Agent i, held at the beginning of period t, that Agent $j \neq i$ cares about variant k. As previously mentioned, this assumption of incomplete information reflects the familiar and universal reality that individuals have only limited knowledge of others' preferences, values, and objectives. However, this knowledge is dynamic; when agents observe others' actions, their beliefs are updated accordingly.

Finally, we make a simple assumption that $\delta \in (0,1)$ is an inter-period discount factor common to both agents. This completes the description of the payoffs achieved in the game. The events of each period unfold as follows:

1) Some variant k of the social interaction is drawn from a uniform distribution over Ω, and one of the agents, each with probability ½, becomes the affected party, while the other becomes the contributing party;
2) The game represented in Figure 6 is played – the agents simultaneously decide whether they make effort or not;
3) Payoffs are distributed, and the period ends.

Agents' objectives can now be specified. We assume that they maximize their expected utilities at their current point in the game. This means that in any period T, an agent maximizes the expected utility over the remainder of the game, represented by:

$$\sum_{\infty}^{t=0} \delta^{T+t} \upsilon_{T+t}$$

where υ_{T+t} is the utility expected in period $T+t$.

The model is now fully specified. Before proceeding further, one remark about collective enforcement is necessary. Regardless of the specific interpretation chosen, the model is based on the necessity of collective action as its conceptual foundation. It assumes the absence of any external (i.e., third-party) enforcement agency, such as state organizations, and therefore rules out the possibility of externally manipulating agents' payoffs. Cooperative outcomes can only be achieved through the direct efforts of the parties involved in the social interaction.

The reliance on collective action is a significant assumption, but it can be justified both historically and from the abstract legal theory perspective. Historically, collective action within security networks – such as the cognatic groups, lineages, clans, or voluntary alliances – served

as the primary means of enforcing rights, freedoms, and other legal positions in pre-modern societies (see, e.g., Unger, 1976; Murray, 1983; Westbrook, 1988). A major part, if not most of the pre-modern history was marked by a

> "need for self-help and the usefulness of belonging to either a strong family or kin group or enjoying the protection of a more powerful individual, whether he was your lord and you his humble servant or tenant, or whether he was the lord who accepted your free service and provided protection in return." (Drew, 1995:1585)

Thus, there is little doubt that the extensive involvement of specialized and organized enforcement agencies in law enforcement is a relatively recent phenomenon, dating back only to the beginning of the industrial era (Friedman, 1995; Allen and Barzel, 2011; Koyama, 2014). This observation justifies the assumption of collective enforcement, as long as we are interested in studying "old," "primitive", or pre-modern institutions. Additionally, owing to the anarchic structure of the international system, collective action has historically been and still remains the key enforcement mechanism in international law (Hathaway and Shapiro, 2011; see also Chapter 1). Together, these points suggest that models assuming collective action as the sole means of enforcement are applicable to a wide range of historical contexts as well as to the international sphere.

More abstractly, it can be argued that all enforcement, even that which is supposedly centralized and organized, requires some degree of coordination among independent agents. Thus, all enforcement is collective at its core. For instance, Basu (2018) suggests that because enforcers are always numerically many, enforcement action is inherently collective action. Consequently, the difference between state-enforced rules and socially enforced rules is one of degree rather than kind. Similarly, Postema (1982) contends that the coordination provided by law operates on three levels: coordination among regular agents, between regular agents and representatives of the legal machinery, and among the representatives of this machinery, including enforcers, who must act jointly for the law to be effective.

2.4. Complete information and payoff-dominant equilibria

Before we proceed to a more nuanced analysis, a few elementary but important remarks should be made. They will further illustrate the previously discussed difference between the incentive problem, the coordination problem, and the discovery challenge.

First, it should be noticed that regardless of the state of agents' knowledge, the repeated game has one obvious equilibrium: no cooperation, always. A strategy of no cooperation simply means always opting for no effort, regardless of the other party's actions.[17] This strategy guarantees an expected per-period payoff of 0. Importantly, the no-cooperation strategy is risk-free. An agent who never makes effort avoids certain utility losses in the name of obtaining uncertain utility gains, now or in the future. It is also evident that no cooperation is a best response to itself: if one agent never contributes, the other agent has no incentive to contribute either. Therefore, a pair of no-cooperation strategies constitutes an equilibrium, irrespective of the agents' beliefs about each other's private classification schemes.

However, cooperative equilibria may also exist, and when they do, they are preferable to the no-cooperation equilibrium in terms of utility (i.e., they are Pareto-superior to the no-cooperation equilibrium). To illustrate, consider a complete information scenario where the classification schemes Ω_i for $i = 1, 2$ are public knowledge. There is no discovery challenge in this scenario: agents know each other's private preferences and thus are fully aware of how their actions affect each other's utility. This means that from the outset of the game, agents know the viable patterns of mutual effort – i.e., possible equilibria. Intuitively, these patterns consist of all pairs of strategies with the following property: at all times, a threat of pivoting to the no cooperation strategy by one agent is sufficient to force the other agent to continue executing his current strategy – a basic Folk theorem construction (see, Friedman, 1971).

However, equilibrium play remains ambiguous in such cases because multiple equilibria with this property are conceivable. For example, agents might cooperate in every variant of the social interaction they both care about, and only those. Alternatively, they may trade effort in some variants for effort in others. In order for an equilibrium to materialize, parties need to have

[17] This is a so-called Cournot strategy: a strategy in which players choose a non-cooperative strategy in each individual period, treated as a separate one-off game. Here, in a single period, it would be optimal for both agents to exert no effort.

a common notion of "right" and "wrong" behavior – this is, a common expectation of when effort should be taken. Therefore, the coordination problem persists.

Limited knowledge about agents' private classification schemes – or more broadly, about the game's structure – adds an important element to the picture: the common notion of "right" and "wrong" needs to develop over the course of the game. We will now systematically consider two stylized mechanisms for establishing such a notion: custom and deliberate design. These two represent the broader categories of unintentional and intentional institution-making. Customary rules develop when agents rely on past experiences to form expectations about others' actions and adjust their own accordingly. In contrast, designed rules are supplied upfront by an external party party, so that agents' expectations can be explicitly set in advance. The comparison will be carried out in the next two chapters of the thesis. Before that, we provide a brief non-technical summary of the argument.

2.4.1. Argument in a nutshell

The next two chapters introduce two types of equilibria distinct from the no cooperation equilibrium. In both cases, agents' long-run actions are guided by rules: ambient rules in one case and predefined, designed rules in the other. These equilibria will be defined, characterized, and compared. We will demonstrate how each type of rule overcomes the discovery challenge in a unique way. However, since some parts of the discussion may be technical, this subsection begins with an informal summary.

In essence, we propose that customary mechanisms for rule creation typical for ambient rules resemble finitely revealing equilibria known from game theory. In this type of equilibrium play, agents share private information in a limited number of steps. This information sharing may involve actions such as making a cooperative effort for the first time, punishing behaviors when the offending party does not expect punishment, and taking risks before others follow suit; it can also take the form of forgoing actions, withholding cooperation, or deliberately ignoring transgressions. These actions lead other agents to update their beliefs about the private, decision-relevant characteristics of others, which shapes their expectations about their likely future actions. In other words, when devising customary rules, agents use trial-and-error methods to identify viable ways of long-term cooperation. This process is called the "formative phase." In the formative phase, the common notion of desirable and undesirable behavior is gradually shaped by the involved parties.

After the formative phase, beliefs and expectations no longer change, and customary rules become firmly established. This results in a long-term, stable pattern of cooperation. Agents expect others to comply with historically established rules and also comply with them themselves.

The formative phase is consequential for the content of ambient rules. Generally, the formation of more complex customs requires more investment in the acquisition and interpretation of relevant information. Therefore, we claim that simpler (i.e., more general, less case-specific) ambient rules have the advantage of offsetting this requirement and therefore are more likely to develop. Because they are cheaper to communicate, observe, and interpret, simple rules have a higher chance of being widely adopted. Moreover, because they change gradually, customary rules are "slow-moving" in the sense of being able to gradually adapt to changing circumstances.

In turn, establishing rules through deliberate design – for example, through methods such as legislation or codification – is likened to constructing focal points known from the game theory (Schelling, 1980 [1960]). A focal point is a third-party coordination device that agents use to structure their interactions. In this scenario, strategies used in the long-term equilibrium are not discovered gradually by the agents but deliberately pre-configured before the social interaction begins. As long as agents expect each other to follow these unique pre-configured strategies, none has an incentive to defect.

Compared to ambient rules, using preconfigured focal points offers an important informational advantage: agents can observe whether a focal point induces an equilibrium in a number of steps that is independent of rule complexity. This means that designing rules can generally allow for greater complexity and sophistication of institutions. At the same time, designed rules, precisely because they are deliberately constructed, are more prone to manipulation by vested interest or to error. They are also "fast-moving": they may persist in a stable form for a prolonged period, only to collapse abruptly and unexpectedly.

3. Ambient rules and designed rules

Summary

This chapter uses the model to represent ambient rules and designed rules. Ambient rules will be understood as rules of conduct that emerge without intentional design and in the absence of purposeful external coordination. They are common in areas such as archaic law, customary international law, commercial usages, and social norms, where custom plays an important role as a source of rules. In contrast, designed rules are deliberately constructed and communicated in advance. We suggest that the lifecycle of ambient rules can be divided into two phases: a formative phase and a long-run phase. During the formative phase, agents actively shape the rules for future cooperation and adjust their expectations about the long-term behavior of others. In the long-run phase, expectations stabilize and patterns of cooperation become predictable. On the other hand, designed rules function as pre-game focal points – unique reference points that help agents coordinate their actions. Because they are unique, incentives to engage in strategies based on focal points are indivisible: agents cannot selectively employ these strategies without undermining the entire coordination process, and thus upsetting the equilibrium altogether.

3.1. Introduction

In this chapter, we take a closer look at two stylized types of institutions: ambient rules and designed rules. We use the model introduced earlier to formally characterize both types. The chapter begins by outlining the historical and philosophical background of distinguishing between institutions often labeled "customary," "informal," or "spontaneous" – which we jointly put under the umbrella name "ambient rules" – and those described as "deliberate," "formal," or "designed." This differentiation is shown to have a solid foundation in legal theory and legal philosophy and to be relevant to legal history and social anthropology. This historical and philosophical background is presented in Section 3.2, right after the Introduction.

The next two sections explore ambient rules and designed rules as two possible equilibria of the game introduced in the previous chapter. Each section discusses the conditions for the existence and the characteristic features of these two equilibria. This analysis sets the stage for Chapter 4, which will provide a more detailed side-by-side comparison of ambient rules and designed rules.

Before we proceed to the remainder of this chapter, one important note must be made upfront. As explained in the previous chapter, the enforcement structure of the institutions examined in this chapter (and the next) is always decentralized. This means that parties must enforce rules of behavior themselves, without relying on external agencies such as a police force or sovereign entity. It must be reiterated that while this design choice reduces the model's generality, it also makes our discussion particularly relevant to two social contexts: primitive society and international law (see Chapter 1). Not surprisingly, these two context will consistently recur throughout the text as illustrations and examples.

3.2. Philosophical and historical background

Conventional wisdom suggests an obvious mechanism for establishing social rules: rules can be deliberately made. A neutral third party or the interested parties themselves may create and announce commonly known standards of behavior. For example, how should a rancher compensate a farmer when cattle destroys crops? A lawmaker can specify a liability regime, such as holding ranchers strictly responsible for damaged crops and requiring specific

monetary payments for any resulting harm. Notably, designed rules are subject to rational control: they are intentionally constructed to govern future interactions between individuals and can be directed to serve the purposes envisioned by the rulemakers.

However, a significant strand of scholarship downplays the social significance of rule design. This perspective proposes that functional social order may (and often does) emerge without recourse to predefined, rationally construed rulesets. Instead, it emphasizes the prevalence and efficiency of customary, or spontaneously emergent rules (e.g., Hayek, 2013 [1973]; Sugden, 1989; Taylor, 1982). In this context, custom can be understood broadly to represent rules derived from past experiences in the same social dilemma. So-defined custom encompasses a variety of traditions, social norms, or historically established conventions. Scholars tend to agree that rule production mechanisms based on custom are especially common in areas like international law, social orders of pre-literate men (i.e., primitive law), and everyday social norms (see, e.g., Barkun, 1968; Guzman, 2008; Young, 2015; see also Chapter 1). Consistently with Chapter 1, we use the technical term "ambient rules" to refer to such institutions emerging from decentralized processes.

The tension between these two stylized models of rule emergence – i.e., design and custom – has been observed by legal scholars and legal philosophers. For instance, Raz (1994:371) describes an institution-oriented model of the rule of law, which "requires elaborate bureaucratic machinery with meticulously observed and policed procedures, [...] and which require for their success anonymous impartial institutions, inhabited by impartial strangers." This model assumes that publicly announced, prospective, and general rules are necessary for its implementation. Raz contrasts institution-oriented models with models of the rule of law that treat legal rules as commonly shared traditions. These can be called customary models of the rule of law.

Likewise, in't Veld (2023:265-266) distinguishes between the bureaucratic and tradition-oriented ideals of the rule of law. Crucial to the bureaucratic ideal is its purposeful future orientation; it requires that "[t]he law should be publicly laid down so that people are able to plan their lives accordingly." In turn, the tradition-oriented ideal can be conceptualized as "a set of practices which evolved over time and withstood the test of time." This model views the rule of law as an accumulation of practices that are legitimized through their longevity and practical efficacy, rather than through formal rule-making processes. In a related vein, Hadfield (2017) writes about social control in preliterate societies. She contrasts tacit rules that emerge from long-standing traditions, which are organic and informally upheld by community

consensus, with those that arise from deliberate legal orders that are supported by sophisticated legal infrastructure, including by professional lawmakers.

A similar tension is present in Hayek's theory of spontaneous order and its antitype, "made" order (Hayek, 2013 [1973]). While Hayek's concept of "order" is broad, it can be concretely applied to social rules. According to Hayek, rules may arise as products of unsupervised evolutionary processes. Through the interactions of individuals each pursuing their own interests, social rules can emerge without being purposefully constructed. This spontaneous emergence results from the gradual convergence of agents' actions and beliefs into a consistent pattern of group behavior. Conversely, rules can also be deliberate products of rule-making activities. This rule creation mechanism involves intentional design by third parties, with a specific social objective in mind. In short, spontaneous orders emerge when third parties do not intentionally design rules; "made" social orders are characterized by third-party entanglement with a specific purpose in mind (Luban, 2020).[18] Further examples of the contradistinction between broadly conceived customary and deliberate methods of establishing institutions can be found in institutional economics (Knight, 1992) and legal theory (Parisi, 1995; Hart, 1961).

Historical evidence suggests that customary and deliberately designed rules often played complementary roles in regulating social behavior. Both rule-making mechanisms were used to address the same social problems, often in immediate succession. For example, consider the prevention and repression of various forms of sexual misconduct in antiquity. On legislative grounds, it might seem that pre-imperial Rome was lenient toward sexual offenses. During the Roman Kingdom (753-509 BC) and most of the Roman Republic (509-27 BC), rape was not criminalized (Gardner, 2008). Moreover, it is uncertain whether any civil remedy designed specifically for rape was provided in republican legislation. Most likely, rape was not treated as a standalone civil wrong. Depending on the circumstances, it could be addressed under categories such as assault or abduction, for which there were established legal remedies in written law (Lintott, 1968; Gardner, 2008). In short, protection against (specifically) sexual offenses appears to have been largely absent from Roman law until the late Republic.

[18] In a way corresponding to the Hayek's argument and openly inspired by it, Leoni (1993:7) considers custom a viable rule-making method for a society: "fewer and fewer people now seem to realize that just as language and fashion are the products of the convergence of spontaneous actions and decisions on the part of a vast number of individuals, so the law too can, in theory, just as well be a product of a similar convergence in other fields." Leoni insists it is not only a viable but a desirable method: "The law-making process ought to be reformed by making it mainly, if not only, a spontaneous process, like that of trading or of speaking." (Leoni; 1993:134)

The absence of explicit legal statutes against rape suggests a reliance on a different, non-codified, societal mechanisms to manage issues related to sexual misconduct. Indeed, unwritten customary rules pertaining to unwelcome sexual behavior seem to have been robust, rigorously upheld, and often enforced through severe and violent means. Private initiative played a major role in preventing and punishing sexual misconduct. As Lintott (1968:26) claims, "the crimes of rape and adultery were (…) the subject of private revenge throughout the Republic." While such revenge was not explicitly recognized in legislation, it was considered a legitimate response by the general public and thus played a role in preventing and repressing offenses of sexual nature.[19]

When legislation finally interfered, it is possible that it primarily aimed to legitimize preexisting customary rules while occasionally delegitimizing others. This pattern can be seen in legislation addressing sexual harassment. Sexual harassment becomes recognized as a private wrong in the late Roman Republic; remedies are granted against following or soliciting women and, most interestingly, against drawing a female's guardian away. This legislative intervention appears to have been initially driven by a desire to reinforce existing social norms regarding female chastity, in particular the norm that women should be accompanied by an adult male protector (Gardner, 2008). In turn, when legislation against various forms of vice, including rape, adultery, and other sexual misconduct is enacted in the early Empire, which is several generations later, previously strong social norms of sexual conduct seem to have largely vanished or become significantly relaxed, as evidenced by the preoccupation of the legislation with concubinages (Nguyen, 2006; Gardner, 2008).

The substitution of designed rules for preexisting customs was often done explicitly through the acts of writing down (black-lettering) of previously prevalent customary rules. For example, in the era of Greek colonization, customary laws became "codified and written down with the addition of new laws framed by the lawgiver to correspond to the exigencies of his time." (G. Smith, 1922:187) Likewise, some scholars consider the Law of XII Tables, traditionally believed to be the first piece of Roman legislation, to have been "nothing more than a codification of such [customary] law, with perhaps a few legislative innovations." (Schiller, 1938:275; cf. Westbrook, 1988; Crawford, 2012) While contemporary researchers

[19] The extensive reliance on self-help seems to have been one of the general features of the social order of the early Roman Republic. Norms regulating the legitimate use of self-help have been gradually incorporated into written law. According to Lintott (1968:34), "Roman legal procedure was originally modeled on ritualized self-help, and for its successful functioning it relied on self-help. Private action was its foundation, and so it cannot be surprising that it allowed the individual so much scope to right his own wrongs."

challenge the established view that the Law of XII Tables was the first ever legislation in Ancient Rome, they still maintain that the process of writing down old customary laws occurred and was a transformative experience for the early Roman society, but simply place it earlier in the history of the city. For example, Perelló notices that

> "According to the *communis opinio*, the mid-fifth century BCE represented a turning point where legislation would either be invented or imported into Roman culture, with the Twelve Tables being hailed as a model. Thus, all reports of earlier legislation were written off as later annalistic inventions which retrospectively projected the situation of subsequent centuries onto the previous, formative era, which should be considered as part of the customary order, rather than the legal one." (Perelló, 2020:57)

In line with our previous remarks about the lack of control over customary norms, he simultaneously observes that

> "Whatever virtues one might ascribe to custom, it lacks the specific ability to self-reform. As a source of law, custom always preserves the laws of the past, and even when it creates new institutions, it usually connects them with older material to create a smooth transition from older laws to new ones. Custom is the result of blind historical developments, and therefore it lacks a rational texture. Where custom has no planning and no conscious aim, legislation is the result of deliberation, contains objectives, and implies design. Custom does not usually reform custom for the simple reason that to amend means to plan, and that can only be done through legislation." (Perelló, 2020:59)

In other words, the transition from the "customary order" to legislation represented a qualitative change toward a social order that could, within limits, be consciously planned and modified.

The same was true of the transition away from orally transmitted customs in Medieval Europe. Since at least the 12th century, written inventories of customs specific to a given place

(e.g., town or province) have begun to surface. They originated as mere attempts to provide a tangible list of rules binding within a given community and initially were not treated as authoritative sources of law. Writing down customary rules originally happened in an unofficial, private manner. However, such private inventories are soon replaced by *official* statements of customary rules authorized by appropriate authorities. Eventually, they become the only valid sources of rules, replacing living customs altogether (Glenn, 1997; Rossi and Spagano, 2018). According to Glenn (1997), these developments empowered a small group of learned lawyers with tools to deliberately influence social rules through the selection, written formulation, and amendment of customs. Where customs were uncertain, difficult to put into writing, or seemingly without clear purpose, those who recorded them either "corrected" or omitted them altogether. While the stated goal was to formalize conventionally followed practices into the language of law, the process of writing them down gradually turned customary law into something customary in name only, transforming it into the domain of a certified profession of lawyers who were driven by conscious purpose.

Finally, international law offers another context where the contrast between the spontaneous emergence and the deliberate design of rules is relevant. According to received wisdom, there are three sources of international law: custom, international treaties, and general principles of law. The interaction between the first two sources is interesting: scholars have noted that custom gradually gives way to written treaties (or international conventions) as a primary source of international law. This shift is another example of substituting customary institutions with broadly conceived designed rules.

Indeed, the number of multilateral treaties in force has been increasing significantly since 1945, especially in areas such as security (e.g., UN Charter), trade (e.g., WTO agreements), environment (e.g., Paris Agreement), and human rights (e.g., International Covenant on Civil and Political Rights) (see, e.g., Inoguchi and Le, 2020). This tendency was noted already in the 1980s:

> "The United Nations Treaty Series comprises well over a thousand volumes, with some 17,000 treaties in force. The League of Nations Treaty Series, before it was concluded upon the advent of the United Nations Treaty Series, consisted of 205 volumes with close to 5,000 treaties. Each treaty usually incorporates numerous stipulations, sometimes in the hundreds." (Dinstein, 1986:7)

It is sometimes argued that treaties offer precise, codified rules that can be quickly negotiated and adapted to modern challenges, which is unlike the slower evolution of customary international law. Moreover, especially in today's globalized world, the wide range of state practices makes it difficult to ascertain uniform practices that are the basis of customary international law. These reasons might have incentivized states to increasingly rely on explicit multilateral treaties instead of customary international law (Lim and Elias, 1997). Regardless of the underlying reasons, the transition from custom to treaties in international law seems to be an undeniable development.

3.3. Ambient rules

Economists have long recognized that patterns of cooperative behavior may spontaneously develop in repeated interactions (e.g., Sugden, 1989; Coleman, 1990; Parisi 1995; Young, 2001 [1998]). Such patterns have been typically modeled with the use of game theory: either as evolutionarily stable equilibria in the evolutionary setting (e.g., Axelrod and Hamilton, 1981; Parisi, 1995; Sugden, 2005 [1986]; Morsky and Akçay, 2019) or subgame-perfect equilibria in the perfectly rational setting (e.g., Mahoney and Sanchirico, 2003; Bicchieri and Sontuoso, 2020).

