Mathematical Models of the Evolution of Institutions

Sergey Gavrilets [1] *and Thomas E. Currie* [2]

Abstract

We review existing mathematical models describing the evolution of social institutions. We discuss the role that formal models can play in understanding institutional change. We highlight the main features of such models by focusing on three related questions: What is the function of an institution? What social processes are involved in establishing institutional rules? How does institutional change occur? We illustrate these points by discussing some existing examples of how institutions have been modeled. Inspired by some major patterns of human history, and reflecting the themes of this book, we separately consider institutions in small-scale societies, the historical emergence of complex sociopolitical institutions, and political and economic institutions in more recent societies. Explicit models of institutional evolution have been relatively few in number but offer the potential to provide important insights into a key aspect of human social organization. We highlight some of the important issues that have not received much attention in current models, such as the coevolution of institutions with other aspects of culture.

1. Introduction

Human societies are fundamentally structured by socially generated and culturally inherited rules that set expectations about how individuals are supposed to behave in different situations and establish the consequences of not following those rules (Currie et al. 2016; North 1990; Ostrom 1990). We refer to these rules, and the

[1] Department of Ecology and Evolutionary Biology, Department of Mathematics, Center for the Dynamics of Social Complexity, University of Tennessee, Knoxville, TN 37996 USA

[2] Human Behaviour and Cultural Evolution Group, Centre for Ecology and Conservation, Department of Biosciences, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9 FE, UK

This chapter is drawn from the following book:

Richerson, Peter J., Jenna Bednar, Thomas E. Currie, Sergey Gavrilets, and John Joseph Wallis, eds, *Institutional Dynamics and Organizational Complexity* : *How Social Rules Have Shaped the Evolution of Human Societies Throughout Human History*. Open Access Book, Cultural Evolution Society, 2023. institutionaldynamicsbook.culturalevolutionsociety.org

social interactions and processes that establish and shape them, as "institutions." Describing the great diversity of institutions across time and space and understanding the processes that have generated this diversity is of interest to a number of disciplines and forms the focus of this book. In this chapter, we discuss the role that formal mathematical models of institutions and institutional change can play in this endeavor. We begin by discussing why modeling is important for understanding institutional change. We then move on to describing the main issues involved in modeling institutions and some of the approaches that have been taken to date. We next illustrate existing models that have been applied to understand some key features of institutional evolution in human history. We finish by discussing some important issues that have not received much attention in current models, such as the coevolution of institutions with other aspects of culture, and highlight areas for future research.

2. The Role of Formal Models in Understanding Institutional Change

Developing formal models of institutions and institutional evolution is important for a variety of reasons. First of all, as discussed elsewhere in this volume (see Bednar and Wallis), the term institutions can be used to refer to a variety of phenomena, just as other terms like evolution can often mean different things to researchers working in different disciplinary traditions (Currie et al. 2016). In developing mathematically based descriptions of the things we are interested in, we are often forced to be more explicit about some assumptions. Secondly, institutions are complicated things involving interactions between different actors and different situations. When faced with such complexity, our best bet is to try to break things into constituent parts and develop relatively simple models that help explain certain situations or certain parts of the bigger problem. Simple models are not designed to capture all the processes that are important in the real world and often involve high levels of abstraction. However, work in complexity theory illustrates that even relatively simple mechanisms involving feedback loops and non-linear dynamics can potentially produce complex outcomes or emergent phenomena. Furthermore, as we build our understanding of simple models and establish firmer foundations, we can begin to combine models and add greater complexity or realism.

The complexity inherent in understanding institutions and institutional change has not prevented researchers in different disciplines from developing ideas to explain the things they are interested in. Formal models can help us assess such hypotheses in two ways. Firstly, they allow us to assess the logic of our ideas and help establish whether they make sense "in theory," i.e., do they represent possible explanations? This is particularly important when hypotheses are developed informally (i.e., "verbal models") as there is more ambiguity and a greater chance that logical slips might be made. Additionally, developing formal models can help reveal new, specific predictions about the phenomena we are interested in that can then be assessed with other lines of evidence. Moreover, formal models can offer insights into complex processes, help develop or strengthen our intuition, provide a general framework for synthesizing accumulated knowledge and generating hypotheses to test, and identify key components as well as relevant spatial and time scales in their dynamics (Servedio et al. 2014).

3. Main Features of Models of Institutions

Before describing the specific mathematical approaches that have been used to model institutions, we will first describe some of the important conceptual issues that such models need to engage with. Following the general approach taken in this book, we can consider institutions as rules of social interactions. In developing models of institutional evolution, we need to think about (1) Why does a particular institution exist? What is its function? (2) What social processes are involved in establishing institutional rules, and who is involved

in setting these rules? and (3) How does institutional change occur? These three topics provide a guideline for thinking about the logically distinct features of models. Furthermore, in line with the need to simplify when developing models discussed previously, in some cases one of these issues may be more of a focus for a particular model than another issue.

3.1 What Is the Function of the Institution?

Institutional rules shape the interactions between individuals in a way that enables them to act collectively and provides some kind of benefit to at least some of those individuals. Institutional rules can help solve some of the challenges that such collective action situations present (Anderies and Janssen 2016; Powers, van Schaik, and Lehmann 2016). In some situations, all individuals would benefit if they could coordinate their actions and create "win-win" situations. A classic example is driving on one side of the road. It does not matter which side of the road people drive on as long as everyone does the same thing. Someone attempting to drive on the "wrong" side of the road is worse for all involved. Furthermore, having widely understood and accepted rules in these situations sets expectations and reduces costs of constantly having to assess and coordinate actions on a case-by-case basis.

Often coordination problems can be solved through communication, as it is in the interests of all parties to settle on complementary behaviors. More challenging situations involve social dilemmas, in which there are incentives for individuals to cooperate with others but also incentives for individuals to perform in ways that provide benefits to themselves at a cost to others. Common situations where such social dilemmas arise are in the creation of public goods (e.g., roads, public libraries, armies), which are costly to produce but can benefit anyone, and the use of common-pool resources (e.g., fisheries, forests, pastures), which are open to everyone but will be degraded if over-exploited. A key feature of both such situations is that it is difficult to exclude free riders, i.e., those who fail to contribute to the cost of producing the public good but are still able to benefit from using it, or those who take more than their fair share of the common resource. Despite these challenges, cooperation in such scenarios is seen in nature and in human societies. A significant effort in evolutionary biology, cultural evolution, and economics has focused on