The conventional law and economics approach suggests that the demand for cooperative rules arises in social dilemmas: situations where individuals responding to short-term incentives produce socially suboptimal collective outcomes. In such cases, there is an opportunity to achieve a cooperative surplus if individuals refrain from acting solely in their immediate self-interest. However, efforts to achieve this surplus can take different forms. One of them will be called "ambient rules," consistently with the typology from Chapter 1. Ambient rules refer to cooperative solutions that arise organically among participants within the existing structure of the game, without external intervention (see, e.g., Aoki, 2001; Parisi, 2000). This characteristic conventionally distinguishes them from legal rules, which are said to involve deliberate modifications to the game's structure. Legal rules typically reallocate rights in ways that intentionally reshape incentives and thus alter the payoff matrix (e.g., Picker, 1994).

The last point carries special importance. The textbook approach rules out the involvement of a third party in the creation or change of ambient rules. Consequently, the

discovery challenge must be resolved without external involvement; rather, it must be addressed through a confluence of expectations arising from independent decision-making. Take the example of rules of trading in a marketplace. Neither sellers nor buyers can learn and understand the rules of trade by reading a code of using a commonly known textbook. A shared understanding of permissible and impermissible behavior must organically develop within the group of interacting traders.

In line with this idea, we suggest a perfect Bayesian equilibrium wherein agents gradually develop a shared understanding of cooperative behavior and consistently abide by this understanding in the long run. Using the model presented in the previous chapter, the next subchapter represents ambient rules as an equilibrium of the repeated game.

The equilibrium concept can be summarized as follows. Agents develop a shared notion of cooperation by initiating cooperative behavior when their expectation of reciprocity is sufficiently strong. Moreover, they may contribute to building this notion even without an initial expectation of reciprocity. For instance, agents might exert costly effort to signal their intention to cooperate in the long term, hoping the other party will follow through not immediately, but at some point in the future. A combination of reciprocity-based cooperation and signaling allows agents to jointly construct a common notion of rightful and wrongful behavior. This common notion develops so that both parties have long-term incentives to abide to it. However, the process of forming this shared notion, or the formative phase of ambient rules, has its limitations. Efforts to cooperate are inherently risky when reciprocity is likely but not guaranteed; on top of that, signaling entails immediate, direct costs to the signaling agent.

3.3.1. Reciprocity

Theoretical studies of norm emergence emphasize the expectation of long-term reciprocity as the key factor for creating and sustaining cooperation when no third-party authority is in place (see, Axelrod and Hamilton, 1981; Boyd and Richerson, 1988; Fon and Parisi, 2003; Elster, 2011). Through reciprocity-based cooperation, agents trade like-for-like activities in such a way that their exchange is stochastically symmetrical in the long run: they are equally likely to be receivers of benefits or contributors to the creation of benefits to others (see, Fon and Parisi, 2003). For example, automatic reciprocity is considered one of the primary meta-rules in international law (Guzman, 2008).

We begin by evaluating how strong the expectation of reciprocity must be to incentivize an agent to exert costly effort. Assume that the agents are facing variant k of the social interaction in the current period. Consider the following expectation held by Agent i ($i = 1, 2$), who is the contributing party: with probability p_k, the affected party will contribute every time variant k occurs, provided that the contributing party continues doing the same; with probability $1 - p_k$, the affected party will not contribute. A rational Agent i will exert effort in the current period if they care about variant k and the following condition is satisfied:

$$-(1-\delta)c + \delta p_k \frac{g-2c}{2N} > 0 \tag{1}$$

The left-hand side represents the contributing party's (i.e., Agent's i) expected utility value of exerting cooperative effort on behalf of the affected party in the current period. It consists of two parts: the cost of contribution that will be incurred in the current period and the continuation payoff, which will be realized in the future with probability p_k. The continuation payoff reflects the expected value accumulated over all future periods in which the agents cooperate in variant k of the social interaction, which constitutes a fraction $\frac{1}{N}$ of all periods. In half of those periods, Agent i is the affected party and enjoys a utility gain g.

Conversely, the right-hand side of Inequality (1) is nil; it represents a scenario where Agent i refuses to exert effort and thus the agents never cooperate in variant k in the future.

Condition (1) gives the following specification of the minimum p_k, i.e., the minimum belief by Agent i that the other party will infinitely reciprocate effort when the agents face variant k:

$$p_k > \frac{1-\delta}{\delta}\frac{2Nc}{g-2c} \equiv p_k^* \tag{2}$$

3.3.2. Signaling and extended cooperation

However, Condition (2) is relevant only if it simultaneously holds for *both* agents. If it holds only for Agent i but not for Agent $j \neq i$, then the expectation that Agent j would exert

effort with probability p_k cannot be rational: Agent j himself does not believe that reciprocal cooperation would materialize, and thus has insufficient incentives to exert effort.

Does it mean that in asymmetrical cases – i.e., when one party believes that there is a potential for cooperation but the other party does not share this belief – cooperation cannot materialize?

It can, but it needs to be initiated differently. This can be done through signaling: agents need to update each other's beliefs by showing their intentions to cooperate. Signaling is a method of starting mutually beneficial endeavors well known from the law and economics literature. For example, the literature on norm entrepreneurship exploits this idea. According to the norm entrepreneurship theory, an agent may take the risk of starting a new group practice in hope that others would follow suit. In case they do, future benefits accruing to the originator of the new practice would outweigh the cost of risk-taking. Such first movers, or norm entrepreneurs, often exhibit superior situational awareness, technological knowledge, or another kind of comparative advantage that makes them particularly suitable for ushering in new modes of conduct (Ellickson, 2001). Likewise, Leeson (2008) shows that heterogeneous agents can use costly signals to manifest their intention to cooperate in the long run and thus reduce the "social distance" that impedes beneficial transactions.

We consider the simplest case of signaling. Assume that Agent i does not believe that Agent $j \neq i$ will exert effort in variant k of the social interaction in the current period. However, he also believes that with probability p_k, once Agent j observes Agent's i effort, he will cooperate in all subsequent periods in which variant k occurs.

In other words, Agent i may signal the intention to cooperate in variant k of the social interaction in the future by unilaterally exerting effort in the current period. He believes that, with probability p_k, this signal will be sufficient to initiate reciprocity-based cooperation in all future instances of variant k. Signaling involves incurring the cost of effort twice without certainty that cooperation had materialized: first, in the current period as a costly signal, and then in the next period in which agents face variant k again. The outcome of this next period will reveal whether signaling has proven successful: with probability p_k, both agents will exert effort from this next period onwards. Conversely, with probability $1 - p_k$, Agent j does not respond to the signal and exerts no effort in the next period instance of variant k. Therefore, the following criterion must be met for a rational Agent i to engage in signaling:

$$-(1-\delta)c + \delta p_k \frac{g-2c}{2N} - (1-\delta)(1-p_k)\frac{\delta c}{\delta + (1-\delta)N} > 0 \quad (3)$$

The left-hand side of Inequality (3) represents a signaling scenario. In the current period, Agent i incurs the cost of effort c. This opens two possible futures. With probability p_k, Agent j will cooperate in variant k in all future periods, resulting in a discounted continuation payoff $\delta\frac{g-2c}{2N}$ for Agent i. However, with probability $1-p_k$, Agent j refuses to make effort the next time variant k occurs. In this case, Agent i incurs the cost c one more time and, upon realizing that Agent j does not cooperate, ceases to contribute. The next occurrence of variant k is discounted by the factor $\frac{\delta}{\delta+(1-\delta)N}$.[20] The right-hand side of Inequality (3) represents the scenario where Agent i decides not to signal.

Inequality (3) yields the condition specifying the minimum probability p_k that makes signaling rational:

$$p_k > p_k^{**}(c, \delta, g, N) \equiv p_k^{**} \quad (4)$$

Given the values of c, δ, g, and N, p_k^{**} is the probability of cooperation for which the left-hand side and the right-hand side in Expression (3) are equal. Values of p_k greater than p_k^{**} justify signaling in the eyes of rational agents.

We do not write Condition (4) explicitly like we did in the case of Condition (2), even though it can be done with some additional difficulty. It is sufficient to notice that Condition (4) is stricter than Condition (2) in the relevant range of $p_k \in [0, 1]$. This is because Inequality (3) "includes" Inequality (1) but adds and additional summand to its left-hand side: $-(1-\delta)(1-p_k)\frac{\delta c}{\delta+(1-\delta)N}$. As long as $p_k \in [0, 1]$, this summand is non-positive. In fact, it is negative for all values of p_k other than 1. Therefore, a higher value of p_k is needed to satisfy Condition (4) than Condition (2) – in other words, $p_k^{**} > p_k^{*}$.

[20] This is a sum of an infinite series of probabilities that the next occurrence of variant k is exactly in period T, times the discount factor of this period. This series covers all future periods from the next one until infinity.

The interpretation is intuitive: initial signaling is a costly way to initiate cooperation. It requires an additional costly step before reciprocity-based cooperation may be expected. However, there is an important difference between reciprocity-based cooperation and signaling. Unlike reciprocity-based cooperation, signaling does not require symmetrical expectations. Condition (4) only needs to be satisfied by the agent who assumes the role of risk-bearer, regardless of the other agent's initial belief.

This conclusion squares with the conventional account of norm entrepreneurship that can be encountered in the literature. This conventional account highlights that individuals take risks by initiating changes to prevalent social practices with limited initial expectation of success. For this reason, successful norm entrepreneurs tend to exhibit comparative advantages in taking such risks, e.g., they have greater stakes in social change, have the capacity to accommodate the associated costs, or possess information unavailable to others. We represent this comparative advantage as an advantage in knowledge: an unusually strong belief p_k that others will eventually adopt new, cooperative modes of behavior if they observe first originators.

The logic of signaling can be extended further. While our framework outlines a specific type of signaling procedure, the concept can be generalized to encompass all costly signaling and cueing methods used to identify long-term cooperation opportunities. For instance, agents may trade efforts in variants of the social interaction that only one agent cares about for efforts in variants that the other agent cares about. Naturally, the hypothetical participation thresholds $p_k^{(a)}$ for singnaling that would convey the necessary information would be even higher than p_k^{**}. In other words, the idea is straightforward: signaling procedures become more costly with increased complexity of the underlying social interaction, as measured by the total number of possible variants n.

3.3.3. Equilibrium

An equilibrium in which agents organically create ambient rules can be imagined as a superposition of reciprocity-based cooperation and cooperation initiated through signaling. How the rule-creation process unfolds depends on the structure of initial beliefs (i.e., $\beta_i^k(1)$ for $i = 1, 2$ and $k = 1, \dots, N$). In other words, the extent of cooperation achieved through the formation of customary rules depends crucially on the initial *familiarity* with each others' private preferences and objectives. This degree of this familiarity determines whether, when

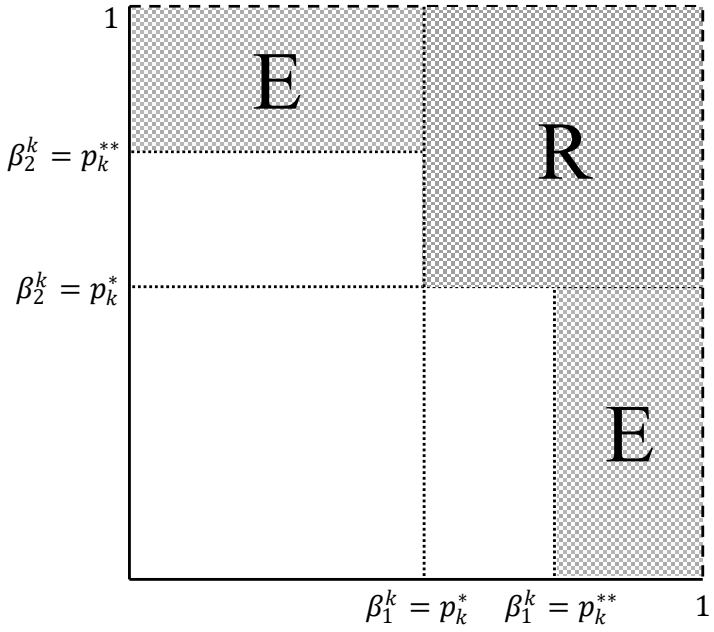
variant k occurs for the first time, agents will engage in reciprocity-based cooperation, signaling, or disregard the possibility to exert effort altogether. Those initial responses shape the scope of ambient rules and define the content of the common notion of desirable and undesirable behavior that guides collective action.

Figure 7 represents a belief space about variant k of the social interaction. The figure can be used to illustrate the logic of how initial beliefs affect the behavior in the formative phase. When $\beta_1^k(1) > p_k^*$ and $\beta_2^k(1) > p_k^*$, reciprocity-based cooperation may begin. The agents are sufficiently convinced that they can both benefit from exerting mutual effort in variant k. Once cooperation begins, agents subsequently update their beliefs $\beta_i^k(t)$ to 1, allowing cooperation to continue. All configurations of initial beliefs that support reciprocity-based cooperation are located in the dark grey field R in Figure 7.

In turn, in asymmetrical cases when $\beta_j^k(1) \leq p_k^*$ but $\beta_i^k(1) > p_k^{**}$ for $i \neq j$, Agent i $(i = 1, 2)$ faces an incentive to signal. For those configurations of initial beliefs, one agent unilaterally exerts effort when variant k occurs and thus signals an intention to cooperate in the future. This signal causes the other agent to update their belief $\beta_i^k(t)$ to 1. After the belief is updated, Condition (2) is satisfied, allowing reciprocity-based cooperation from the next period onward. Configurations of initial beliefs in which signaling can be supported are located in the light grey fields E. Needless to say, Agent i with the higher value of β_i^k is the one to take the initiative to signal.

Figure 7

Combinations of beliefs necessary for the formation of ambient rules: R – reciprocity-based cooperation; E – signaling



Signaling may take more complex forms beyond the one characterized above. If any parties finds additional ways to signal further cooperation opportunities, these ways are also used if similar conditions in the form $\beta_i^k(1) > p_i^{(a)}$ are met. In general, reciprocity has the lowest threshold of initial familiarity with other parties' private preferences and objectives; simple signaling that enables reciprocity in the next period has a higher threshold; and more complex signaling would have even higher thresholds. However, when the beliefs are insufficient for agents to rationally engage in reciprocity-based cooperation, ambient rules do not develop at all.

Importantly, agents also update beliefs when someone who is expected to contribute in any given period *fails* to do so. For example, when $\beta_1^k(1) > p_1^*$ and $\beta_2^k(1) > p_2^*$, the conditions for reciprocity-based cooperation are satisfied. Both agents are expected to exert effort when they face variant k of the social interaction. A failure to contribute when expected conveys private information about the failing agent – i.e., that this agent, in fact, does not care about variant k.

The equilibrium can be now characterized. For the sake of simplicity, we restrict the equilibrium to include only reciprocity and basic signaling. The equilibrium strategies and beliefs can be described as follows:

Agent's $i\ (=1,2)$ strategy:

1) If variant k occurred for the first time, $k \in \Omega_i$, and Agent i is expected to exert effort, Agent i makes effort; otherwise, they ignore variant k;
2) If variant k occurred at least once in the past, $k \in \Omega_i$, Agent i is expected to exert effort, and all agents who were expected to make effort in the past when k was occurring did so, Agent i exerts effort; otherwise, they ignore variant k.

Agent's $i\ (=1,2)$ beliefs $\beta_i^k(t)$ for $\omega_k \in \Omega$:

1) $\beta_i^k(t) = 1$ if Agent $j \neq i$ always succeeded in exerting effort in variant k when expected to do so until period $t-1$;
2) $\beta_i^k(t) = 0$ if Agent $j \neq i$ failed to exert effort in variant k at least once when expected to do so until period $t-1$;
3) $\beta_i^k(t) = \beta_i^k(1)$ otherwise.

Although the equilibrium may seem technical, it can be interpreted in broader and relatively simple terms. The lifecycle of ambient rules is divided into two phases. Initially, before each possible variant in Ω occurs more than twice, a trial-and-error process unfolds, which can be referred to as the formative phase of ambient rules. During this phase, agents engage in trial-and-error procedures to establish mutually beneficial long-run patterns of group behavior through reciprocity or signaling. They search for cooperation possibilities by revealing private information about their preferences to each other.

The second phase of the lifecycle of ambient rules can be termed the long-run equilibrium. In this phase, no additional private information is revealed, and beliefs remain no longer change. Agents systematically cooperate in those variants of the social interaction in which successful cooperation has been established in the past; the common notion of "right" and "wrong", or a common, public notion of acceptable and unacceptable behavior, is thus fully developed and stabilized. The remaining variants receive no collective responses and are infinitely ignored. Importantly, the equilibrium also involves punishment of opportunistic defectors. Agents who were expected to exert effort but failed to do so are penalized: in a "grim trigger" manner, those who did not contribute in variant k would be infinitely denied assistance

when they become the affected party in the same variant. This incentive prevents contributing parties from shrinking.

The equilibrium idea presented above can be generalized. Technically, it constitutes a special case of a well-known equilibrium design for infinite games with incomplete information: a finitely revealing equilibrium. In this type of equilibrium, players reveal private information to others only within a finite number of periods. After a finite number of periods, the game continues stably without beliefs being updated any further (see, e.g., Koren, 1992; Pęski, 2014). The formative phase in the current framework is analogous to the first part of the above-mentioned equilibrium. As suggested, it can be abstractly envisioned as a phase in which agents attempt to discover mutually advantageous cooperation possibilities that would be exploited in the long run. To borrow a phrase from the customary international law literature, the formation of custom is a "struggle (…) reflected in exchanges of signals, cues, bids, and responses" (Bederman, 2010:165), until the rule stabilizes in its final form.

In other words, the discovery is made through trial-and-error procedures where agents engage in risky actions and use costly signals, thus communicating future cooperation possibilities. For example, imagine that agents are buyers who may punish malperforming sellers. After a seller commits a serious wrongdoing against one of them – such as failing to deliver goods or delivering them far later than agreed – the buyers may attempt to boycott for the first time. Once they observe that the boycott is widespread, their expectations stabilize and boycotts become an expected norm. Sometimes, when a seller's wrongdoing is less apparent, a single buyer may initiate an unilateral boycott, hoping that the other buyer will join in next time. Eventually, these actions result in the creation of a common set of expectations about when buyers should boycott sellers. In the long, buyers consistently act on the basis of these expectations, which eventually become self-fulfilling.

3.4. Designed rules

We turn to designed rules. As was mentioned previously, design is a method of establishing rules distinct from custom, primarily due to its deliberate nature and the involvement of a third party in their creation. Postema (1982:178), who considers all social rules that assist in coordination "conventions", points that "conventions may be codified (i.e.,

given a canonical formulation), and codification may actually make coordination more efficient, but explicit formulation is not necessary." We will now analyze rules that are, indeed, given explicit formulation.

According to the focal point theory of law, legal rules function as focal points that coordinate agents' strategic interactions. They are a coordination mechanism such that "the solution to the coordination problem is sought not by the parties (…) but by an uninvolved third party" (Postema, 1982:184) before the game unfolds. This third party's role is to preordain "a solution by going through the process of reasoning which the parties should have used in the situation" (Postema, 1982:183-4). The parties themselves may consult this solution provided by an external party when facing a strategic interaction with multiple equilibria.

To use the previous example, buyers may consult a legal statue that specifies sellers' performances that are considered wrongful. This statute helps buyers determine when a collective boycott should be enacted. Or, in the problem of liability rules for animal trespass, a third party may announce that liability rests with the ranchers when a farmer cultivates expensive crops, but ranchers are relieved of liability when a farmer cultivates basic crops. This third party may also specify how damage resulting from trespass should be compensated (e.g., through monetary payments, payments in kind, or another method). McAdams elegantly summarizes this idea in the following way:

> "Of greatest significance to law, Schelling proposed that third-parties can create focal points in the more common mixed motive games that involve both coordination and conflict. (…) Several theorists have noted that law can work in this manner, as legal rules are a form of third-party expression (of legislators, judges, or executive branch officials) making focal the form of behavior the law demands." (McAdams, 2009:29)

In other words, designed rules provide a unique, third-party public classification scheme deliberately supplied by the rulemaker and recognized as salient by social actors.[21]

[21] Ultimately, the coordination account of law considers legal systems elaborate conventions. Those conventions assign a special status to classification schemes originating from sources socially regarded as "legitimate" while disregarding other classification schemes as contingent and thus irrelevant.

Legislation is a prime example of coordination through focal points. It provides a unified ex-ante characterization of the decision-making logic that should govern a specific class of social interactions. For example, a legislated statute may state liability rules for various types of animal trespass. With these statements in mind, ranchers and farmers can develop shared expectations regarding the appropriate actions to take in future instances of animal-inflicted crop damage. In the context of international relations, designed rules may be associated with international treaties, though not without qualification. What domestic legislation and international treaties have in common is that they are both made known in advance, which provides immediate information about future rules.

However, domestic legislation and international treaties differ in the degree to which they are negotiable. Domestic legislation is typically a creation of an external rulemaker and the rulemaker's relatively narrow circle of experts, consultants, and other influences. In contrast, parties to international treaties often negotiate their contents beforehand. For this reason, the sets of rulemakers and rule-followers overlap.[22] While this feature influences, whose interests can be incorporated into designed rules more easily (see Chapter 4), it does not change their fundamental characteristic: they are publicly communicated in advance and can serve as unique focal points.

Our analytical framework represents designed rules as a public classification scheme that is preannounced before the game unfolds. The act of preannouncement is "public" in the sense of being common knowledge: the content of the classification scheme is known to all agents, all agents know that others know it, and so on. Apart from its public nature, the preannounced classification scheme is no different from the private schemes that characterize the individual preferences of particular agents. Formally, it can be represented as a subset Ω_L of all variants of the social interaction. The set Ω_L includes the variants where, according the public classification scheme, agents are required to exert effort. Importantly, we ignore the possibility of a negotiation phase by assuming that no information about the agents' private preferences has been revealed, even if negotiations there were negotiations. As a result, the agents' initial beliefs about each other remain unchanged.

The publicness of the set Ω_L is particularly important. We assumed that agents are not familiar with each other's private classification schemes. However, because Ω_L is public knowledge, the agents are familiar with a single, unique classification scheme and are also

[22] Of course, counteries may also join an already drafted and operational international agreements, in which case they play no role in their design.

aware that others are familiar with it. Consequently, the unique classification scheme can potentially serve as a focal point around which agents align their expectations about how the game should be played (Schelling, 1980 [1960]).

3.4.1. Incentives to follow designed rules

We examine the conditions under which focal points provided to agents before the game starts can guide their behavior under incomplete information. These conditions can be characterized intuitively. Third-party coordination requires that the agents believe that the public classification scheme is sufficiently aligned with their private preferences and with the private preferences of other agents. This means that it demands effort in sufficiently many variants about which agents care and in sufficiently few about which agents do not care.

Consider the situation of a farmer who prefers strict liability rules for cattle trespass. Whenever cattle enter a farmland, a cattleowner should be liable to compensate for the damage inflicted. However, the applicable statute consists almost exclusively of negligence rules, and the stipulated negligence threshold is set high. In other words, the rule designer determines that a cattle owner needs to take modest precautions against cattle trespassing (e.g., by erecting a frugal and fragile fence) in order to be relieved from liability. In many cases where cattle damages the crops while fences are in place, no compensation for the farmer is called for. In this example, the classification scheme embodied in the legislation is poorly aligned with farmer's private preferences. However, even when it was well aligned, the farmer would still need to know, whether the classification scheme is aligned with the preferences of other community members who enforce it.

The alignment between a public classification scheme and agents' private preferences can be measured with two magnitudes: the scope of the public classification scheme and its convergence with agents' private preferences. The scope is the fraction of all possible variants in which the public classification scheme demands effort. We use the notation s_L to represent this share. Convergence, on the other hand, indicates how often the scheme demands effort in variants that agents care about. This can also be captured by a single parameter, v_i, $i = 1, 2$ representing the fraction of variants in Ω_L that Agent i cares about (and for which effort is requested by the public classification scheme because they are in Ω_L). Thus, the share $s_L v_i$ represents the portion of *all* variants of the social interaction that Agent i cares about and that are considered wrongful by the public classification scheme. Consequently, there is a fraction

of $s_L(1 - v_i)$ of all variants that the public classification scheme considers wrongful but Agent i considers neutral.

It is now possible to characterize the incentives to use focal points, i.e., to contribute whenever the public scheme demands. We call this behavior Ω_L-directed cooperation. The incentives to participate in Ω_L-directed cooperation will be characterized in several steps. First, consider the perspective of Agent i who assumes that Agent $j \neq i$ would always partake in Ω_L-directed cooperation, provided that Agent i never fails to do so. Imagine the following scenario: variant k occurs, Agent j is the affected party, and the public classification scheme Ω_L demands collective effort in variant k of the social interaction. In this case, a rational Agent i has an incentive to participate in Ω_L-directed cooperation in the current period if:

$$-(1 - \delta)c + \delta s_L \left(\frac{v_i g}{2} - c\right) > 0 \tag{5}$$

Expression (5) has a simple interpretation. The right-hand side represents a situation where Agent i does not exert effort in the current period. Consequently, Agent j ceases to participate in Ω_L-directed cooperation, and both agents turn to the no cooperation strategy for the rest of the game, which is their optimal choice.

The left-hand side of Expression (5) represents the expected utility from participating in Ω_L-directed cooperation. This utility can be divided into two parts: the cost of participating in collective action during the current period, and the continuation payoff, which is the expected utility achieved throughout all future periods. In the current period, Agent i is required to bear the cost of effort c. Subsequently, Ω_L-directed cooperation continues infinitely, with Agent i expecting to receive a per-period utility of $s_L(\frac{v_i g}{2} - c)$. This is because Agent i knows that the public classification scheme identifies the share s_L of all variants as requiring collective action. Every time one of these variants occurs, agents collectively contribute, which entails a cost c to Agent i. Moreover, Agent i cares about a fraction v_i of these variants, and since Agent i assumes the role of the affected party on average in every second period, this yields an average utility gain $\frac{v_i g}{2}$.

Note that $s_L(\frac{v_i g}{2} - c)$ can be negative. In such a case, Condition (5) will never be satisfied, and thus Agent i will never participate in Ω_L-directed cooperation. If $s_L(\frac{v_i g}{2} - c)$ is

positive, Condition (5) translates into the following condition specifying the minimum convergence between the public classification scheme and Agent's i private preferences:

$$v_i > \frac{2c}{g}\frac{1-\delta(1-s_L)}{\delta s_L} \equiv v_{min} \tag{6}$$

Expression (6) specifies a participation threshold. The value v_{min} is the minimum convergence between the public classification scheme and Agent's i private preferences that justifies Agent's i participation in Ω_L-directed cooperation, provided that the other agent participates as well. The participation threshold has intuitive properties: agents with higher δ, meaning higher value attached to future utility relative to present utility, have lower partipation threshold. Likewise, higher benefit g as well as lower cost of effort c also decrease v_{min}.

It is important to reiterate that the participation threshold described in Expression (6) is valid only under the strong reciprocity constraint. This threshold specifies Agent's i incentives to participate in Ω_L-directed cooperation, assuming that this agent knows that the other party *already* faces similar incentives. However, because information about private preferences is incomplete, it is unrealistic to assume such knowledge in all circumstances. As a result, neither agent can directly observe whether the public classification scheme Ω_L aligns sufficiently with *the other agent's* private preferences.

While unable to directly observe each other's private preferences, both agents entertain beliefs about them. We have previously characterized those beliefs by specifying that $\beta_i^k(t)$, $k = 1, \dots, N$ denotes Agent's i belief that Agent $j \neq i$ cares about variant k of the social interaction. Agents can also use $\beta_i^k(t)$ to infer the probability of convergence between the public classification logic and private preferences of the other agent. We write $\xi_i(\kappa)$ to denote the function (technically: a cumulative distribution function) that characterizes Agent's i initial belief that Agent's j convergence level amounts to at least κ. Put differently, $\xi_i(\kappa)$ is a probability, as assessed by Agent i before the game starts, that the convergence v_j between Agent's j private classification logic and the public classification logic is κ or less.

Initial beliefs about each other's convergence levels are critical for the possibility of equilibrium with Ω_L-directed cooperation in which both agents participate. They overcome the difficulty associated with the lack of direct insight into the private classification scheme of the

other agent. Instead of assuming that the other party will participate in Ω_L-directed cooperation, agents attach a subjective probability to such a scenario. From the perspective of Agent i, this probability is given by:

$$\pi_i = 1 - \xi_i(v_{min}) \tag{7}$$

Expression (7) represents Agent's i initial belief that the other agent's private preferences are sufficiently aligned with the public classification logic to engage in Ω_L-directed cooperation. In other words, it expresses, from Agent's i perspective, the probability that Agent j has sufficient incentives to be coordinated by the public classification scheme Ω_L in all future periods, provided that Agent i continues doing the same. Expression (7) is important because it can be used to make Condition (6) operational in a world where agents' knowledge of others' private preferences is incomplete. It allows expressing a participation threshold adjusted for imperfections in agents' mutual knowledge.

Note that π_i is an initial belief: it is constructed before the game starts. It depends only on the initial beliefs about the other agent's private preferences and the content of Ω_L, which is public knowledge. In fact, agents update their belief about the alignment between the public classification scheme and private preferences of the other agent only once in the course of the game. This alignment can be fully confirmed if Ω_L-directed cooperation by the other agent is observed when this agent is the contributing party. In this case, π_i becomes 1; if such cooperation is not observed, the alignment is disproven and π_i becomes nil.[23]

We can now incorporate π_i into the model by slightly correcting Condition (5). Assume that the situation is identical to the one characterized when Condition (5) was derived: variant k occurred in the current period, Agent $j \neq i$ is the affected party, and the public classification scheme calls for exerting effort in variant k. In this scenario, it is rational for Agent i to follow the direction of the public classification scheme Ω_L in the current period, if the following condition is met:

[23] When an agent whose preferences are not aligned with the public classification scheme is the affected party, this agent may exert effort if he expects effort from the other agent as well. In such a case, exerting effort is beneficial to this agent even when he does not intend to engage in Ω_L-directed cooperation in the long term. This agent "pretends" to play by the rules for the time being. However, as soon as he is required to contribute without any immediate gain, effort will stop.

$$-(1-\delta)c + \pi_i \delta s_L \left(\frac{v_i g}{2} - c\right) > 0 \tag{8}$$

Condition (8) is stricter than Condition (5) yet it resembles the latter in all but one respect. Agent i still needs to bear the cost of effort in the current period. However, they now believe that the payoff from the continuation game, i.e., the payoff from Ω_L-directed cooperation in all future periods, is expected to be achieved with probability π_i. With probability $1 - \pi_i$, Agent's j private classification logic is not sufficiently aligned with the public classification scheme, and thus Agent j will be unwilling to follow the directions given by the public classification logic now and in the future.

Condition (8) can be rewritten to obtain the minimum initial belief π_i:

$$\pi_i > \frac{1-\delta}{\delta} \frac{2c}{s_L(v_i g - 2c)} \equiv \pi_i^* \tag{9}$$

The value π_i^* specifies the minimum belief of Agent i about Agent's j preference alignment with the public classification scheme Ω_L that rationally justifies Ω_L-directed cooperation.

3.4.2. Equilibrium

With Condition (9), we can characterize the equilibrium in which both agents participate in Ω_L-directed cooperation, i.e., follow the set of rules that has been predefined before the game began.

Agent's i $(= 1, 2)$ strategy:

1) Adapt the actions demanded by the public classification scheme Ω_L (i.e., engage in Ω_L-directed cooperation) if $\pi_i(t) > \pi_i^*$ and if both agents never failed to adapt any of those actions in the past;
2) Always choose no effort if either $\pi_i(t) > \pi_i^*$ is not satisfied or if any agent failed to adapt an action demanded by the public classification scheme Ω_L in the past.

Agent's i $(= 1, 2)$ beliefs $\pi_i(t)$:

1) $\pi_i(t) = \pi_i(1) = 1 - \xi_i(v_{min})$ as long as Agent $j \neq i$ was never a contributing party when the classification scheme Ω_L demanded effort and never failed to engage in Ω_L-directed cooperation in the past;
2) $\pi_i(t) = 1$ if Agent j always engaged in Ω_L-directed cooperation in the past;
3) $\pi_i(t) = 0$ otherwise.

The equilibrium idea is simple. The public classification scheme serves as the common notion of desirable and undesirable behavior. It includes behaviors that should be met with a costly collective reaction. The advantages of Ω_L, as opposed to customary rules, lie precisely in its unique and public character. All agents are presented with the same classification scheme, and all know that this (and no other) scheme has been presented to others. Thus, rule design can resolve the coordination problem: it can overcome the ambiguity associated with selecting one of many possible equilibria.

Two questions remain. The first question is whether agents are incentivized to use the third-party classification scheme to coordinate their activities. Each agent will do so only if they find doing so in their long-term interest, meaning that Condition (5) needs to be satisfied. This interest can be understood as the degree to which the public classification scheme overlaps with this agent's private preferences. Agents will consistently apply the logic embodied in the public classification scheme only if they expect that they would benefit from it over the entire duration of the game.

Importantly, coordination based on the public classification scheme is possible only if agents believe that both will participate in Ω_L-directed cooperation. This belief requires common knowledge about how the game should be played. For coordination to work, the classification logic must be accessible, unique, and salient. These requirements are met: the contents of Ω_L are communicated before the game begins.

However, beside the problem of coordination (which equilibrium to play?), there is still the discovery challenge (is this an equilibrium at all?). While reliance on the public classification scheme may sufficiently benefit one agent, does it also benefit others? This is the second question that must be addressed; it corresponds to the discovery challenge discussed in the previous chapter. The equilibrium with Ω_L-directed cooperation is only possible if both agents find Ω_L-directed cooperation more beneficial than the no cooperation strategy. Under incomplete information, this cannot be ascertained from the outset. Therefore, the ongoing

participation of the other party is initially uncertain. Agents need to take into account the other agent's incentives to participate in Ω_L-directed cooperation. At the beginning of the game, they are unsure whether such incentives are in place.

However, based on their beliefs about the private preferences of the other party, agents are able to infer the probability that their participation in Ω_L-directed cooperation would be reciprocated. These initial beliefs are soon confirmed or disproved. When the other agent begins exerting costly effort without receiving any immediate benefit, their incentive to take part in in Ω_L-directed cooperation becomes evident. At the same time, consistently following the directions embodied in Ω_L is the only way through which agents can sustain the belief by the other party that they are genuinely interested in continuing this cooperation. Once Ω_L-directed cooperation is launched, each agent becomes confident that others share their interest in sustaining it in the long run. In other words, they have discovered an equilibrium and also understand what form this equilibrium will take long-term.

Third-party classification schemes have an important property: inflexibility. Either the agents' beliefs are sufficient to sustain Ω_L-directed cooperation, in which case this cooperation is carried out to the fullest extent, or they are insufficient, in which case the cooperation cannot materialize at all. It must be emphasized that the full extent is crucial. Ω_L-directed cooperation is only possible if agents believe that others share their interest in sustaining this cooperation in the original, pre-configured form already known to both agents. Agents cannot selectively ignore a part of the determinations included in Ω_L without undermining others' confidence in the entire coordination mechanism. This stems from the very structure of the incentives that support the use of third-party classification schemes. The belief that *each and every* variant included in Ω_L will be met with coordinated effort in the future, provided that an agent continues to participate in Ω_L-directed cooperation, is indispensable for the existence of an equilibrium with third-party coordination. Without such a belief, Conditions (6) and (8) cannot not hold, and adhering to the classification logic embedded in Ω_L cannot be individually rational.

For example, imagine that agents are buyers in the marketplace. The code of the marketplace calls for boycotting sellers who delay deliveries, refuse to use the scale in the middle of the square, or have not paid their subcontractors. Buyers are confident that other buyers would abide by the code. However, one day, some buyers declare that they have no interest in supporting subcontractors and will continue purchasing from sellers who owe money to their business associates, while still following the rest of the code. Would it be rational for other buyers to still boycott sellers who violate the code against these defecting buyers? It would

not, because it undermines the certainty of cooperation based on the code. If one party can ignore parts of its determinations without consequence, any party would expect to be allowed to do the same. The coordination logic becomes uncertain again: how agents will behave in the future is once again an open question.

Therefore, the expectation that agents will coordinate their actions based on the coordination logic embedded in Ω_L – i.e., the expectation that underlies Conditions (6) and (8) – could not be rational. Consequently, agents would update their belief $\pi_i(t)$ to zero. This would collapse their own incentive to participate in Ω_L-directed cooperation, leading to a breakdown in coordination. To prevent this dismal outcome, rational agents, for whom the participation constraint is satisfied, are not tempted to defect in any part of the effort required by Ω_L. The fear of the decay of the cooperation pattern is a "grim trigger" that effectively prevents defection, thus ensuring the ongoing reliance on predesigned rules.

3.5. Concluding remarks

In this chapter, we presented ambient rules and designed rules as two types of equilibria in the game introduced previously. In both cases, the outcome is the development of institutions: "public," interpersonal rules that are consistently followed and enforced. These rules can be described as conditional propositions in the form "in situation X, do Y", so that agents can predict what is expected from them in specific situations. Moreover, these rules are socially implementable: they are enforced, meaning that behaviors that do not conform with rules are treated differently than behaviors that conform. If agents defy the rules, they are punished with non-cooperation by the other agent.

In each type of equilibrium, these rules originate differently. Ambient rules are customary. They arise and evolve spontaneously, with agents' understanding of desirable and undesirable behavior emerging from historical precedents that eventually stabilize into a predictable pattern of behavior. This common understanding is created when agents engage in trial-and-error processes. In contrast, designed rules provide a common understanding of desirable and undesirable behavior upfront to all parties. A unique, predefined decision-making logic serves as an external coordination mechanism. The systematic use of this logic can also

create a cooperative equilibrium: agents follow it to coordinate their efforts because they fear a collapse of coordination in the long run.

Customary rules, social norms, and similar rule-creation mechanisms have been previously modeled in the law and economics literature. However, our approach distinguishes itself from earlier contributions in several respects. Unlike most earlier models, our framework assumes incomplete information, representing an environment in which agents are uncertain about the private preferences and objectives of others. This situates our work within the literature on rules that emerge under conditions of "private assessment" (see, e.g., Okada et al., 2018; Okada, 2020), a literature that until recently has been almost exclusively anchored in theoretical biology.

Trial-and-error representations of norm emergence have been developed within the framework of evolutionary game theory (e.g., Sugden, 1986; Young, 2001; 2015; Aoki, 2001). These frameworks often depict norms as "sticky" behavioral predispositions that remain constant even when new situations arise or fresh information becomes available.[24] In these models, norms are much like genes in biological populations. They compete with each other, being transmitted across generations at varying rates. In contrast, our model adopts a rational actor perspective. Unlike the fixed behavioral schemas assumed in evolutionary models, our approach allows for agents who dynamically adjust their actions based on what they consider optimal at any given time.

Likewise, the depiction of designed rules as a tool for third-party coordination builds upon earlier contributions within the coordination accounts of law. Our analytical framework is inspired by the semi-rigorous formalizations by Postema (1982) and McAdams (2000, 2009). To an even greater extent, it follows the repeated game model used by Hadfield and Weingast (2012). However, the approach used in this chapter differs in several respects. First, we envision social rules as patterns of behavior where actions deemed socially desirable or undesirable are systematically met with collective responses, such as collective punishment or mutual aid. In contrast, the model by Hadfield and Weingast conceptualizes social order based on rules primarily as an ever-present threat of punishment, which functions off-equilibrium. In their framework, the emphasis is placed on credible deterrence, where the mere possibility of punishment for rule violation is sufficient to maintain social order.

[24] A similar notion of social norms as relatively stable behavioral predisposition can be found in empirical institutional economics. For a survey, see, Voigt, 2024a; 2024b.

4. Comparative discussion

Summary

Building on the two equilibrium concepts from the previous chapter, this chapter discusses and compares ambient rules and designed rules based on several criteria. Specifically, we consider: (i) rule specificity, (ii) capacity for compromise between conflicting preferences, as well as the capacity to accommodate vested interest and error (iii) flexibility of rules and the performance under changing cooperative environment, and (iv) performance under deferent degrees of group cohesion. We highlight important differences between these two types of institutions across each criterion. The discussion is further supplemented by examples from academic literature in anthropology, legal history, international law, and general international relations.

4.1. Introduction

While there are numerous studies on "spontaneous" institutions (see, e.g., Powell and Stringham, 2009), and several theoretical accounts on the coordinative role of law, relatively few have conducted comparative analyses between spontaneous and deliberate institution-making from the law and economics perspective. One notable exception is Parisi (2001), who compares legislation and custom as two "institutional designs of rulemaking." Parisi identifies legislation with political procedures that can be analyzed through public choice theory, suggesting that democratic lawmaking is burdened with numerous agency problems and adverse incentives, which are likely to render the resulting rules inefficient. Simultaneously, Parisi argues that custom suffers from the collective action problem: agents may face inadequate incentives to create socially efficient customary rules. Rossi and Spagano (2018) attempt to explain the codification of customary rules used among early modern merchants. They argue that the cost of disseminating rules played a major role in this process. Because it is more difficult to familiarize oneself with customary practices in larger social networks, written text began to enjoy an advantage as the relevant markets expanded.

We believe that by seriously considering incomplete information, the comparison developed in this chapter can improve upon previous studies. This chapter systematically examines how these two mechanisms for creating rules compare in light of four criteria: (i) rule specificity, (ii) capacity for compromise between conflicting preferences, as well as the capacity to accommodate vested interest and error (iii) flexibility of rules and the performance under a changing cooperative environment, and (iv) performance under deferent degrees of group cohesion.

Among these criteria, rule specificity means the internal differentiation within a rule, or the number of distinctions it makes between individual cases; more distinctions imply greater specificity. For example, rules that differentiate between cases based on *mens rea* – the intention of wrongdoing – would be more specific than those that do not. The capacity for compromise, and the potential for serving vested interests refer to the way in which situations like "agent A wants event X to happen, but agent B does not want event X to happen" can be dealt with. Flexibility is understood as adaptiveness to changing circumstances – i.e., how rules may adapt to changes in the underlying parameters of cooperation, such as parties' preferences. Finally,

cohesive groups are defined as groups in which the interests and preferences of group members are more strongly aligned, and group members are more aware of this alignment.

In short, we suggest that ambient rules tend to be simple, in the sense of distinguishing between fewer types of cases. Simpler customs are easier to communicate, understand, and imitate, and as a result they disseminate easier among agents and become expected standards of behavior. In turn, designed rules are created in advance, which makes them less constrained by the need for simplicity. Moreover, ambient rules and designed rules handle compromises differently. Conflicting interests can be accommodated in ambient rules by making them more general (less specific). Designed rules, in contrast, can incorporate compromises more openly. However, they may also embed the vested interests of certain parties or the interests of rulemakers themselves. We also portray ambient rules as "slow-moving institutions" – i.e., institutions that change gradually as the socio-economic conditions change. Designed rules, by contrast, tend to remain in a relatively constant form until the cooperation they facilitate abruptly breaks down. Finally, we suggest a direct relationship between group cohesion and the specificity of ambient rules: they emerge more complex in more cohesive groups, whereas less cohesive social environments tend to produce simpler ambient rules.

In the course of the discussion, we occasionally illustrate our theoretical findings with examples from legal anthropology, legal history, and international law.

4.2. Specificity

There are multiple ways to understand the specificity of social rules. One way is to interpret specificity as internal differentiation: the degree to which rules adjust depending on circumstances. In this sense, general rules apply uniformly across a wide range of situations, while specific rules are tailored to particular conditions (Schuck, 1992; Kaplow, 1995; Mahoney and Sanchirico, 2005).

For example, consider tort law. Strict liability is a relatively general rule because it assigns liability based solely on causation. Whoever caused the delict is liable for the damage. In contrast, negligence is more specific, as it takes into account an additional factor: the amount of care exercised by the tortfeasor. It distinguishes between cases where sufficient care was exercised and those where it was not. Under the rule of negligence, the person who caused the

delict is liable only if they were negligent; otherwise, they are not. Contributory negligence is even more specific. It differentiates between cases based on the diligence – or lack thereof – separately exercised not by one but two parties. If one party was negligent, the rule states that it must be considered whether the other party was also negligent. This added layer of differentiation further complicates the liability rule, making it more specific in the sense described above.

The specificity of rules can also be represented within the formal framework developed in Chapter 2. This can be done by "bundling" variants. Bundling means that a group of variants is treated as equivalent: if the rule requires exerting effort in one variant within the bundle, it also requires effort in all other variants in the same bundle. For example, imagine that variant k_1 is a slightly late delivery of goods by a seller. On the other hand, variant k_2 may denote a delivery with a more significant delay. If a rule develops in such a way that it always demands that both buyers respond the same way to variants k_1 and k_2, it can be considered more general (less case-specific) than a rule which, for instance, requires that sellers are collectively punished by the buyers when they deliver with significant delay but not with minor delay.

How does specificity relate to ambient rules and designed rules? Based on the analysis in the preceding sections, we propose that ambient rules and designed rules differ in their capacity to handle specificity. Our argument is straightforward: the processes through which ambient rules are formed tend to promote rules that are general rather than specific. By contrast, deliberate design allows for the development of rules with higher levels of specificity and sophistication.

We begin with ambient rules. The argument will use a slightly modified equilibrium presented in the previous section. Instead of focusing on an equilibrium where each variant k of the social interaction is treated separately, we now consider a class of similar equilibria. The argument is simple: imagine that agents can cooperate by treating multiple individual variants as elements of a single bundle. If this is possible, ambient rules in which agent bundle individual variants have lower informational requirements: they are easier to emerge and disseminate, even at the expense of the efficiency of the long-term equilibrium. This, as we argue, makes such general ambient rules more likely to develop than specific ones.

The argument can be expressed in terms of the model. Assume that $K \geq 2$ variants of the social interaction can be viewed as special cases within a larger bundle of variants. Further, assume that Agent i cares about a fraction ρ of these variants. This fraction is straightforward

to interpret. For example, a buyer might only care about certain types of delivery delays, such as significant delays or delays involving goods critical to the buyer's trade. The cases in which such delays are expected from sellers are the fraction ρ of all delivery delays.

We now generalize the logic used previously in Chapter 3. Assume that one of the K variants within the bundle occurred in the current period and Agent i is in the role of the contributing party. Moreover, Agent i believes there is a probability p_B that the other agent will cooperate by exerting effort across the entire bundle of K variants once Agent i begins doing the same. In other words, Agent i believes that reciprocity-based cooperation for the bundle of K variants will materialize with probability p_B. Under this assumption, it is rational for Agent i to exert effort in the current period if the following condition is satisfied:

$$-(1-\delta)c + \delta p_B \frac{\rho g - 2c}{2}\frac{K}{N} > 0 \tag{1a}$$

The interpretation of Condition (1a) is identical to the interpretation of Condition (1) earlier. Condition (1a) can be split into the immediate cost of effort and the continuation payoff. Once the cost of effort c is incurred, Agent i expects that, with probability p_B, both agents will cooperate in the fraction of $\frac{K}{N}$ variants in the remainder of the repeated game. Agent's i expected pay-period payoff in the periods when one of the K variants occur amounts to $\frac{\rho g - 2c}{2}$. Unlike in the case in which agents cooperated in making effort in a single variant, the benefit g is multiplied by ρ because Agent i cares only about the fraction ρ of the variants within the entire bundle. Therefore, only values of ρ such that $\rho g - 2c > 0$ are admissible. As was the analogous condition in the previous chapter, the current period and the continuation payoff are weighted with $1-\delta$ and δ, respectively, which represents the value Agent i attaches to the future relatively to the present moment.

Condition (1a) can be translated into a condition for p_B:

$$p_B > \frac{1-\delta}{\delta}\frac{2Nc}{K(\rho g - 2c)} \equiv p_B^* \tag{2a}$$

Condition (2a) is analogous to (and is a generalization of) Condition (2). As was argued previously, this inequality must be satisfied for both agents in order for reciprocity-based cooperation to emerge.

It is clear at this point that we apply the same reasoning to bundles of variants that was used for individual variants. This method can also be extended to analyze signaling. As in Chapter 3, assume that condition (2a) is not met for Agent $j \neq i$, but Agent i believes that, once it is satisfied, Agent j will cooperate in all K variants with probability p_B. Based on this assumption, we can derive the necessary condition that motivates Agent i to engage in one-period signaling:

$$-(1-\delta)c + \delta p_B \frac{\rho g - 2c}{2}\frac{K}{N} - (1-\delta)(1-p_B)\frac{K\delta c}{K\delta + N(1-\delta)} > 0 \qquad (3a)$$

If Agent i makes effort in the current period, this agent needs to incur the cost c. Starting from the next period, he then expects a per-period payoff of $\frac{\rho g - 2c}{2}$ in the fraction of $\frac{K}{N}$ of all periods – the utility gain g is multiplied by ρ because Agent i is interested only in the fraction ρ of variants in the bundle. This expectation materializes with probability p_B. However, with probability $1 - p_B$, Agent i expects that the other agent will not cooperate in any of the K variants within the bundle, in which case he needs to expend the cost c in vein once again. The period in which this cost will be incurred is discounted with $\frac{K\delta}{K\delta + N(1-\delta)}$.[25] On the other hand, in case Agent i decides not to signal, no utility is gained or sacrificed, as shown in the right hand-side. Like in the original case without bundling the limiting value p_B^{**} that satisfies (3a) must be greater than p_B^*.

It is easy to see that expressions (1a)-(3a) closely resemble expressions (1)-(3) derived in the previous chapter. In fact, expressions (1)-(3) are special cases of (1a)-(3a) when $K = 1$, i.e., when there is no bundling. However, the question remains: is bundling indeed a superior strategy compared to treating variants individually? The answer to this question will be affirmative if bundling makes the conditions necessary for the emergence of ambient rules less restrictive – in other words, if $p_B^* < p_k^*$ and $p_B^{**} < p_k^{**}$.

[25] I.e., $\sum_{t=0}^{\infty} \frac{K}{N}\delta^{t+1}(\frac{N-K}{N})^t$

We start by showing that the condition necessary for reciprocity-based cooperation indeed is less restrictive with bundling, i.e., $p_B^* < p_k^*$. First, notice that for both (1a) and (3a) to hold, the condition $\rho > \frac{2c}{g}$ must be necessarily satisfied. If this is not the case, then $\rho g - 2c \leq 0$, which makes the left-hand sides of (1a) and (3a) negative. This requirement means that an agent must have a sufficiently high interest in reciprocity-based cooperation across the entire bundle of K variants to even consider doing so.

It can be shown that the sufficient condition that makes reciprocity-based cooperation with bundling preferable is either identical to the necessary condition indicated above, or is only slightly more restrictive. Notice that if $\rho > \frac{2c}{g}$, then $p_B^* < p_k^*$ translates to the following condition:

$$\rho > \frac{2c}{g} + \frac{g - 2c}{gK} \tag{10}$$

To see that Condition (10) is, in fact, only slightly more restrictive than mere $\rho > \frac{2c}{g}$, we can rewrite ρ as $\frac{R}{K}$, where R is the number of variants within the bundle of K variants that Agent i cares about. Then Inequality (10) can be written as:

$$R > \frac{2c}{g}K + \frac{g - 2c}{g} \tag{11}$$

Note that $\frac{g-2c}{g} < 1$, and R is an integer. If R satisfies $\frac{R}{K} > \frac{2c}{g}$, it either also satisfies Condition (11), or $R + 1$ does. In other words, if the bundling strategy meets the basic criterion of rationality (i.e., $\rho > \frac{2c}{g}$), then most likely reciprocity-based cooperation with bundling can be executed in conditions in which it could not be executed without bundling.

Moreover, in some cases, bundling also relaxes the condition necessary for signaling. This means that in many cases (but not all), $p_B^{**} < p_k^{**}$. We will use the fact that Condition (3a) is a generalization of Condition (3). Just like it was the case with Condition (3), the left-hand

side of Condition (3a) can be separated into two additive parts. The first part is identical to the left-hand side of Condition (1a), while the second part is $-(1-\delta)(1-p_B)\frac{K\delta c}{K\delta+N(1-\delta)}$. It has just been shown that the values of p_B for which the left-hand side of Condition (1a) becomes zero is lower than the value of p_k for which the left-hand side of Condition (1) was zero. However, the opposite is true for the right-hand side of these conditions. The right-hand side is an increasing function of p_B that takes non-positive values. With higher K, these values become lower, and larger p_B is needed to compensate for it.

In sum, bundling has two opposite effects on signaling requirements: on the one hand, it makes long-term cooperation viable beyond the options available without bundling. This makes the conditions for signaling less restrictive. Such is the interpretation of the left-hand side of Condition (3a). On the other hand, bundling increases the cost of a mistake, represented by the right-hand side of Condition (3a). If the other agent does not respond to the initial signal, the signaling party incurs the cost of being mistaken sooner, which tightens the condition necessary for signaling. In general, it is uncertain which effect prevails. However, even when signaling is not more viable with bundling, reciprocity-based cooperation still is.

The previous couple of paragraphs can be now informally summarized. The generalization of the model of ambient rules that permits bundling leads to an important conclusion: bundling enables agents to cooperate in scenarios where they have partially shared, but partially divergent private preferences. The original strategy described in Chapter 3 assumed that agents can cooperate only when they are homogeneous; the obstacle to cooperation lies in only the fact that the agents do not know that they are homogenous. They do not know, and thus need to discover, the cases in which they share the assessment of specific variants of the social interaction.[26] In other words, in situations where variants of the social interaction are taken separately – such as those depicted originally in the previous chapter – cooperation was limited to cases involving homogeneous parties. Such parties already share many normative views and will cooperate if they become sufficiently convinced of this shared understanding. However, this is no longer true when bundling is considered. Through bundling, agents overcome their heterogeneity and can cooperate despite their differences.

The reasoning behind the last few paragraphs can also be summarized in less technical terms: it requires more trial-and-error to learn a specific customary rule from observing and

[26] I am grateful to an anonymous reviewer from the *Review of Law and Economics* for this observation.

emulating agents' behavior than to learn a simpler one. Therefore, such lengthier, more burdensome, and costlier procedures are less likely to be undertaken. In other words, we suggest that the emergence of ambient rules may be facilitated by reduced norm specificity – i.e., by bundling. This is true even if such a reduction makes collective action less efficient in a narrow sense, meaning that it allows for achieving only a fraction of economic surplus. The view that customary rules tend to be general finds support in the scholarship inspired by evolutionary game theory (e.g., Alexander, 2007) and is consistent with our theoretical exposition.

As H. Smith (2009) noted, community customs often evolve in informationally efficient ways by favoring repeatable practices that require minimal situation-specific knowledge. This evolutionary tendency is exemplified by the emergence of the social norm of strict liability for cattle trespass among ranchers and farmers in Shasta County, California (Ellickson, 1991). Implementing more complex negligence rules would require information about situation-specific circumstances, such as the precautions taken by the rancher or farmer, which would impose significant informational costs on those seeking to learn and apply the same social norm in the future. In contrast, strict liability rules are easier to formulate, communicate, and observe in practice. Thus, Smith argues that custom tends towards simplification:

> "[T]he message to keep off, the default regime of possession, is easier to communicate to more far-flung parties. (…) The general, formal default – here the norms of trespass and exclusion – have a gravitational pull, beyond the benefits in terms of the relative importance of farming compared to ranching. Even where a more nuanced flip of the rule from fencing in to fencing out would be efficient in the narrow sense of maximizing the value of the two activities, in close cases there is an additional reason to stick with the general rule of fencing in, because it comes along with the general exclusion regime for ownership of land, which is undemanding from an informational standpoint." (H. Smith, 2009:27)

In other words, Smith points out that customary rules often evolve in a way that minimizes the amount of information required from agents to familiarize with the relevant practices. This means that they involve practices that are repeatable and require low amount of situation-specific knowledge. Even when such practices could be in principle replaced with

more efficient ones, the value of replicability prevails. However, the same argument suggests that simplification can lead to the establishment of less efficient rules than other methods could. For example, a well-known result from the economic analysis of accidents indicates that negligence rules are often preferred over strict liability rules on efficiency grounds (Dari-Mattiacci and Parisi, 2006). Relying on simpler strict liability rules may forfeit the efficiency benefits that could be achieved through more fine-grained rules.

The same logic can be applied to customary international law. Customary international law can be broadly defined as a set of rules that states perceive as binding because they are supported by sufficiently widespread or accepted practice. Thus, at least in its pure definitional form, customary international law structurally resembles ambient rules (see Chapter 1). Precisely for this reason, scholars have noted that customary international law may fail to adequately guide states' behavior toward cooperative equilibria when the number of possible equilibria is large (e.g., Goldsmith and Posner, 2005). In other words, customary international law may fail to successfully address the coordination problem – i.e., to unequivocally indicate one of the many possible ways in which states are expected to cooperate – when too many such ways are conceivable. This potential failure has led Goldsmith and Posner to conclude that:

> "[Customary international law] permitted states to attack a fishing vessel that is a threat, but what counts as a threat? A vessel that contains weapons that are being transported from here to there? A vessel whose sailors might be spies, or who might simply report the whereabouts of the enemy navy to their own forces? (…) These problems of ambiguity (that is, multiple equilibria) have always made customary international law very weak and have spurred states to use treaties to clarify customary international law." (Goldsmith and Posner, 2005:41)

Their point is that customary international law is incapable of providing the specific details necessary to address case-specific situations. Instead, countries that wish to establish more specific rules need to rely on treaties. The natural corollary to this statement – though not explicitly voiced by Goldsmith and Posner – would be the implication that customary international law, if effective at all, consists of rules that are less specific than the provisions of international treaties. Rule of customary international law – as long as they remain customary

instead of being written down – should be skewed in favor of norm generality over norm specificity. This argument is sometimes made directly, although not by law and economics scholars. Varella (2013) associates the rise in complexity in international law with its progressive institutionalization, formalization, and codification, which, among other things, means that custom gradually plays a less prominent role than it did in the past (see Chapter 3, Section 3.2).

Compared to customary rules, deliberately designed rules appear better equipped to accommodate specificity. In principle, they can differentiate between cases across any number of types or classes. As long as the public decision-making logic is announced in advance, allowing the parties to reconstruct the implied common notion of desirable and undesirable behavior, complexity a limiting factor is less important. This idea is corroborated in our analytical framework. The incentives to use designed rules (i.e., Conditions (5) to (9)) do not depend on the specificity of these rules (i.e., on N) but only on how closely they align with agents' private classification schemes.

The ability of designed rules, such as those found in legislation, to accommodate high degrees of specificity has been used to justify making use of this capacity in crafting legal rules. For example, De Geest (2013) formulates a "N instruments for N problems" principle in contract law by claiming that:

> "[I]f we want to solve – say – 8 different problems (creating incentives for optimal breach, reliance, precaution, mitigation, incentives to reveal unusually high potential losses, incentives to promise carefully, incentives not to opportunistically renegotiate, and optimally allocating risks) we need 8 separate rules or doctrines that each address one of these problems rather than trying to solve them all with the choice of a single remedy." (De Geest, 2013:43)

Applying this principle would often require highly case-specific rules that treat individual cases differently based on a large number of efficiency demands. In the example of 8 problems, the law should recognize that cases differ across at least 8 dimensions, meaning at least 8 independent rationales for differentiating rules applicable to individual cases. In other words, a rational legal design should recognize "the strong disadvantages of using fewer

instruments" because "compromise instruments (i.e., single instruments that are meant to solve multiple problems) cannot be fully effective at solving two or more problems when there is tension between them." (De Geest, 2013:43,45)

The operation of a similar principle can sometimes be observed in legal history. In describing archaic legal systems based on unwritten customary sources, Parisi notes that:

> "(…) liability was to arise as a consequence of a loss that was causally related to a human act. The subjective elements of the injurious behavior were factors of a second moment. Imputability of the actor and voluntariness of the act were presumed or simply ignored." (Parisi, 1994:319)

However, legislative intervention often leads to increased specificity within social rules that were originally based on custom. A notable example is "blood money" – i.e., a liability rule for homicide, which was common in primitive legal systems in many parts of early-to-mid medieval Europe. The institution of blood money required the killer to pay a predetermined amount of money – often referred to as the value of blood spilt – to the victim's kin, allies, or other support group. The payment of blood money effectively prevented a feud that would otherwise ensue or ended one that was already underway. Importantly, the compensation amount was typically based on one variable only: the victim's status, giving no consideration to factors like the needs of the victim's kin, the perpetrator's ability to pay, or, arguably, even culpability.

However, the history of culpability for homicide among Anglo-Saxons in England (Robinson, 1980) and the history of Scottish feuding practices (Wormald, 1982) suggest that these additional factors began to play a significant role as a result of royal legislative interventions. In the Scottish case,

> "[r]oyal writs of the thirteenth and fourteenth centuries (…) offered protection to the man who had killed in self-defence. (…) And a provision in *Regiam maiestatem* carefully distinguished between the amount of compensation owed when a horseman in a village rode down and killed a pedestrian, and when he killed one by backing his horse into him; compensation for the first

(…) was far greater than for the second, when the pedestrian should have been more careful." (Wormald, 1982:111-112)

However, the discussion above needs several qualifications. First, the assertion that rule efficiency can be enhanced by increasing specificity does not imply that greater specificity *always* results in improved efficiency. The relationship between rule specificity and efficiency has been a subject of debate within the law and economics scholarship. Scholars have often concluded that there exists a nuanced balance or "sweet spot" of optimal specificity, beyond which it becomes either excessively high or inefficiently low (e.g., Kaplow, 1995; Wright, 2000).

Moreover, our discussion deliberately omits the cost of designing rules. Researchers point out that more case-specific rules are more costly to make because lawmakers face an increasing "difficulty of specifying the contingencies of a complex environment." (Fon and Parisi, 2007:152) Thus, taking the rulemaking cost into account would allow to paint a fuller picture of the costs and benefits associated with deliberate and customary mechanisms for rule creation. However, because we are primarily interested in the *qualities* of rules that may arise under both regimes, designed rules are treated as endogenous, and the cost of rule provision borne by the rule provider is disregarded.

Also, we abstract from the learning burden associated with excessively specific or complex rules. Agents who are unfamiliar with the intricate network of rules, legal doctrines, and methods of normative reasoning may find it difficult to acquire the necessary knowledge, or they may do so only at prohibitive costs. In extreme cases, the information burden may undermine the common knowledge assumption – that all parties know the rules, know that others know them, and so forth. Since such common knowledge is essential for the equilibrium in which agents coordinate their actions based on a third-party classification scheme, excessive rule specificity may render the equilibrium impossible. While this possibility must be recognized, it has not been included in the earlier considerations.

Finally, the discussion above requires a point of clarification. The claim about the tendency of ambient rules toward "generality" (i.e., low specificity) is limited to of specificity as characterized at the beginning of this section. It should not be mistaken for a claim that ambient rules tend to be general in the sense of embodying "standards" rather than "rules." A prominent perspective on the complexity of rules frames it as an alternative between "rules"

and “standards” (e.g., Kaplow, 1992; Fon and Parisi, 2007), in which both represent opposite poles of precision. In this sense, a rule entails “an advance determination of what conduct is permissible, leaving only factual issues for the adjudicator” (e.g., by mandating that a seller who delivered goods with delay must pay a compensation to the buyer).

On the other hand, a standard leaves “both the specification of what conduct is permissible and factual issues for the adjudicator” (Kaplow, 1992:559-560) (e.g., by requiring that a seller who seriously hampered buyer’s trade with his actions must pay a compensation). The model from Chapter 3 does not directly address the distinction between rules and standards. However, if the logic of information costs was to be applied again, the previous argument about ease of demonstration and imitation would suggest that ambient rules are more likely to appear in the form of highly precise (and thus easy-to-imitate) rules rather than vague standards.

4.3. Compromise, susceptibility to manipulation by vested interest, and susceptibility to errors

It has long been recognized that negative externalities are one of the primary sources of inefficiency of custom. Parties who do not participate in rule creation can suffer disutility when their interest is not taken into account by those who create them (Coleman, 1990; Posner, 2000). For example, a customary rule developed by landowners regarding waste disposal, such as dumping the waste into the river, might impose environmental costs on riverside communities that were not involved in establishing this rule. Nonetheless, as long as only participating parties are considered, the potential for the inclusion of rule components that benefit one group of agents at the expense of another is limited. Before an emerging custom solidifies, agents may disengage their interests would be impaired otherwise.

The realities of customary international law seem to corroborate this suggestion. The influential (though sometimes controversial) persistent objector doctrine in international law allows each state to voice an objection to an emerging practice during its formative phase. By doing so, the objecting state can exempt itself from being bound by an undesirable custom (Green, 2016). This corresponds to the possibility, identified in our formal framework, of deliberately withholding cooperation during the formative phase of customary rules. Withholding cooperation in variant k of the social interaction during this variant’s first and

second historical occurrence conveys a message that the agent does not want effort in this variant to become socially expected in the future.[27] This mechanism allows agents to directly influence the content of ambient rules. By refusing to participate in practices they find objectionable, agents ensure that such practices do not become established as behavior expected from agents in the long run.

However, even if ambient rules evolve is such a way that they benefit all parties, the important issue that remains is how these benefits are distributed – i.e., which party gains more than others. Knight (1992) claims that the distributional effects of spontaneously developed institutions depend on the bargaining positions of the involved agents. The party with a safer fallback option, i.e., the party that can do relatively better without cooperation, is in a position to shape such institutions in its favor. Technically, this implies an equilibrium selection mechanism: which rule is observed in reality depends on what happens when agents fail to coordinate on one of the mutually beneficial institutional outcomes.

Rules regulating intergenerational transfer of property serve as an example of Knight's logic. Typically, all family members must contribute to their joint economic well-being, and failing to respect inheritance norms could lead to a collapse of familial cooperation. If this happens and a family member is alienated or cast out, this members needs to survive on their own. However, whose interests are favored by rules of inheritance depends on the relative bargaining positions of the family members, i.e., on how well they can manage without the familial support. For instance, children with sources of income other than the family estate may secure more favorable succession rights, such as a guaranteed share of inheritance for every child. Their fallback option, which they can rely on if familial cooperation breaks down, allows them to make demands more comfortably. Conversely, children whose material well-being throughout their adult lives strongly depends on the family estate are in a weaker position and may be subject to less favorable rules, such as primogeniture, which favors the eldest son – the child whose effort typically has the highest immediate value to the family head.

The model presented earlier does not currently incorporate a similar reasoning, but this could be done after a slight modification. Consider a similar model with three agents instead of

[27] According to the persistent objector doctrine, objections must be explicitly voiced; silence of a state is insufficient for this state to be recognized as an objector (Green, 2016). Our formal framework cannot perfectly recreate explicitly voiced opposition. The framework includes only two agents who take part in interactions in every period, without inactive observers. Therefore, withholding cooperation in the formative phase, before the other agent develops a solid expectation of cooperation in a specific variant of the social interaction the future, is the only option to communicate that the customary rule should not evolve into demanding cooperation in such cases.

two. In the new setup, cooperative outcomes can be achieved in any given period not only when all agents exert effort, as previously assumed, but when the total sum of efforts exceeds a critical mass of ½. Further, assume the agents have unequal contributions to make: Agent 1 can contribute ½ of the total effort, while the remaining two agents can each contribute only ¼. This disparity in effort may result from differences in economic productivity, power of the agent's security network, political power, or other exogenous factors.

Given this disparity, it is natural to expect that the common notion of desirable and undesirable behavior that develops on a customary basis would lean toward one preferred by Agent 1. This is because Agent's 1 effort can be effectively combined with that of *either* of the other two agents, providing a favorable fallback if one of them chooses not to cooperate in some variants of the social interaction. In contrast, Agents 2 and 3 need Agent 1 to achieve cooperation in any variant they care about and have no reasonable fallback option. However, if the underlying economic or political factors change, giving more power to Agent 2 and less to Agent 1, ambient rules should adjust accordingly, now in favor of Agent's 2 preferences. An important conclusion is that a change in exogenous economic conditions is both sufficient and necessary to trigger a change in ambient rules: these rules respond to such conditions and, after the formative phase is complete, they do not change unless those conditions shift. More on this topic will be said in the next section.

On the other hand, designed rules can be exogenously changed even if the underlying economic forces remain constant. The change occurs simply by reshaping the publicly accessible content of Ω_L (e.g., through announcing a new legislative act). This recreates the unique focal point around which the common notion of desirable and undesirable behavior is centered. In this way, the incorporation of rule components that favor one party or group but disadvantage another seem to be compatible with deliberate mechanisms for creating rules. This is caused by the already mentioned systemic and inflexible nature of strategies that rely on focal points. Agents will engage in Ω_L-directed cooperation if they have *sufficient* incentives to do so, and have no option of selectively rejecting its particular elements.

The possibility to recreate focal points, combined with the systemic nature of strategies that rely on focal points, widens the possibility of asymmetrical rule manipulation. The content of Ω_L can be changed to align more closely with the interests of one party or group at the expense of another, even when all the underlying economic forces are stable. The possibility of incorporating such asymmetries into designed rules allows for designing rules that improve efficiency in the Kaldor-Hicks sense: total utility is improved, but at least one party is

disadvantaged in the process.[28] Thus, when satisfying the preferences of one party or group at the expense of another one is socially productive, design becomes a superior mechanism for rule production from the efficiency standpoint.

Likewise, design can be superior when agents are highly heterogeneous and their conflicting interests cannot be easily reconciled. In this case, as was argued above, customary methods of rule-making sacrifice the utility resulting from rule specificity for the ease of communication and dissemination of rules. Design as a rule-making method does not need to make this sacrifice. Helfer and Wuerth make a similar suggestion in the context of international treaties negotiated beforehand by the states that are subsequently bound by them:

> "Most of these agreements contain carefully crafted compromises, often hashed out in exquisite detail among cross-cutting alliances. A group of states may give up a preferred position in one section of a treaty (or in one treaty in a nested treaty regime, such as the WTO) in exchange for benefits in another section." (Helfer and Wuerth, 2016:576)

Interestingly, the same idea has been voiced by a legal historian, but the context is slightly different. In his description of the first codifications of Greek laws, G. Smith acknowledges that drafting of legal codes in Greek colonies, which is where the first Greek legal codes were developed, was frequently an attempt to reconcile conflicting interests of colonists of diverse cultural backgrounds:

> "In many cases codification was imperative because the members of a colony were recruited from different cities. Hence no single set of customary laws could be entirely satisfactory even if it proved suitable to the new conditions. The first codes were made in the western colonies which were farther from

[28] The term "total utility" should not be interpreted in an additive sense, as if it were the summation of individual utilities. This interpretation is, of course, impossible – utility is the degree to which subjective desire satisfaction is fulfilled, not an extensive or measurable magnitude. Instead, the improvement of "total utility" can be understood as a scenario in which parties, under no costs for information, negotiation, and contract enforcement, could choose between outcomes and make side payments afterward. If, under such conditions, the parties would choose distribution A over distribution B, then a movement from distribution B to A represents an improvement in total utility.

> Greece than the eastern colonies both in distance and in the difficulty of the voyage." (G. Smith, 1922:188)

However, the possibility of manipulating the content of designed rules, regardless of the underlying economic fundamentals, also allows for socially inefficient rule manipulation. If a subgroup of agents can influence the content of Ω_L, they may change designed rules in their favor, no matter the efficiency considerations. This scenario can be identified with regulatory capture. Of course, such capture is more likely in the domestic context, where only a small subgroup of agents can be involved in the creation of rules. Moreover, beyond a regulatory capture by a subgroup of agents, there is also a hypothetical possibility of rule manipulation that leaves *all* interested parties worse off. In terms of the model, while both agents would suffer a utility loss as a result of such rule manipulation, it would still not undermine an equilibrium with Ω_L-directed cooperation; sufficient incentives to cooperate based on the public classification scheme still exist.

A manipulation of the rule content that leaves all parties worse off may happen when the rule creation mechanism itself becomes captured by the agents operating it – e.g., by legal professionals benefiting from the existence of overly complicated and troublesome rules. The possibility that self-serving representatives of the legal system may, once this system becomes sufficiently bureaucratized and centralized, distort it to the detriment of the general population has been seriously considered by prominent legal scholars. For example, Raz (1994) claims that

> "[A]s a result of the growth of a legal profession and a highly articulated legal culture, legal issues are formulated in technical terms, caught in legal categories which are far removed from the way ordinary people understand their conduct and interactions with others. The law becomes financially inaccessible and conceptually remote and alienating." (Raz, 1994:372)

Likewise, Ogus (2002:434) argues that "[t]o the extent that they have monopolistic power, lawyers can exploit the key features of legal culture to extract rents: the law used can be more formalistic, more complex and more technical than is optimal". In a similar vein, Ruhr and Ruhr (1997:456) notice a trend toward increasing legal complexity and a movement away from the idea of law as a tool for regulating relations between persons towards a tool for

accomplishing predefined policy objectives. In the United States, this trend is observable most prominently in three main fields of law. In environmental law, the was a gradual departure from "a nuisance-based common law system to a highly developed federal statutory structure." The authors suggest that the innovation in environmental law was beneficial for particular professions:

> "[W]hile investment in legal structure as a strategy unquestionably has solved many problems of environmental degradation, it also has spawned large new regulatory bureaucracies, breathtakingly complicated regulatory schemes, masses of specialized legal and engineering consultants, new curriculums in law and engineering schools, new technologies and methodologies, and so on." (Ruhr and Ruhr, 1997:460)

All in all, the very nature of coordination by focal points, whose ongoing viability depends on designed rules being *sufficiently* (but not necessarily perfectly) efficient, and which can be manipulated easier than the content of customary rules, invites the possibility of asymmetrical treatment. This possibility is double-edged: on the one hand, it enables Kaldor-Hicks efficiency improvements that are difficult to incorporate into customary rules. On the other hand, it allows socially inefficient manipulation that serves vested interests. In short, the capacity for efficiency that goes beyond the limitations of customary rule-creation mechanisms can simultaneously accommodate significant inefficiency.

The argument presented in the previous paragraphs can also be extended to include unintentional errors committed in the design of rules. A mistake in devising rules may lead to outcomes that differ from, or even contradict, the intended results. While occasionally such design errors may produce socially beneficial outcomes, in most typical cases such "unintended consequences" frustrate the original, socially beneficial intention (see, Baert, 1991).

It can be inferred from the previously presented model that social rules developed through deliberate design are more susceptible to errors resulting in adverse outcomes compared to ambient rules. This is because the participatory nature of custom formation allows the interested parties to express objections verbally or communicate them through acts of non-cooperation. This process reveals the dispersed knowledge of potential negative consequences, even if the exact nature of these consequences is not communicated directly. Conversely,

deliberate rule-making lacks such feedback mechanisms, which impairs its ability to filter out mistaken ideas. Quite the contrary: because the function of focal points is unhindered even with (some) mistakes, they are easily overlooked and difficult to correct. In other words, despite the presence of significant mistakes, coordination by focal points will still be preferred to completely rejecting this coordination method.

4.4. Flexibility of rules and stability of the cooperative environment

One of the qualitative features of spontaneously evolved institutions, such as customary rules, often discussed in legal scholarship, is their flexibility (e.g., Kadens 2012a). In this context, flexibility refers to the capacity of institutions to adapt to changing circumstances. Institutions are considered inflexible when they remain unchanged despite shifts in socio-economic conditions. For instance, a liability rule for animal trespass is inflexible if it remains intact despite significant changes in the relative market value of farming and ranching. Conversely, institutions that adjust to such changes are considered flexible.

There are more sophisticated accounts of institutional flexibility. Roland (2004a; 2004b) proposed a well-known theory distinguishing between slow-moving and fast-moving institutions. According to this theory, institutions can be grouped into two main categories: slow-moving and fast-moving. They differ in two respects: the first is the nature of how they change – i.e., continuously or abruptly. The second is how often they change – i.e., frequently or rarely. Slow-moving institutions change gradually over time in a continuous and steady way. The rules and behavioral expectations conveyed in such institutions remain relatively stable from one period to the next, with little noticeable change over short periods, even though change does occur. Prime examples of slow-moving institutions are those embedded in culture, such as social norms, values, and usages. On the other hand, fast-moving institutions change abruptly and significantly, though these changes happen infrequently and at irregular intervals. After long periods of stability, fast-moving institutions may experience sudden and dramatic transformations, sometimes even overnight. These transformations may be non-continuous, with little similarity between subsequent forms of the same institution. In this context, domestic political systems are key examples.

While the classification into slow-moving and fast-moving institutions has been portrayed as binary for expository purposes, it is more plausible to think of institutions as existing on a spectrum between the slowest-moving and fastest-moving extremes. For instance, norms shaped by evolution, such as the desire for social connection and the value people attach to reciprocity, can be considered among the slowest-moving, possibly permanent, features of the social realm. Values determined by culture and upbringing but not evolutionary programming of the human mind, while still slow-moving, can be placed closer to the other extreme, and so forth.

Building on this idea, we propose that ambient rules are relatively slow-moving compared to designed rules, which are faster-moving in the abovementioned sense of this term. As we argued, ambient rules can gradually adapt to changes in underlying socio-economic conditions, such as the preferences of agents or the cost-benefit ratio. As these conditions evolve, the content and specificity of ambient rules adjust to match the features of the environment in which individuals interact. Our analysis in Chapter 3 and the previous subchapter supports this characterization and provides a rationale. As was shown in Chapter 3, the incentives for agents to cooperate in any variant k of the social interaction vary with parameters such as g (the utility benefit from exerting joint effort), c (the cost of effort), and δ (a discount factor incorporating things like the rate of time preference, frequency of contact, and the likelihood of future interactions). These parameters represent the socio-economic conditions influencing cooperation. For instance, if Agents 1 and 2 are buyers in the marketplace, a higher parameter g means that successful purchases without product flaws or imperfections are more valuable.

We have argued throughout that the strategy of forming ambient rules leads to rules where both the scope of cooperation and the specificity of rules depend on these parameters. Generally, higher utility benefit g from successful cooperation, lower cost of effort c, and higher value attached to the future relative to the present, represented by δ, encourage a broader scope of cooperation but also the emergence of more situation-specific rules. Likewise, a lower utility benefit g, higher cost of effort c, and lower discount factor δ result in less cooperation and the development of more general ambient rules. This translates to the following suggestion: customary rules in societies populated by members with low individual productivity (i.e., low g) should be numerically few and "crude" in the sense of recognizing relatively few special cases, exemptions, and sub-cases. Once the productivity of group members increases, the content and specificity of ambient rules should adjust accordingly. Likewise, customary rules

of international law among states who heavily depend on each other in trade or security should be more sophisticated and detailed than analogous rules within a group of states whose economic well-being and physical security is less dependent on other states it the group.

In contrast, we characterize designed rules as slow-moving. The reason is simple: whether they can function or not depends on meeting a single fixed condition: $\pi_i(t) > \pi_i^*$. As long as this condition holds, agents find it beneficial to follow these rules in their preannounced form. If the cooperative environment changes – for example, if Agent i receives new knowledge that affects this agent's valuation of certain outcomes – there are only two possibilities. Either the condition $\pi_i(t) > \pi_i^*$ still holds despite the change caused by new knowledge and Ω_L-directed cooperation continues, or this conditions no longer holds, Agent i stops behaving in the way demanded by the public classification logic Ω_L, and Ω_L-directed cooperation collapses entirely. In short, this rigidity, or "all-or-nothing" characteristic, situates designed rules more closely to fast-moving institutions.

So far, we have discussed the flexibility of ambient rules and designed rules in the face of *unexpected* shocks to the cooperation parameters. We discussed how flexible both institutional solutions are when the cost-benefit ratio of activities changes without agents' knowing that the change is coming; or when an agent acquires new knowledge without expecting that some new knowledge can be acquired, etc. Now, we turn to the effects of another type of change: anticipated change. The conditions relevant to agents' decision-making may occasionally shift, and agents may have an understanding of how stable or unstable these conditions typically are. For example, a new manufacturing technology might make timely delivery less critical by tolerating longer input delays. As a result, punctuality could lose some or all of its value. While the general rate of technological development might be known to the involved parties, they cannot predict the specific effects of future technologies on their activities until the technologies are discovered.

The anticipated degree of instability – or the anticipated rate of change in the underlying conditions of cooperation – plays a key role in the analysis. Using the model from Chapter 3 again, assume that, while the agents cannot know the future states of the cooperative environment, they have an idea about how often this environment reshuffles. They do not expect the specifics of any particular change, only that changes will occur. To the extent that agents can anticipate the rate of environmental change, they understand that the durability of their current knowledge is limited. Information about possible variants of the social interaction, the preferences of other parties, and the benefits and costs associated with various outcomes, as

well as other decision-relevant circumstances, is expected to become obsolete at some point in the future.

Importantly, the anticipated rate of change is higher when the cooperative environment becomes more unstable. Consequently, the probability that current knowledge will cease to be useful in the next period increases. If the perceived probability that current knowledge becomes obsolete in the next period is denoted as $1 - \xi$, it can be directly incorporated into the discount rate: the rate at which future periods are valued relative to the present. Thus, the instability-adjusted discount rate can be written as:

$$\delta = \check{\delta}\xi \tag{12}$$

Expression (12) represents the discount rate as a product of two factors. $\check{\delta} < 1$ denotes the pure rate of time preference, reflecting how much the utility decreases when its enjoyment is postponed by one period. ξ is a factor reflecting the belief that the knowledge useful in the current period remains useful in the next period. It represents the rate at which relevant knowledge is expected to be carried over between two subsequent periods. Obviously, more volatile conditions of cooperation result in a lower value of ξ. Therefore, if agents expect their knowledge to become obsolete sooner, the total discount rate δ declines.

It is now possible to hypothesize how the anticipated stability of the cooperative environment interacts with the structure of ambient rules and design rules. We previously noted that lower δ – i.e., a lower value attached to the future – sharpens the conditions necessary for the spontaneous formation of ambient rules. In line with the logic used in the previous paragraphs, this makes the formation of less specific ambient rules more likely. In other words, we suggest that the way to remedy the effect of higher instability within the cooperative environment is to simplify the content of ambient rules, i.e., make them less specific in the sense characterized in the previous section.

On the other hand, the stability of the cooperative environment plays no direct role in shaping designed rules. However, it influences the conditions necessary for initiating Ω_L-directed cooperation which, in turn, can affects the rulemakers' design strategy. Note that Condition (9) becomes stricter as δ increases – meaning that with a higher discount rate δ, agents are less likely to find it rational to follow a given common classification logic Ω_L.

Consider a rulemaker who aims to design rules (i.e., the content of Ω_L) in a way that ensures a certain probability of agents adopting them as a focal point in their interactions. To achieve this, the rulemaker must counteract the destabilizing effects of a less stable cooperative environment: the effect of higher δ. This induces drives the rulemaker to create more equitable rules – i.e., rules that are less biased toward the perspective of any single agent or the perspective of the rulemaker himself. More equitable rules increase the likelihood that Condition (9) is satisfied, thereby offsetting the negative impact of an unstable cooperative environment and helping to maintain the probability that Ω_L-directed cooperation will materialize.

4.5. Group cohesion; personality and impersonality of rules

Finally, the discussion can progress to the role played by group cohesion. Conventional law and economics wisdom is suggest that, when third-party enforcement is absent, the establishment of social orders requires close-knit and multifaceted relationships among community members (e.g., Taylor, 1982; Ellickson, 1991). The same insights are sometimes transplanted to the study of international relations. Some world politics scholars argue that shared identities, values, and virtues among political leaders – or the affiliation with the "international community" with a distinct sense of belonging – are vital for rules-based international order (Bull, 1977; Cronin, 1999).

These close-knit relationships play two important roles. First, they enhance the flow of information, enabling the effective use of enforcement mechanisms based on reputation. In close-knit groups, members interact often, giving them more chances to observe each other and share knowledge with other actors. Strong in-group ties also make punishments more severe by providing multiple ways for members to sanction for misbehavior. The converse proposition, i.e., that lawless environments *tend to produce* communities characterized by these attributes, is less frequently asserted (e.g., Weiner, 2013; Greif and Tabellini, 2017).

Group identities and group membership also play an important role in legal anthropology. It is a recurring motif that such membership is often cascading: someone belongs to a family, which belongs to a lineage, which is part of a broader kinship structure, which is part of a tribe, etc. In the absence of third-party enforcement, the hierarchical system of group identities allows for enforcement based on mutual aid. Family members unite in disputes against

other families; lineages support their members in disputes against members of different lineages, and so on. Sometimes, group membership may be based not on birth but rather on the deliberate inclusion of a new member, such as through adoption or other forms of admission. Regardless of how membership is acquired, the general expectation is that members of each group provide aid to one another in disputes with non-members. This can be done in many ways, including coordinated fighting, refusal to share, refusal to give shelter, ostracism, etc. (Evans-Pritchard, 1940; Weiner, 2013). However, the role of group membership is not limited to the provision of mutual aid. It is equally important that group members typically bear responsibility for transgressions committed by other members. In other words, the responsibility is collective: redress (e.g., liability payments) may be obtained from any member of the group to which the original transgressor belongs, not only from the transgressor (Parisi and Dari-Mattiaci, 2004; Greif, 2004).

However, enforcement is not the only context in which group identity matters. Group affiliation also determines the rules applicable between the parties. In interactions with others, social standing vis-à-vis others plays a key role in shaping the expectations of how a given interaction should be conducted (see, e.g., S. Diamond, 1971; Pospíšil, 1971; Pershits, 1977). In cases of tort or crime, the harm inflicted upon members of the same family, lineage, clan, or other social group is treated differently from harm directed at outsiders, marking a "distinct opposition of intragroup and intergroup norm" (Pershits, 1977:410). Such dependence of social rules on the relationship between the parties, a common feature of pre-modern social systems, is often referred to as "personality" of rules (e.g., Wallis, 2011; Greif, 2004). Moreover, personality can be contextual. Group identities themselves may be defined differently, depending on the situation in question. For example, according to Bullough, among the Germanic peoples in the early Middle Ages,

> "[t]he likelihood is that the circle of kinsmen – *cognatio*, *consanguinitas* or *genealogia*, *propinqui* or *parentes* – was differently conceived not only among different Germanic peoples but locally according to custom and differently when the issue was one of inheritance of land or a monastery, vendetta and composition or who should be present at a wedding-feast." (Bullough, 1969:15)

We contribute to this discussion by examining in greater detail how group membership interacts with the content of social rules. We go beyond stressing the importance of group identities for enforcement and observing that variations in "social distance" between parties – i.e., in the minimum common identity shared by the parties – translate to different rules. Our suggestion focuses on *how* the rules differ. We propose that in tightly knit groups, spontaneously formed rules tend to be more specific. Conversely, loosely connected groups develop simpler, more general rules.

The argument supporting this claim again relies on the tradeoff between the specificity of ambient rules and the effort required to establish them. Interactions within groups that are internally connected are characterized by higher δ, because the likelihood of repeated interactions with the same agent is increased. Additionally, such groups are characterized by higher values of $\beta_i^k(1)$ – i.e., greater preexisting familiarity with agents' private characteristics relevant to their decision-making. Simply put, members of close-knit groups have "more in common," which translates to higher $\beta_i^k(1)$. For these two reasons, ambient rules that emerge in cohesive groups cannot emerge in the equally specific form in less cohesive ones. According to the logic used in one of the previous subchapters, specificity must be sacrificed, or the rule formation process will not pass the formative phase.

In other words, the analysis from Chapter 3, combined with the discussion of specificity earlier in this chapter, provides a link between the degree of group cohesion and specificity of ambient rules developed within these groups. Not only do different rules apply depending on the relationship between the parties, but their level of specificity should also vary. Closer in-group relationships tend to produce more detailed and elaborate social rules, whereas distant relationships lead to simpler, less nuanced rules.

Once again, legal anthropology can lend a hand in support of this claim. For example, Gluckman examines the customary rules governing violent conflict among the Nuer, a Southern Nile pastoralist tribe previously mentioned in Chapter 1. He concludes that as social distance between combatants decreases, the expectations regarding the conduct of fighting become more detailed and sophisticated. In more tightly connected communities, distinctions between different types of fighting are more refined: combat with certain weapons is treated differently from combat with other types of weapons, the concept of excessive violence and destruction is more complex, and non-combatants are distinguished more clearly from combatants:

> "Nuer recognize certain changes in the rules of war. Men of the same village fight each other with clubs, not spears. Men of different villages fight each other with the spear. There is no raiding within the tribe for cattle, and it is recognized that a man ought to pay cattle as compensation for killing a fellow-tribesman (…). Nuer tribes raid one another for cattle, but not for women and children who must not be killed; nor must granaries be destroyed. When raiding foreign people, women and children and even men can be captured, women and children can be killed, and granaries can be destroyed." (Gluckman, 1965a:8-9)

In a similar vein, Elias discusses the customary law of homicide among the Bantu people of Awemba. Specifically, he examines the role of intentionality: is there a distinction between deliberate and accidental killings? He concludes that a distinction is made only when the killing occurs between individuals with close social ties. In such cases, intent does matter – compensation is required for accidental killings, while murder demands blood retaliation. However, when the victim is a guest or a host from a different village, a higher standard applies automatically: every killing is treated as murder, and considerations of intent are absent:

> "Among the Awemba and some other tribes, (…) whether the offence is murder or manslaughter the penalty, which is death in the case of the former but compensation in that of the latter, is just as much a process of restoring the equilibrium. But there is no mitigation of the extreme penalty where the victim is a guest or a host (…)" (Elias, 1962:134)

These differences are also reflected in the amount of blood money required for accidental and deliberate killings (if the aggrieved family chooses to forgo revenge for the latter). The blood money for accidental killing among the Awemba is half the usual amount, except when the victim is a stranger, in which case the full amount applies regardless of the circumstances of the death. Gluckman generalizes this observation to the whole of customary law in Africa:

> "There is in these records undoubtedly the implication that whatever the motivations of the killer or the circumstances surrounding the deed, blood money had to be paid if a member of one grouping of kinsmen was killed by a member of some other group. The mental element seems to be irrelevant." (Gluckman, 1965b:205)

Until this point, the discussion has revolved around ambient rules. We suggest that, in the process through which such rules arise, there are good reasons for them to be personal and vary in specificity. But what if this personalization and variance lead to dissatisfaction? What if overly general rules applied between strangers stands in the way of the interests of some parties? Or what if identities become blurred, making it unclear which rules are applicable in a given interaction?

The response to the second point has been suggested earlier. As we noted previously, Rossi and Spagano (2018) hypothesize that, in 16th century Europe, customary rules are written down in response to the growing challenge of becoming familiar with them in increasingly diverse populations. They also provide historical evidence from the European merchant community in that era. Similarly, and in line with our reasoning from Chapter 3, the act of writing down rules can serve as a tool against the excessive personalization of rules that historically developed on a customary basis. This idea was recently suggested by the legal historian Armstrong in the context of early Roman legislation. According to Armstrong, early Roman legislation was an attempt to integrate populations living outside the *Pomerium* – i.e., the administrative limits of the city – and those within the city borders under a unique set of laws. Before that, the two populations lived under different sets of rules. Those outside the city limits followed customary law, which placed a strong emphasis on group identity and the sovereignty of the lineage (gens), while different laws applied to the population within the city borders. However,

> "[w]ith the settling down of gentes and other groups in Rome's hinterland, the community's population was no longer easily defined by the *curiae* and the urban area of Rome within the *Pomerium*. Populations and identities were increasingly blurred in the expanding liminal zone around the urban core of Rome. There was now an ever-increasing population outside of the *Pomerium*

> which associated itself with the community and which, it seems, desired to create a set of agreed laws and norms which applied to it as well." (Armstrong, 2020:156)

4.6. Concluding remarks

In the previous Chapter 3, we presented two stylized mechanisms for rule-making and modeled them in game-theoretical terms. These mechanisms have now been compared based on several criteria. On a concluding methodological note, it must be stressed that the comparative analysis of ambient and designed rules treated them as "ideal types" (Weber, 1949). Ideal types are simplified and purified representations of real-world objects that emphasize their core aspects and disregard circumstantial characteristics. Because of this simplification, it is natural to expect that our stylized model will not perfectly correspond to rule-creation mechanisms observed in the real world. However, this does not mean that ideal types are useless; on the contrary, the analysis of ideal types makes it possible to distill essential elements of real-world phenomena, and to pinpoint their key similarities and differences.

In Chapter 3, ambient and designed rules were initially conceptualized as strategies, primarily for demonstration purposes. This chapter extends that perspective by treating them instead as *classes* of strategies: one in which the common notion of desirable and undesirable behavior is discovered, and another in which this notion is defined upfront. This step allowed for studying the relationship between preexisting familiarity with others' private preferences and the equilibrium strategies for developing ambient rules. As such familiarity decreases, only strategies that produce more general rules remain viable.

Building on this, we proceeded to topics such as the flexibility of ambient rules versus the relative inflexibility of designed rules as well as the role of group cohesion in shaping the content of ambient rules. We also examined different ways in which both institutional types can accommodate conflicting preferences of the parties involved, and the degree to which they are susceptible to capture and error.

The chapter is theoretical; its ambition was to enhance our understanding of various social rules from a rational choice perspective. For this reason, its potential applications have only been vaguely indicated in the discussion. Nonetheless, we still believe that our theory can

shed light on historical and contemporary contexts where the two stylized institution-making methods function as alternatives, especially legal history, international law, and international relations in general.

5. Ending

In the Introduction, we promised a law and economics inquiry into "spontaneous" institutions. This promise was supposed to be delivered in two steps. First, as a general investigation into what spontaneous institutions are, and how they can be defined and understood from a law and economics perspective. This task has been carried out in Chapter 1 – the longest chapter in the thesis. Chapter 1 explained various reasons why certain institutions are considered "spontaneous" in political science, economics, law, sociology, and other fields. It also argued that, while determining what qualifies as a spontaneous institution is ultimately a matter of legal philosophy, existing approaches to this question can be combined into a useful classification. This classification can be used to delineate structurally different ideal types of institutions.

Second, the focus was shifted to a more detailed analysis of custom and design as two mechanisms for rule-making, in a setting where collective action is necessary for enforcement. This was done in Chapters 2 to 4 – first by identifying three general challenges all institutions must overcome: the incentive problem, the coordination problem, and the discovery challenge; then by examining how the discovery challenge is addressed differently in customary rules and in rules made through deliberate design. Chapter 4 specifically analyzed the consequences of these differences for the shape of rules, their stability and flexibility, the potential to accommodate asymmetric interest, and the personality of rules.

The contents of the thesis can be interpreted in several ways. In most general terms, they can be understood as an abstract exercise in formal institutional theory that utilizes the toolkit of rational choice methodology. In this sense, the thesis is an attempt to understand various aspects of the emergence and stability of institutions, especially institutions that earned the "spontaneous" label. Naturally, it cannot match the level of seminal works in the field, such as *Evolution of the Social Contract* (1996) by Brian Skyrms or *Toward a Comparative Institutional Analysis* (2001) by Masahiko Aoki. By paying special attention to incomplete information, the thesis only slightly expands the analysis beyond the fundamentals brilliantly laid down in these and other great textbooks. Still, the point stands: according to this interpretation, the purpose of the current thesis was purely theoretical. The illustrations that recur throughout the text would

be barely supplementary and could be omitted by an impatient reader who is interested in game-theory-oriented institutional analysis.

The interest in spontaneous institutions has been the primary motivation behind the thesis, and especially behind its first chapter. Accordingly, the research on spontaneous institutions is the field in which the thesis makes its primary contribution. This research is typically done from a certain axiological angle. Scholars who, from a moral standpoint, criticize the monopolistic role of contemporary states in shaping social order search for alternative solutions. On the one hand, these scholars hypothesize about a world in which law-making and order-enforcing functions of contemporary states are carried out in a competitive, non-exclusive, non-compulsory setting. For example, philosopher David Thunder has recently published a book entitled *The Polycentric Republic. A Theory of Civil Order for Free and Diverse Societies* (2025), in which a moral criticism of the present-day model of "sovereign, monocentric states" is accompanied by a suggestion to replace it (or at least attempt to replace it) with a "polycentric republic." One year prior, Thunder co-edited (with Pablo Paniagua) a book titled *Polycentric Governance and the Good Society* (2024). Some of the chapter titles in this work speak for themselves: "An Ethical Case for Bottom-Up, Polycentric Governance in a Complex Society," "The Problem of Complexity and the Emergence of Polycentric Political Order," "Self-Governance Solutions to Social Dilemmas: A Polycentric Approach," and "Polycentrism, the Rule of Law, and the Intelligibility of Human Rights Law." They all offer rationales for and proposals to replace the currently predominant legal and social "monocentrism" with a social order whose institutions are spontaneous, not crafted by monopolistic political authorities.

Similar works have been published by other philosophers as well as by economists and legal scholars. Edward Stringham compiled a large collection of papers and essays, entitled *Anarchy and the Law: The Political Economy of Choice* (2007), that present the case for the rule of law understood as an essentially "private" enterprise. Some of these papers have been cited in the thesis. Likewise, Anthony de Jasay wrote a book with the telling title *Against Politics. On government, anarchy, and order* (1997), in which he uses state-of-the-art economic theory to argue for the feasibility of the rule of law without government. Unlike many other theorists, de Jasay believes that, while "ordered anarchy" is feasible, it is also economically inefficient compared to first-best solutions. However, he simultaneously claims that the same must be said about institutions devised and imposed by governments.

It was argued that these and similar considerations often suffer from insufficient conceptual clarity. The problem of the origin of the rules (who makes the rules? where are the rules derived from, or how are they discovered?) often overlaps with the problem of enforcement (who, or what, makes agents comply with the rules? how is non-compliance punished?), which again overlaps with the problem of centralized interpretation and validation (when conflicts over interpretation and validation of rules emerge, who has the final say?). Seen in the light of this intellectual tradition, the thesis is not only an attempt to bring clarity to the aforementioned debates, but also an attempt to point out that these debates can, and should, be supplemented with state-of-the-art institutional theories (such as the coordination account of law or the theory of institutions as rules-in-equilibrium) and with the historical analysis of real-world institutions. In other words, it suggests that the problems considered within the (often speculative) theory of spontaneous institutions have often been addressed by institutional economists and law and economics scholars in other contexts.

The thesis makes another modest theoretical contribution. First, it revisits contemporary theories of institutions, placing special emphasis on the coordination aspect of institutions. Institutions are often conceptualized as coordination mechanisms: they are rules that govern human interactions in situations where multiple hypothetical solutions are conceivable. Importantly, these rules must be jointly implemented by multiple actors. From this point of view, a rule for the prisoner's dilemma that says "Always defect" is not an institution – it can be carried out by any agent in isolation. Conversely, a rule that stipulates "If the other party cooperates, cooperate" can be considered an institution, because actions of both parties are required for it to function.

As mentioned in the Introduction, this approach to institutions has recently been summarized as "rules-in-equilibrium" (Hindriks and Guala, 2015). On the one hand, institutions are "rules" because they command a specific course of action depending on particular conditions. They take the form: "In circumstances X, do Y." These conditions vary: "If the other party cooperates, cooperate" is a conditional rule, as are "If you are a man, let the lady through the door first" or "If your right-hand side is free, continue driving." These rules act as coordination devices: they prescribe a specific course of action for designated agents under specific circumstances, and therefore ensure that everyone understands what they are expected to do and what to expect from others. A given rule is distinguished from other possible coordination devices (i.e., other rules), such as "If the other party cooperates *sometimes*, cooperate," "If you are a man, go through the door before a lady does," or "If your left-hand

side is free, continue driving." However, institutions are not just any rules, but rules "in equilibrium." This means the actions and beliefs embedded in the rules must be sustainable as equilibrium behaviors among the involved parties. In other words, they must create a system of incentives such that agents' actions, driven by these rules, do not collapse or deviate from the stipulated behavior. If the other party cooperates, there must be an incentive to cooperate as well.

Because institutions are "rules," they address the coordination problem; because they are "in equilibrium," they address the incentive problem outlined in Chapter 2 of the thesis. What the thesis adds to this picture is an emphasis on how "rules" are discovered to be "in equilibrium" or not. For example, it is well established that the rule "If the other party cooperates, cooperate" can only be implemented in a two-person prisoner's dilemma under certain conditions – players need to strongly value future utility relative to present utility. But do they really value the future so highly? Or, in other words, are their private discount factors high enough? In conventional models, discount factors are assumed to be known in advance, and this question, along with similar ones, is typically assumed away. However, in reality, the emergence of different types of institutions depends on how agents discover whether their intended "rules" are "in equilibrium." The thesis illustrates this claim through a comparison between what it calls "ambient rules" and "designed rules." It suggests that ambient rules tend to be more general and equitable, while designed rules can accommodate more complexity (i.e., case-specificity) but also more vested interests.

However, the interpretation of the thesis can go beyond a mere intellectual exercise in institutional analysis. Beyond the abstract study of institutions, and especially spontaneous institutions, it is hard to overlook that two types of applications and examples were particularly present in all thesis chapters: legal anthropology (or early legal history) and customary international law. This fact certifies to the possibility of an alternative, more empirically-oriented interpretation of the thesis. It can be treated as a formal attempt to analyze early ("primitive") law, international law, and international rules-based order in general from the law and economics perspective. The thesis emphasizes the possibility of cross-fertilization between the study of legal history or legal anthropology on the one hand and international law or international cooperation on the other. While this possibility has been recognized by legal scholars long ago (e.g., Barkun, 1968; Dinstein, 1986; Campbell, 1988), the similarities between both have been only incidentally used in the law and economic analysis. More

generally, the parallels between all kinds of "spontaneous" institutions characterized in Chapter 1 seem to be too rarely exploited in the analysis of their structural counterparts.

These parallels between legal history and international relations (including the study of international law) can be demonstrated with two examples taken from contemporary scholarship on international relations and the law-and-economics approach to legal history.

The first of the examples concerns the key promise of international law and the institutionalization of the global scene: the promise of alleviating tensions between states, decreasing the fierceness of international conflict, and generally civilizing hostilities among nations. It is well known among international relations scholars that the prevalence of war has not decreased since the establishment of the United Nations and its associated institutions. Despite the hopes of enthusiasts of international institutions, their increasing involvement in international affairs has failed to bring more peace between nations (e.g., Braumoeller, 2019). However, a more modest ambition remains valid: even though there is no less war globally than was historically the norm, the conduct of hostilities can be more predictable, more civilized, and safer for non-combatants thanks to international humanitarian law.

International humanitarian law, historically known as the "laws of war" (*ius in bello*), is a branch of international law concerned with acceptable and unacceptable ways of conducting military action. As such, this body of norms is as old as warfare itself. Bederman (2001) identifies international norms governing military conflict as one of the four main branches of international law in antiquity. Likewise, Henckaerts and Doswald-Beck (2005), writing on behalf of the International Committee of the Red Cross, trace many customary humanitarian rules back to medieval or even ancient practices of warfare.

Importantly, international humanitarian law underwent a qualitative transformation throughout the 20th century. At the beginning of the century, it was still largely a system of customary rules: practices often considered valid since time immemorial and formulated in general or even vague terms. In 1898, there were only three international treaties that codified the rules of international humanitarian law: the Declaration Respecting Maritime Law of 1856, the Red Cross Convention of 1864, and St. Petersburg Convention of 1868. One hundred years later, the number had grown to 33. In other words, international humanitarian law underwent substantial and comprehensive codification. It began as a system of ambient rules, in the terminology used in this thesis, and was transformed into a system of complex relational agreements. This change was associated with features expected of relational agreements in this thesis: "new" international humanitarian law is more detailed than the "old" customary *ius in*

bello and provides extensive protections for third parties – i.e., non-belligerents. This development "presents states with much greater costs of compliance without concomitant increases in benefits" (Fazal, 2012:565).

This transformation had significant consequences for the way the laws of war are invoked and complied with within the international community. As Fazal (2012) convincingly argues, states avoid being fully subordinated to the new regime of humanitarian law by refusing to declare war on each other. This is a deliberate strategy: states choose to forgo the protections afforded by the laws of war in order to claim that they are not bound by those laws, since no formal war is being fought.

Here we see several features previously discussed in the comparison of ambient and designed rules. First, codified international humanitarian law has characteristics that customary laws of war did not: greater detail and stronger protections for non-combatants. The designing of international humanitarian law appears to have crossed a threshold beyond which states prefer to avoid it altogether. Because they cannot selectively avoid only the provisions they find problematic, they need to reject it in its full scope.

As another example of how the same structural factors can explain both international law and primitive law, consider the role of reputation in each. A key insight of modern game theory is that, absent centralized enforcement, reputation can serve as an effective enforcement mechanism. This becomes even clearer when we assume imperfect information: actors have limited memories of past events or tend to forget distant ones. In such circumstances, reputation cannot be established once and for all by a single action, nor can it be maintained solely off-equilibrium – i.e., by threatening future responses if cooperation breaks down. Instead, reputation must be actively nurtured and maintained. Recent actions need to be striking enough to leave a strong impression on others.

This general structural similarity between the international system and primitive societies has recently inspired new theories of reputation. In international relations, much has been written about the so-called "reputation for resolve": the idea that states confronting a crisis consider how willing other states have been to resort to violent means when resolving similar crises in the past. If potential rivals are known to resort to violence, there is a stronger incentive to back off. However, this incentive is limited. If states care about their reputation for resolve, they may resort to violence even when they expect others to back down. The usual motive to avoid inefficient, mutually destructive violence can thus be counterbalanced by reputational

considerations: If I do not act decisively now, will others remember me as a determined actor in the future? (see, Mercer, 2010; Crescenzi, 2018).

The same factor – and indeed the same logic – may offer a powerful explanation for the evolution of early law. Remedies in ancient law are often described as evolving through four phases: (i) unlimited revenge; (ii) limited revenge, or measure-for-measure; (iii) optional damage payments; and (iv) mandatory damage payments (Daube, 1947; Parisi, 2004). While several law-and-economics explanations exist for the emergence of the later stages, unlimited revenge poses a particular puzzle. Why would families, lineages, and clans choose to engage in costly and dangerous acts of revenge, often inflicting more harm than they originally suffered, thereby magnifying perceptions of injustice? Why would they not settle disputes through compensation and avoid the risk of spiraling feuds?

Reputational considerations are one possible explanation. In an environment of self-help, maintaining a consistent and forceful reputation for resolve is crucial for deterring predatory aggression. Moreover, reputation fades over time. If members of a family, lineage, or clan have not recently engaged in retaliatory violence, their willingness to resort to force may be doubted. To sustain a reputation for resolve, such acts may therefore be necessary. For this reason, more dramatic revenge may be preferred over milder responses: more brutal acts taken in response to relatively small transgressions are more likely to be noticed and remembered. As communities begin living in closer settlements and forming multi-faceted relationships (e.g., through frequent commercial contact), information flows improve, reducing the need for such resonant acts of revenge. At that point, damage payments, being far less wasteful from a resource perspective, can begin to take precedence over violent retaliation.

All in all, both primitive law and international rules-based order share several features, including (i) an emphasis on sovereignty of actors (state sovereignty, autonomy of family, lineage, clan, etc.), (ii) the conciliatory nature of adjudication, (iii) the central role of custom as a source of rules, with a limited role of deliberate, prospective rule-making, (iv) the locality and personality of rules, and (v) the coordinative, rather than peremptory, role of legal institutions like courts, tribunals, and rule-making bodies. Some of them have been mentioned throughout the four main chapters of the thesis. A more detailed study of this parallel, and one that fully explores the cross-fertilization between these two domains, is the author's future research goal – if given the chance.

List of references

Abbott, K. 2008. “Enriching Rational Choice Institutionalism for the Study of International Law,” 1 *University of Illinois Law Review* 5-46.

Acemoglu, D., and Jackson, M. 2014. “History, Expectations, and Leadership in the Evolution of Social Norms,” 82 *The Review of Economic Studies* 423-456.

Acemoglu, D., and Wolitzky, A. 2020. “Sustaining Cooperation: Community Enforcement versus Specialized Enforcement,” 18 *Journal of the European Economic Association* 1078-1122.

Aldashev, G., Chaara, I., Platteau, J.-P., and Wahhaj, Z. 2012. “Using the law to change the custom,” 97 *Journal of Development Economics* 182-200.

Alexander, J. 2007. *The Structural Evolution of Morality*. New York: Cambridge University Press.

Alexandrov, S. 2006. “The Compulsory Jurisdiction of the International Court of Justice: How Compulsory Is It?,” 5 *Chinese Journal of International Law* 29-38.

Allen, D., and Barzel, Y. 2011. “The Evolution of Criminal Law and Police during the Pre-modern Era,” 27 *The Journal of Law, Economics, and Organization* 540-567.

Anderson, T., and Hill, P. 1978. “An American Experiment in Anarcho-Capitalism: The Not So Wild, Wild West,” 3 *The Journal of Libertarian Studies* 9-29.

Aoki, M. 2001. *Toward a Comparative Institutional Analysis*. Cambridge, MA, and London: MIT Press.

Armstrong, J. 2020. "Beyond the *Pomerium*: Expansion and Legislative Authority in Archaic Rome," in Bell, S. and Du Plessis, P., eds. *Roman Law before the Twelve Tables, An Interdisciplinary Approach*, Edinburgh: Edinburgh University Press.

Assier-Andrieu, L. 1983. "Custom and Law in the Social Order: Some Reflections upon French Catalan Peasant Communities," 1 *Law and History Review* 86-94.

Austin, J. 1832. *The Province of Jurisprudence Determined*. London: John Murray.

Axelrod, R., and Hamilton, W. 1981. "The Evolution of Cooperation," 211 *Science* 1390-1396.

Axelrod, R. 1986. "An Evolutionary Approach to Norms," 80 *The American Political Science Review* 1095-1111.

Baert, P. 1991. "Unintended consequences: a typology and examples," 6 *International Sociology* 201-210.

Barkun, M. 1968. *Law without Sanctions. Order in Primitive Societies and the World Community*. New Haven: Yale University Press.

Barsalou, O. 2010. "The History of Reprisals Up to 1945: Some Lessons Learned and Unlearned for Contemporary International Law," 49 *Military Law and The Law of War Review* 335-371.

Basu, K. 2018. *The Republic of Beliefs. A New Approach to Law and Economics*. Princeton and Oxford: Princeton University Press.

Bederman, D. 2001. *International Law in Antiquity*. Cambridge: Cambridge University Press.

Bederman, D. 2010. *Custom as a Source of Law*. New York: Cambridge University Press.

Bellomo, M. 1995. *The Common Legal Past of Europe. 1000-1800*. Washington, D.C.: The Catholic University of America Press.

Ben-Shahar, O. 1999. "The Tentative Case against Flexibility in Commercial Law," 66 *The University of Chicago Law Review* 781-820.

Benda-Beckmann, K. 1981. "Forum Shopping and Shopping Forums: Dispute Processing in a Minangkabau Village in West Sumatra," 13 *The Journal of Legal Pluralism and Unofficial Law* 117-159.

Benson, B. 1989. "Enforcement of Private Property Rights in Primitive Societies: Law without Government," IX *The Journal of Libertarian Studies* 1-26.

Benson, B. 1990. *The Enterprise of Law: Justice without the State*. Pacific Research Institute.

Bernstein, L. 1992. "Opting out of the Legal System: Extralegal Contractual Relations in the Diamond Industry," 21 *The Journal of Legal Studies* 115-157.

Bernstein, L. 1996. "Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms," 144 *University of Pennsylvania Law Review* 1765-1821.

Bernstein, L. 2001. “Private Commercial Law in the Cotton Industry: Cooperation through Rules, Norms, and Institutions,” 99 *Michigan Law Review* 1724-1790.

Bertolini, D. 2016. “On the Spontaneous Emergence of Private Law,” 29 Canadian Journal of Law & Jurisprudence 5-36.

Bicchieri, C. 2006. *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Bicchieri, C., and Sontuoso, A. 2020. “Game-theoretic accounts of social norms: the role of normative expectations,” in C. Capra et al., eds. *Handbook of Experimental Game Theory*. Cheltenham: Edward Elgar Publishing.

Bicchieri, C. et al. 2023. “Social Norms,” in E. Zalta and U. Nodelman, eds. *The Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/social-norms/

Boyd, R., and Richerson, P. 1988. “The evolution of reciprocity in sizable groups,” 132 *Journal of Theoretical Biology* 337-356.

Boyd, R., and Richerson, P. 2009. “Culture and the evolution of human cooperation,” 364 *Phil. Trans. R. Soc. B*. 3281-3288.

Bradford, A. (2008). “Regime Theory,” in *The Max Planck Encyclopedia of Public International Law*. Oxford University Press.

Braunmoeller, B. 2019. *Only the dead. The persistence of war in the modern age*. New York: Oxford University Press.

Buchanan, J., and Tullock, G. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.

Bull, H. 1977. *The Anarchical Society. A Study of Order in World Politics*. New York: Palgrave Publishers.

Bullough, D. 1969. "Early Medieval Social Groupings: The Terminology of Kinship," 45 *Past & Present* 3-18.

Calvert, R. 1995. "Rational Actors, Equilibrium and Social Institutions," in Knight, J., and Sened, I. eds. *Explaining Social Institutions*. Ann Arbor: University of Michigan Press.

Campbell, A. 1988. "International Law and Primitive Law," 8 *Oxford Journal of Legal Studies* 169-196.

Campbell, D. and Harris, D. 1993. "Flexibility in Long-Term Contractual Relationships: The Role of Co-Operation," 20 *Journal of Law and Society* 166-191.

Cardenas, J. 2011. "Social Norms and Behavior in the Local Commons as Seen Through the Lens of Field Experiments," 48 *Environmental & Resource Economics* 451-485.

Carugati, F. et al. 2015. "Building Legal Order in Ancient Athens," 7 *Journal of Legal Analysis* 291-324.

Cheung, S. et al. 2006. "How Relational are Construction Contracts?," 132 *Journal of Professional Issues in Engineering Education and Practice*.

Coase, R. 1960. "The Problem of Social Cost," 3 *Journal of Law and Economics* 1-44.

Coleman, J. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Cooter, R. 1994. "Structural adjudication and the new law merchant: A model of decentralized law," 14 *International Review of Law and Economics* 215-231.

Cooter, R. 1996. "Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant," 144 *University of Pennsylvania Law Review* 1643-1696.

Cooter, R., and Ulen, T. 2016. *Law and Economics, 6th edition*. Boston: Addison-Wesley.

Crawford, M. 2012. "Twelve Tables," in: Hornblower, S, Spawforth, A., and Eidinow, E. *Oxford Classical Dictionary*, 4th ed. Oxford and New York: Oxford University Press.

Cronin, B. 1999. *Community under Anarchy: Transnational Identity and the Evolution of Cooperation*. Columbia University Press.

Dalhuisen, J. 2008. "Custom and Its Revival in Transnational Private Law," 18 *Duke Journal of Comparative & International Law* 339-370.

Dari-Mattiacci, G., and Parisi, F. 2006. "The Economics of Tort Law: A Precis," in J. Backhaus, ed. *The Elgar Companion to Law and Economics* (2nd ed.). Cheltenham and Northampton, MA: Edward Elgar Publishing.

Daube, D. 1947. *Studies in Biblical Law*. Cambridge: Cambridge University Press.

Diala, A. 2017. "The Concept of Living Customary Law: A Critique," 49 *Journal of Legal Pluralism and Unofficial Law* 143-165.

Dinstein, Y. 1986. “International Law as a Primitive Legal System,” 19 *New York University Journal of International Law and Politics* 1-32.

De Geest, G. 2013. “N problems require N instruments,” 35 *International Review of Law and Economics* 42-57.

Diamond, A. 1971. *Primitive Law, Past and Present*. London: Methuen & Co.

Diamond, S. 1971. “The Rule of Law Versus The Order of Custom,” 38 *Social Research*, 42-72.

Dixit, A. 2004. *Lawlessness and Economics. Alternative Modes of Governance*. Princeton: Princeton University Press.

Drew, K. 1995. “Public vs. Private Enforcement of the Law in the Early Middle Ages: Fifth to Twelfth Centuries,” 70 *Chicago-Kent Law Review* 1583-1592.

Druzin, B. 2016. “Social Norms as Substitute for Law,” 79 *Albany Law Review* 67-100.

Durkheim, E. 1984. *The Division of Labour in Society* (1893). London: The MacMillan Press.

Elias, T. 1962. *The Nature of African Customary Law* (1956). Manchester: Manchester University Press.

Ellickson, R. 1986. “Of Coase and Cattle: Dispute Resolution Among Neighbors in Shasta County,” 38 *Stanford Law Review* 623-687.

Ellickson, R. 1989. “A Hypothesis of Wealth-Maximizing Norms: Evidence from the Whaling Industry,” 5 *Journal of Law, Economics, & Organization* 83-97.

Ellickson, R. 1991. *Order Without Law: How Neighbors Settle Disputes*. Cambridge, MA: Harvard University Press.

Ellickson, R. 2001. “The Market for Social Norms,” 3 *American Law and Economics Review* 1-49.

Elster, J. 1989. “Social Norms and Economic Theory,” 3 *Journal of Economic Perspectives* 99-117.

Elster, J. 2011. “Reciprocity and Norms,” in M. Fleurbaey et al., eds. *Social Ethics and Normative Economics. Studies in Choice and Welfare*. Berlin and Heidelberg: Springer.

Epstein, R. 1992. “The Path to ‘The T. J. Hooper’: The Theory and History of Custom in the Law of Tort,” 21 *The Journal of Legal Studies* 1-38.

Epstein, R. 1998. “Customary Practices and the Law of Torts,” in Newman, P., ed.. *The New Palgrave Dictionary of Economics and the Law*. New York: Macmillan.

Evans-Pritchard, E. 1940. *The Nuer: A description of the modes of livelihood and political institutions of a Nilotic people*. Oxford: Oxford University Press.

Fazal, T. 2012. “Why States No Longer Declare War,” 21 *Security Studies* 557-593.

Fon, V., and Parisi, F. 2003. “Reciprocity-Induced Cooperation,” 159 *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft* 76-92.

Fon, V., and Parisi, F. 2006. “International Customary Law and Articulation Theories: An Economic Analysis,” 2 *BYU International Law & Management Review* 201-232.

Fon, V., and Parisi, F. 2007. “On the optimal specificity of legal rules,” 3 *Journal of Institutional Economics* 147-164.

Friedman, D. 1979. “Private Creation and Enforcement of Law: A Historical Case,” 8 *The Journal of Legal Studies* 399-415.

Friedman, D. 1989. *The Machinery of Freedom: Guide to a Radical Capitalism* (1973). Chicago: Open Court Publishing Company.

Friedman, D. 1995. “Making Sense of English Law Enforcement in the Eighteenth Century,” 2 *The University of Chicago Law School Roundtable* 475-505.

Friedman, D. et al. 2019. *Legal Systems Very Different From Ours*. Independently published (self-published).

Friedman, J. 1971. “A Non-cooperative Equilibrium for Supergames,” 38 *The Review of Economic Studies* 1-12.

Gardner, J. 2008. *Women in Roman Law and Society*. Bloomington and Indianapolis: Indiana University Press.

Gavrilets, S., and Richerson, P. J. 2017. “Collective action and the evolution of social norm internalization,” 114 *PNAS Proceedings of the National Academy of Sciences of the United States of America* 6068-6073.

Geloso, V., and Leeson, P. T. 2020. “Are Anarcho-Capitalists Insane? Medieval Icelandic Conflict Institutions in Comparative Perspective,” 130 *Revue d'économie politique* 957-974.

Ginsburg, T., and McAdams, R. 2004. “Adjudicating in Anarchy: An Expressive Theory of International Dispute Resolution,” 45 *William and Mary Law Review* 1229-1339.

Glenn, P. 1997. “The Capture, Reconstruction and Marginalization of "Custom",” 45 *American Journal of Comparative Law* 613-620.

Gluckman, M. 1965a. *Custom and Conflict in Africa*. Oxford: Basil Blackwell & Mott.

Gluckman, M. 1965b. *The Ideas in Bartose Jurisprudence*. New Haven and London: Yale University Press.

Greif, A. 2004. “Impersonal Exchange without Impartial Law: The Community Responsibility System,” 5 *Chicago Journal of International Law* 109-138.

Greif, A., and Kingston, Ch. 2011. “Institutions: Rules or Equilibria?,” in Schofield, N. and Caballero, G., eds. *Political Economy of Institutions, Democracy and Voting*. Berlin and Heidelberg: Springer Verlag.

Greif, A., and Tabellini, G. 2017. “The clan and the corporation: Sustaining cooperation in China and Europe,” 45 *Journal of Comparative Economics* 1-35.

Greuel, P. J. 1971. “The Leopard-Skin Chief: An Examination of Political Power Among the Nuer,” 73 *American Anthropologist* 1115-1120.

Green, J. 2016. *The Persistent Objector Rule in International Law*. Oxford: Oxford University Press.

Grieco, J. 1988. "Anarchy and the Limits of Cooperation: A Realist Critique of the Newest Liberal Institutionalism," 42 *International Organization* 485-507.

Griffiths, J. 1986. "What is Legal Pluralism?," 18 *The Journal of Legal Pluralism and Unofficial Law* 1-55.

Gruter, M. 1986. "Ostracism on Trial: The Limits of Individual Rights," 7 *Ethology and Sociobiology* 271-279.

Guala, F. 2016. *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton, NJ: Princeton University Press.

Gutmann, J., and Voigt, S. 2020. "Traditional law in times of the nation state: Why is it so prevalent?," 16 *Journal of Institutional Economics* 445-461.

Guzman, A. 2008. *How International Law Works. A Rational Choice Theory*. New York: Oxford University Press.

Hadfield, G., and Weingast, B. 2012. "What is Law? A Coordination Account of the Characteristics of Legal Order," 4 *The Journal of Legal Analysis* 471-514.

Hadfield, G., and Weingast, B. 2013. "Law without the State. Legal Attributes and the Coordination of Decentralized Collective Punishment," 1 *Journal of Law and Courts* 3-34.

Hadfield, G. 2017. *Rules for a Flat World. Why Humans Invented Law and How to Reinvent It for a Complex Global Economy*. New York: Oxford University Press.

Hallaq, W. 2004. *The Origins and Evolution of Islamic Law*. Cambridge: Cambridge University Press.

Hart, H. 1994. *The Concept of Law* (1961). Oxford: Clarendon Press.

Hathaway, O., and Shapiro, S. 2011. "Outcasting: Enforcement in Domestic and International Law," 121 *The Yale Law Journal* 252-349.

Hayek, F. 1945. "The Use of Knowledge in Society," 35 *The American Economic Review* 519-530.

Hayek, F. 2002. "Competition as a Discovery Procedure," 5 *The Quarterly Journal of Austrian Economics* 9-23. Originally: "Der Wettbewerb als Entdeckungsverfahren," a 1968 lecture.

Hayek, F. 2013. *Law, Legislation, and Liberty. A New Statement of the Liberal Principles of Justice and Political Economy* (1973). London and New York: Routledge.

Helfer, L., and Wuerth, I. 2016. "Customary International Law: An Instrument Choice Perspective," 37 *Michigan Journal of International Law* 563-609.

Helper, S., and Henderson, R. 2014. "Management Practices, Relational Contracts, and the Decline of General Motors," 28 *Journal of Economic Perspectives* 49-72.

Henckaerts, J.-M., and Doswald-Beck, L. 2005. *Customary International Humanitarian Law*. Cambridge: Cambridge University Press.

Hindriks, F., and Guala, F. 2015. “Institutions, Rules, and Equilibria: A Unified Theory,” 11 *Journal of Institutional Economics* 459-80.

Hodgson, G. 2006. “What Are Institutions?,” 40 *Journal of Economic Issues* 1-25.

Hodgson, G. 2015. “On Defining Institutions: Rules versus Equilibria,” 11 *Journal of Institutional Economics* 497–505.

Hoebel, E. 1967. *The Law of Primitive Man. A Study in Comparative Legal Dynamics*. Cambridge, MA: Harvard University Press.

Hohfeld, W. 1917. “Fundamental Legal Conceptions as Applied in Judicial Reasoning,” 26 *The Yale Law Journal* 710-770.

Hooker, M. 1975. *Legal Pluralism: An Introduction to Colonial and Neo-Colonial Laws*. Oxford: Clarendon Press.

Igbokwe, V. 1998. “Socio-Cultural Dimensions of Dispute Resolution: Informal Justice Processes among the Ibo-speaking Peoples of Eastern Nigeria and Their Implications for Community/Neighbouring Justice System in North America,” 10 *African Journal of International and Comparative Law* 446-471.

Inoguchi, T., and Le, L. 2020. *The Development of Global Legislative Politics. Rousseau and Locke Writ Global*. Singapore: Springer.

in‘t Veld, M. 2023. “Was There a Rule of Law in Early Modern Amsterdam? Mercantile Customary Law as a Test,” 15 *Hague Journal of the Rule of Law* 263-282.

International Court of Justice 2024. Declarations recognizing the jurisdiction of the Court as compulsory, available at https://www.icj-cij.org/declarations [accessed 03.12.2024]

Jaraba, M. 2020. "Khul' in Action: How Do Local Muslim Communities in Germany Dissolve an Islamic Religious-Only Marriage?," 40 *Journal of Muslim Minority Affairs* 26-47.

Jasay, Anthony de (1997). *Against Politics. On government, anarchy, and order*. London and New York: Routledge.

Kadens, E. 2012a. "Introduction: Lessons from the History of Custom," 48 *Texas International Law Journal* 349-355.

Kadens, E. 2012b. "The Myth of the Customary Law Merchant," 90 *Texas Law Review* 1153-1206.

Kadens, E. 2013. "Custom's Two Bodies," in Jansen, K., Geltner, G., and Lester, A., eds. *Center and Periphery: Studies on Power in the Medieval World in Honor of William Chester Jordan*. Leiden: Brill.

Kadens, E. 2019. "Convergence and the Colonization of Custom in Pre-modern Europe," in Moreteau, O., and Modeer, K. eds. *Comparative Legal History* 167. Cheltenham: Edward Elgar Publishing.

Kaplow, L. 1992. "Rules versus Standards: An Economic Analysis," 42 *Duke Law Journal* 557-629.

Kaplow, L. 1995. "A Model of the Optimal Complexity of Legal Rules," 11 *Journal of Law, Economics, & Organization* 150-163.

Katz, A. 1995. “Taking Private Ordering Seriously,” 144 *University of Pennsylvania Law Review* 1745-1763.

Keohane, R. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, NJ: Princeton University Press.

Kim, M. 2007. “Law and Custom under the Chosŏn Dynasty and Colonial Korea: A Comparative Perspective,” 66 *The Journal of Asian Studies* 1067-1097.

Kim, M. 2009. “Customary Law and Colonial Jurisprudence in Korea,” 57 *The American Journal of Comparative Law* 205-248.

Kim, M. 2021. *Custom, Law, and Monarchy: A Legal History of Early Modern France*. Oxford: Oxford University Press.

Kinsella, S. 1995. “Legislation and the Discovery of Law in a Free Society,” 11 *Journal of Libertarian Studies* 132-181.

Knight, J. 1992. *Institutions and Social Conflict*. Cambridge: Cambridge University Press.

Koren, G. 1992. “Two-Person Repeated Games Where Players Know Their Own Payoffs,” Courant Institute of Mathematical Sciences.

Kornhauser, L. 2004. “Governance Structures, Legal Systems, and the Concept of Law,” 79 *Chicago-Kent Law Review* 355-381.

Kornhauser, L. 2015. “Doing Without the Concept of Law,” NYU School of Law, Public Law Research Paper No. 15-33, Available at SSRN: https://ssrn.com/abstract=2640605 or http://dx.doi.org/10.2139/ssrn.2640605

Kornhauser, L. 2022. “Law as an achievement of governance,” 47 *Journal of Legal Philosophy* 1-23.

Kostritsky, J. 2006. “Judicial Incorporation of Trade Usages: A Functional Solution to the Opportunism Problem,” 39 *Connecticut Law Review* 451-529.

Koyama, M. 2014. “The law & economics of private prosecutions in industrial revolution England,” 159 *Public Choice* 277-298.

Leeson, P. 2008. “Social Distance and Self-Enforcing Exchange,” 37 *The Journal of Legal Studies* 161-188.

Leeson, P. 2009. “The Laws of Lawlessness,” 38 *The Journal of Legal Studies* 471-503.

Lefkowitz, D. 2017. “What makes a social order primitive? In defense of Hart's take on international law,” 23 *Legal Theory* 258-282.

Leoni, B. 1993. *Freedom and the Law*. Los Angeles: Nash Publishing.

Lesaffer, R. 2007. “Siege warfare in the early Modern Age: A study on the customary laws of war,” in Perreau-Saussine A., and Murphy J., eds. *The Nature of Customary Law. Legal, Historical, and Philosophical Perspectives*. New York: Cambridge University Press.

Lintott, A. 1968. *Violence in Republican Rome*. Oxford: Oxford University Press.

Lim, C., and Elias, O. 1997. “The Role of Treaties in the Contemporary International Legal Order,” 66 *Nordic Journal of International Law* 1-21.

Luban, D. 2020. “What Is Spontaneous Order?,” 114 *American Political Science Review* 68-80.

MacCormack, G. 1973. “Revenge and Compensation in Early Law,” 21 *The American Journal of Comparative Law* 69-85.

Mackie, G. 1996. “Ending Footbinding and Infibulation: A Convention Account,” 61 *American Sociological Review* 999-1017.

Macneil, I. 1980. *The New Social Contract: An Inquiry into Modern Contractual Relations*. New Haven: Yale University Press.

Mahoney, P., and Sanchirico, C. 2001. “Competing Norms and Social Evolution: Is the Fittest Norm Efficient?,” 149 *University of Pennsylvania Law Review* 2027-2062.

Mahoney, P., and Sanchirico, C. 2003. “Norms, Repeated Games, and the Role of Law,” 91 *California Law Review* 1281-1329.

Mahoney, P., and Sanchirico, C. 2005. “General and Specific Legal Rules,” 161 *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft* 329-346.

Maine, H. 1883. *Dissertations on Early Law and Custom. Chiefly Selected from Lectures Delivered at Oxford.* London: John Murray.

Malinowski, B. 2017. *Crime and Custom in Savage Society* (1926). Oxon and New York: Routledge.

Marmor, A. 2001. "Legal Conventionalism," in Coleman, J., ed. *Hart's Postscript: Essays on the Postscript to "the Concept of Law."* Oxford: Oxford University Press.

Masferrer, A. 2019. "The decline and displacement of custom in early modern Spain," 84 *Tijdschrift voor Rechtsgeschiedenis / Revue d'histoire du droit / The Legal History Review* 427-472.

McAdams, R. 2000. "Focal Point Theory of Expressive Law," 86 *Virginia Law Review* 1649-1730.

McAdams, R. 2009. "Beyond the prisoners' dilemma: Coordination, game theory, and law," 82 *Southern California Law Review* 209-258.

McAdams, R., and Rasmusen, E. 2007. "Norms and the Law," in Polinsky, M., and Shavell, S., eds. *The Handbook of Law and Economics*. Elsevier Science. Vol. 2.

Mearsheimer, J. 2001. *The Tragedy of Great Power Politics*. New York: W. W. Norton & Company.

Mercer, J. 2010. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.

Merry, S. 1984. "Anthropology and the Study of Alternative Dispute Resolution," 34 *Journal of Legal Education* 277-283.

Milgrom, P., North, D., and Weingast, B. 1990. “The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs,” 2 *Economics & Politics* 1-23.

Miller, W. 1996. *Bloodtaking and Peacemaking. Feud, Law, and Society in Saga Iceland*. Chicago: University of Chicago Press.

Moore, S. 1986. *Social Facts and Fabrications. “Customary” Law on Kilimanjaro 1880–1980*. Cambridge: Cambridge University Press.

Morsky, B. and Akçay, B. 2019. “Evolution of social norms and correlated equilibria,” 116 *Proceedings of the National Academy of Sciences* 8834-8839.

Murray, A. 1983. *Germanic Kinship Structure. Studies in Law and Society in Antiquity and the Early Middle Ages*. Toronto: Pontifical Institute of Mediaeval Studies.

Ndulo, M. 2011. “African Customary Law, Customs, and Women’s Rights,” 18 *Indiana Journal of Global Legal Studies* 87-120.

Nguyen, N. 2006. “Roman Rape: An Overview of Roman Rape Laws from the Republican Period to Justinian's Reign,” 13 *Michigan Journal of Gender & Law* 75-112.

Nicholas, B. 2008. *An Introduction to Roman Law*. New York: Oxford University Press.

Norman, G., and Trachtman, J. P. 2005. “The Customary International Law Game,” 99 *The American Journal of International Law* 541-580.

North, D. 1991. “Institutions,” 5 *Journal of Economic Perspectives* 97-112.

Noyes, J. 1989. “Compulsory Third-Party Adjudication and the 1982 United Nations Convention on the Law of the Sea,” 4 *Connecticut Journal of International Law* 675-696.

Nye, H. 2022. “Does Law ‘Exist’? Eliminativism in Legal Philosophy,” 15 *Washington University Jurisprudence Review* 29-78.

Ogus, A. 2002. “The Economic Basis of Legal Culture: Networks and Monopolization,” 22 *Oxford Journal of Legal Studies* 419-434.

Okada, I. et al. 2018. “A solution for private assessment in indirect reciprocity using solitary observation,” 455 *Journal of Theoretical Biology* 7-15.

Okada, I. 2020. “A Review of Theoretical Studies on Indirect Reciprocity,” 11 *Games* 27.

Ostrom, E. 2005. *Understanding Institutional Diversity*. Princeton: Princeton University Press.

Parisi, F. 1994. “Alterum Non Laedere: An Intellectual History of Civil Liability,” 39 *The American Journal of Jurisprudence* 317-351.

Parisi, F. 1995. “Toward a theory of spontaneous law,” 6 *Constitutional Political Economy* 211-231.

Parisi, F. 1998. “Customary Law,” in Newman, P., ed. *The New Palgrave Dictionary of Economics and the Law*. New York: Macmillan.

Parisi, F. 2000. “The Formation of Customary Law,” paper presented at the 96th Annual Conference of the American Political Science Association.

Parisi, F. 2001. “Sources of Law and the Institutional Design of Lawmaking,” 19 *Journal of Public Finance and Public Choice*, 95-122.

Parisi, F. 2004. “The Genesis of Liability in Ancient Law,” 3 *American Law and Economics Review* 82-124.

Parisi, F., and Dari-Mattiacci, G. 2004. “The rise and fall of communal liability in ancient law,” 24 *International Review of Law and Economics* 489-505.

Parisi, F. Pi, D., Luppi, B., and Fargnoli, I. 2020. “Deterrence of Wrongdoing in Ancient Law,” in Dari-Mattiacci, G., and Kehoe, D. P., eds. *Roman Law and Economics. Volume II: Exchange, Ownership, and Disputes*. Oxford: Oxford University Press.

Perelló, C. 2020. “The Twelve Tables and the *leges regiae*: A Problem of Validity,” in Bell, S., and Du Plessis, P., eds. *Roman Law before the Twelve Tables, An Interdisciplinary Approach*. Edinburgh: Edinburgh University Press.

Peters, P. G. 2000. “The Quiet Demise of Deference to Custom: Malpractice Law at the Millennium,” 57 *Washington and Lee Law Review* 163-205.

Petersen, N. 2017. “The International Court of Justice and the Judicial Politics of Identifying Customary International Law,” 28 *European Journal of International Law* 357-385.

Pęski, M. 2014. “Repeated games with incomplete information and discounting,” 9 *Theoretical Economics* 651-694.

Pershits, A. 1977. “The Primitive Norm and Its Evolution,” 18 *Current Anthropology* 409-417.

Picker, R. 1994. “An Introduction to Game Theory and the Law,” Coase-Sandor Institute for Law & Economics Working Paper No. 22.

Platteau, J.-P. 2008. “The Causes of Institutional Inefficiency: A Development Perspective,” In: Brousseau, E., and Glachant, J.-M. eds. *New Institutional Economics. A Guidebook*, Cambridge: Cambridge University Press.

Posner, E. 1998. “Symbols, Signals, and Social Norms in Politics and the Law,” 27 *The Journal of Legal Studies* 765-797.

Posner, E. 2000. *Law and Social Norms*. Cambridge, MA and London: Harvard University Press.

Posner, R. 1980. “A Theory of Primitive Society, with Special Reference to Law,” 23 *The Journal of Law & Economics* 1-53.

Posner, R., and Goldsmith, J. 1999. “A Theory of Customary International Law,” 66 *The University of Chicago Law Review* 1113-1177.

Pospíšil, L. 1958. *Kapauku Papuans and their law*. New Haven, Conn.: Published for the Department of Anthropology, Yale University.

Pospíšil, L. 1971. *Anthropology of Law. A Comparative Theory*. New York: Harper & Row.

Postema, G. 1982. “Coordination and Convention at the Foundations of Law,” 11 *The Journal of Legal Studies* 165-203.

Powell, B., and Stringham, E. 2009. “Public Choice and the Economic Analysis of Anarchy: A Survey,” 140 *Public Choice* 503-538.

Rasmusen, E., and Hirshleifer, D. 1989. “Cooperation in a Repeated Prisoners’ Dilemma with Ostracism,” 12 *Journal of Economic Behavior and Organization* 87-106.

Raz, J. 1994. *Ethics in the Public Domain. Essays in the Morality of Law and Politics*. Oxford: Oxford University Press.

Renner, M. 2021. “Private Ordering,” in Grundmann, S., Micklitz, H.-W., and Renner, M. eds. *New Private Law Theory. A Pluralist Approach*. Cambridge: Cambridge University Press.

Roberts, S. 2005. “After Government? On Representing Law Without the State,” 68 *Modern Law Review* 1-24.

Robinson, P. 1980. “A Brief History of Distinctions in Criminal Culpability,” 31 *Hastings Law Journal* 815-853.

Roland, G. 2004a. “Fast-moving and slow-moving institutions,” 2 *CESifo DICE Report* 16-21.

Roland, G. 2004b. “Understanding institutional change: Fast-moving and slow-moving institutions,” 38 *Studies in Comparative International Development* 109-131.

Rossi, G., and Spagano, S. 2018. “From Custom to Law, An Economic Rationale behind the Black Lettering,” 52 *Journal of Economic Issues* 1109-1124.

Ruhl, J., and Ruhl Jr. H. 1997. “The Arrow of the Law in Modern Administrative States: Using Complexity Theory to Reveal the Diminishing Returns and Increasing Risks the Burgeoning of Law Poses to Society,” 30 *U.C. Davis Law Review* 405-482.

Sagy, T. 2011. “What's So Private about Private Ordering?,” 45 *Law & Society Review* 923-954.

Savaş, A. 2022. “The Process of Transforming Strict Liability into Liability for Fault in Roman Law, and the Effect This Transformation Has Had on Modern Law,” 80 *Istanbul Law Review* 537-582.

Schauer, F. 2015. *The Force of Law*. Cambridge, MA: Harvard University Press.

Schelling, T. 1980. *The Strategy of Conflict* (1960). Cambridge, MA: Harvard University Press.

Schiller, A. 1938. “Custom in Classical Roman Law,” 24 *Virginia Law Review* 268-282.

Schuck, P. 1992. “Legal Complexity: Some Causes, Consequences, and Cures,” 42 *Duke Law Journal* 1-52.

Sen, A. 1977. “Rational Fools: A Critique of the Behavioral Foundations of Economic Theory,” 6 *Philosophy & Public Affairs* 317-344.

Shaw, M. 2017. *International Law. Eighth edition*. Cambridge: Cambridge University Press.

Skyrms, B. 1996. *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Smith, G. 1922. “Early Greek Codes,” 17 *Classical Philology* 187-201.

Smith, H. 2009. “Community and Custom in Property,” 10 *Theoretical Inquiries in Law* 5-41.

Snyder, F. G. 1981. “Colonialism and Legal Form: The Creation of “Customary Law” in Senegal,” 13 *The Journal of Legal Pluralism and Unofficial Law* 49-90.

Stein-Wilkeshuis, M. 1986. “Laws in medieval Iceland,” 12 *Journal of Medieval History* 37-53.

Streeck, W., and Thelen, K. 2005. “Introduction: institutional change in advanced political economies,” in Streeck, W., and Thelen, L. eds. *Beyond continuity: institutional change in advanced political economies*. Oxford: Oxford University Press.

Stringham, E. 2015. *Private Governance: Creating Order in Economic and Social Life*. New York: Oxford University Press.

Sugden, R. 1986. *The Economics of Rights, Co-operation and Welfare*. New York: Palgrave Macmillan.

Sugden, R. 1989. “Spontaneous Order,” 3 *The Journal of Economic Perspectives* 85-97.

Sunstein, C., Jolls, C, and Thaler, R. 1998. “A Behavioral Approach to Law and Economics,” 50 *Stanford Law Review* 1471-1550.

Taylor, M. 1982. *Community, Anarchy, and Liberty*. Cambridge: Cambridge University Press.

Taylor, M. 1987. *The Possibility of Cooperation*. Cambridge: Cambridge University Press.

Tesler, L. 1980. “A Theory of Self-Enforcing Agreements,” 53 *The Journal of Business* 27-44.

The British Home Office – Siddiqui et al., 2018. “The independent review into the application of sharia law in England and Wales,” report presented to Parliament by the Secretary of State for the Home Department.

Thompson, E. 2015. *Customs in Common. Studies in Traditional Popular Culture* (1992). New York: The New Press.

Thunder, D. 2025. *The Polycentric Republic. A Theory of Civil Order for Free and Diverse Societies*. Milton and New York: Routledge.

Thunder, D., and Paniagua, P. (eds.) 2024. *Polycentric Governance and the Good Society. A Normative and Philosophical Investigation*. London: Lexington Books.

Ullmann-Margalit, E. 1977. *The Emergence of Norms*. New York: Oxford University Press.

Unger, R. 1976. *Law in Modern Society. Toward a Criticism of Social Theory*. New York: The Free Press.

Van Caenegem, R. 1988. *An Historical Introduction to Private Law*. Cambridge: Cambridge University Press.

Voigt, S. 2019. *Institutional Economics: An Introduction*. Cambridge: Cambridge University Press.

Voigt S. 2024a. “Determinants of social norms I – the role of geography,” 20 *Journal of Institutional Economics* E5.

Voigt, S. 2024b. “Determinants of social norms II – religion and family as mediators,” 20 *Journal of Institutional Economics* E10.

Volckart, O., Mangels, A. 1999. “Are the Roots of the Modern Lex Mercatoria Really Medieval?,” 65 *Southern Economic Journal* 427-450.

Wallis, J. 2011. “Institutions, organizations, impersonality, and interests: The dynamics of institutions,” 79 *Journal of Economic Behavior & Organization* 48-64.

Waltz, K. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley Publishing Company.

Weber, M. 1949. “’Objectivity’ in social science and social policy,” in Weber, M., ed. *Essays in the Methodology of the Social Sciences*. New York: The Free Press.

Weiner, M. 2013. *The Rule of the Clan: What an Ancient Form of Social Organization Reveals About the Future of Individual Freedom*. New York: Farrar, Straus, & Giroux.

Weiss, U., and Agassi, J. 2020. “Game Theory for International Accords,” 16 *South Carolina Journal of International Law and Business* 1-23.

Wendt, A. 1999. *Social Theory of International Politics*. Cambridge: Cambridge University Press.

Westbrook, R. 1988. “The Nature and Origins of the Twelve Tables,” 105 *Zeitschrift der Savigny-Stiftung für Rechtsgeschichte: Romanistische Abteilung* 74-121.

Williamson, O. 2002. “The Theory of the Firm as Governance Structure: From Choice to Contract,” 16 *Journal of Economic Perspectives* 171-195.

Williamson, O. 2005. “The Economics of Governance,” 95 *The American Economic Review*, 1-18.

Worby, E. 1997. “Eleven Guilty Men from Goredema: Parallel Justice and the Moralities of Local Administration in Northwestern Zimbabwe,” 39 *Anthropologica* 71-77.

Wormland, J. 1982. “The Blood Feud in Early Modern Scotland,” in: Bossy, J., ed. *Disputes and Settlements. Law and Human Relations in the West*. Cambridge: Cambridge University Press.

Wright, R. 2000. “The Illusion of Simplicity: An Explanation of Why the Law Can’t Just Be Less Complex,” 27 *Florida State University Law Review* 715-744.

Young, P. 2001. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions* (1998). Princeton, NJ: Princeton University Press.

Young, P. 2008. “Social norms,” in Durlauf, S., and Blume, L., eds. *The New Palgrave Dictionary of Economics. Second Edition.* Vol. 7. New York: Palgrave Macmillan.

Young, P. 2015. “The Evolution of Social Norms,” 7 *Annual Review of Economics* 359-387.

Zenker, O., and Hoene, M. 2018. “Processing the paradox. When the state has to deal with customary law,” in Zenker, O., and Hoene, M., eds. *The State and the Paradox of Customary Law in Africa*. Noe York: Routledge.

Summary

There is a field of research dedicated to social order emerging without the apparatus typical of present-day nation states: without legislatures with undisputed competence to make laws, without enforcement agencies capable of coercion into compliance, or even without courts having the final say in disputes. In fact, there is more than one field of this type, or at least this field has been explored from multiple angles across various academic disciplines. Libertarian social theorists study this field, alongside political theorists, legal anthropologists, institutional economists, and international relations scholars – each group from a different angle and with different objectives in mind. Perhaps for this reason, the boundaries of this field can sometimes be blurred. Apparently, there is more than one way to define terms such as "anarchy," "lawlessness," "spontaneous law," "custom," and "private ordering." Therefore, the first part of this thesis outlines the basic building blocks of the field of "spontaneous" social orders. The result is a classification of institutions studied within this field into several ideal types, which are similar in some respects and different in others. It is also shown that, in the rational choice framework, these similarities and differences are "structural": they can be represented with the corresponding designs of games in the game-theoretical sense. The second, longer part of the thesis takes a deeper dive into a specific area of the field in question, examining the different outcomes of two rule-making regimes: custom and design. The focus is on institutions with decentralized, or collective, enforcement, which aligns closely with the realities of societies governed by "early" or "primitive" law, as well as the structural anarchy of the international scene. The novelty of this part lies primarily in incorporating incomplete information into the analysis. It begins with a basic observation: institutions are public knowledge, but also start with much of the relevant information being private knowledge of the parties involved. Against this backdrop of incomplete information, the second part of the thesis analyzes, on the one hand, customary rules – those that emerge from the interactions of agents – and, on the other hand, "designed" rules. The thesis suggests that different methods of institution-making result in varying degrees of specificity, susceptibility to manipulation, and flexibility or adaptability to the broader environment. These conclusions are illustrated with examples from legal anthropology, legal history, and international law.

Samenvating

Er is een onderzoeksgebied gewijd aan sociale orde die ontstaat zonder het apparaat dat kenmerkend is voor de huidige natiestaten: zonder wetgevende macht met onbetwiste bevoegdheid om wetten te maken, zonder handhavingsinstanties die naleving kunnen afdwingen, of zelfs zonder rechtbanken die het laatste woord hebben in geschillen. In feite is er meer dan één gebied van dit type, of in ieder geval is dit gebied onderzocht vanuit meerdere invalshoeken in verschillende academische disciplines. Libertarische sociale theoretici bestuderen dit gebied, naast politieke theoretici, juridische antropologen, institutionele economen en wetenschappers op het gebied van internationale betrekkingen - elke groep vanuit een andere invalshoek en met andere doelen voor ogen. Misschien dat daarom de grenzen van dit veld soms vaag zijn. Blijkbaar is er meer dan één manier om termen als “anarchie”, “wetteloosheid”, “spontaan recht”, “gewoonte” en “private ordening” te definiëren. Daarom schetst het eerste deel van dit proefschrift de basisbouwstenen van het veld van “spontane” sociale ordes. Het resultaat is een classificatie van instituties die binnen dit veld zijn bestudeerd in verschillende ideaaltypen, die in sommige opzichten op elkaar lijken en in andere verschillen. Er wordt ook aangetoond dat, in het kader van rationele keuze, deze gelijkenissen en verschillen “structureel” zijn: ze kunnen worden voorgesteld met de overeenkomstige spelontwerpen in speltheoretische zin. Het tweede, langere deel van het proefschrift neemt een diepere duik in een specifiek gebied van het veld in kwestie, waarbij de verschillende uitkomsten van twee regelgevende regimes worden onderzocht: gewoonte en ontwerp. De nadruk ligt op instellingen met gedecentraliseerde, of collectieve, handhaving, wat nauw aansluit bij de realiteit van samenlevingen die worden geregeerd door “vroeg” of “primitief” recht, evenals de structurele anarchie van het internationale toneel. De nieuwigheid van dit deel ligt vooral in het opnemen van onvolledige informatie in de analyse. Het begint met een fundamentele constatering: instellingen zijn openbare kennis, maar beginnen ook met veel van de relevante informatie die privékennis is van de betrokken partijen. Tegen deze achtergrond van onvolledige informatie analyseert het tweede deel van het proefschrift enerzijds gebruikelijke regels - regels die voortkomen uit de interacties tussen agenten - en anderzijds “ontworpen” regels. De dissertatie suggereert dat verschillende methoden van institutionele vorming resulteren in verschillende gradaties van specificiteit, gevoeligheid voor manipulatie en flexibiliteit of aanpassingsvermogen aan de bredere omgeving. Deze conclusies worden geïllustreerd met voorbeelden uit de rechtsantropologie, de rechtsgeschiedenis en het internationaal recht.

Appendix – doctoral activities and CV

Courses		
Title	*Venue*	*Date*
Introductory Statistics	University of Bologna	Oct 2019
Introduction to Law and Economics	University of Bologna	Oct – Dec 2019
Law and Economic Development	University of Bologna	Nov – Dec 2019
Modeling Private Law	University of Bologna	Dec 2019 – Jan 2020
Introduction to EU Competition Law	University of Bologna	Jan 2020 – Mar 2020
Research Design	On-line/ Erasmus University Rotterdam	Jan 2020 – Jul 2020
Empirical Legal Studies	On-line/ University of Hamburg	Apr 2020 – May 2020
Summer School (4 courses)	On-line/ University of Hamburg	May 2020 – June 2020
Academic Writing	On-line/ Erasmus University Rotterdam	Oct – Dec 2020
Advanced Empirical Methods Research Design	On-line/ Erasmus University Rotterdam	Oct 2020
EGSL course – Communicate your research	On-line/ Erasmus University Rotterdam	Oct 2020
Advanced Empirical Methods Research Design – Applied	On-line/ Erasmus University Rotterdam	Nov 2020
Computational Methods	On-line/ Erasmus University Rotterdam	Dec 2020
Academic Integrity and Responsible Research	On-line/ Erasmus University Rotterdam	July 2024

Conferences, seminars, and other activities (presentations in red)		
Description	*Venue*	*Date*
EDLE 3rd year seminar	University of Bologna	Nov 2019
8th International Conference "Austrian Economics in the 21st Century": presentation "Monetary Equilibrium and the Capital Structure of Production"	Oesterreichische Nationalbank Wien	Nov 2019
EDLE seminars (incl. two chapter presentations)	On-line/ Erasmus University Rotterdam	Sep 2020 – Mar 2021
"The Future of Law and Economics" conference	On-line/ University of Maastricht	Mar 2021
(selected) Lectures in L&E at the Insitute of L&E	University of Hamburg	Every semester starting from Mar 2021 until Jun 2024
(selected) Jour fixe meetings at the Institute of L&E (incl. two presentations)	University of Hamburg	Every semester starting from Mar 2021 until Jun 2024
KASE (Club of Austrian School in Economics in Poland) on-line seminars (incl. giving three talks about classical papers)	On-line (under the auspices of the Ludwig von Mises Institute for Economic Education, Poland)	Since Mar 2021, once in six weeks on average, irregular participation
EDLE 3rd year seminar: chapter presentation (work-in-progress version)	University of Bologna	Nov 2021
"The Future of Law and Economics" conference: presentation "Spontaneous norms in law and economics: a sketch typology"	University of Maastricht	Apr 2022
SIDE-ISLE Annual Conference: presentation "Spontaneous norms in law and economics: a sketch typology"	Università LUMSA, Palermo	Dec 2022
SIDE-ISLE Annual Conference: presentation "Spontaneous norms among heterogeneous agents: a rational-choice model"	University of Brescia	Dec 2023
"The Social Ontology of Sovereignty: Timeless and Timely Perspectives" – LVII Annual Philosophical Meetings: presentation "Popular sovereignty and	Universidad de Navarra, Pamplona	Mar 2024

Hohfeldian analysis" (with Dr. Tanja Porcnik)		
SIDE-ISLE Annual Conference: presentation "Customary rules versus designed rules"	Sapienza Università, Rome	Dec 2024
Polish Association of Law and Economics 2025 Jubilee Conference: presentation (proposal accepted) "Rational-choice sisters: international law and primitive law"	University of Warsaw	To be held in May 2025

Publications and working papers:

- “Nacjonalizm w perspektywie ekonomii politycznej”, 2 *Eryda*, 2015, pp. 31-52. ISSN 2451-2028
- “Anticommons. A very brief introduction with selected applications.” In: *Mengerian Economics*, edited by Jasiński, Ł., Sielska, A., & Turowski, K. Edward Elgar Publishing, 2023, pp. 168-181. https://doi.org/10.4337/9781035302895
- “Spontaneous institutions: a typology." Eur J Law Econ 60, 173–208 (2025). https://doi.org/10.1007/s10657-025-09851-1
- “Customary rules with heterogeneous agents and incomplete information” (Considered for publication in the “Review of Law and Economics”, currently before minor revision. Previous version as ILE Working Paper, No. 78)

Curriculum vitae

Karol Zdybel
Freta 11/5, 00-227 Warsaw, Poland
karolzdybel91@gmail.com zdybel@law.eur.nl
+48 500 088 469

Aspiring law and economics scholar with a strong mathematical/ game-theorical background. Interested in legal theory, early legal history, and international law/ international relations. Management consultant outside of academia.

Professional experience

- 2025-ongoing

Senior consultant/ Project manager at Cognivate, a boutique management consultancy focused on the financial sector

- 2021-2025

Interim senior consultant at Roland Berger, a global consulting firm (on-and-off engagements during the Ph.D. studies)

- 2017-2019

Consultant at Roland Berger (full-time)

- 2015-2017

Consultant at PwC, a global assurance and advisory company

Education

- 2019-ongoing

Universität Hamburg, Erasmus University Rotterdam, Università di Bologna – EDLE: European Doctorate in Law and Economics

Ph.D. thesis: “The Thin Boundary Between Custom and Law: A Law and Economics Approach.” Supervision: Prof. Stefan Voigt (Hamburg), Prof. Michael Faure (Rotterdam)

- 2010-2016

University of Warsaw (2010-2016)

M.A. (Polish: Magister) in Economics. Thesis supervision: Prof. Katarzyna Metelska-Szaniawska

B.A. (Polish: Licencjat) in Philosophy

Awards

- Brenno Galli Award (2022) – awarded by the Italian Society of Law and Economics (SIDE-ISLE), for "the most promising young scholar's paper presented at the SIDE-ISLE annual conference", for an early version of the paper "Spontaneous institutions: a typology"
- International Vernon Smith Prize for the Advancement of Austrian Economics, 1st prize (2016) – awarded by ECAEF foundation in Vaduz, Liechtenstein, in a global essay contest open to individuals up to 30 years of age

Research skills and interests

- Game theory
- Mathematical modeling
- Philosophy of law
- Early legal history
- International relations theory

Karol Zdybel – Propositions for the Ph.D. defense

- The "rules-in-equilibrium" notion of institutions theoretically reconciles the two hitherto opposite definitions of institutions: "rules governing human behavior" and "equilibria in human interactions"
- There are three main ways to define "spontaneous" institutions in the law and economics literature – (i) customary institutions, (ii) institutions sustained through decentralized enforcement, and (iii) institutions sustained without third-party coordination provided ex post – which are mostly logically independent from one another
- Including incomplete information as one of the central assumptions of institutional analysis can substantially enrich that analysis
- Customary institutions, especially when sustained without third-party coordination provided ex post, are simpler (i.e., less case-specific) than institutions originating through conscious design, which can be seen both in archaic law and international law
- Institutions that are deliberately designed, even if sustained through decentralized enforcement, are less likely to adjust gradually but more likely to fall abruptly than customary institutions
- Law and economics and utilitarianism do not need to go hand in hand; it is possible to do law and economics without striving for economic efficiency as a policy objective or moral imperative
- It would be conducive for research if social scientists in all fields relied more on economic models (in particular: game theory models) while formulating their theories, especially in international relations, sociology, and political sciences
- Realism and liberalism, two main orientations in the study of international relations, can be understood as two special cases of a unified rational-choice model of the international system
- With a few exceptions, nationalism, with all its flaws, is essentially a theory of peace rather than an inherently aggressive, war-mongering ideology it is often portrayed as today
- The ambition of policymaking, pervasive among law-and-economics scholars, diminishes the quality of research and leads the field astray. A more proper ambition in science would be to understand the world, not to change it
- (not to be defended) The AI revolution will likely be a social failure, even to a greater extent than the smartphone revolution has proven to be in hindsight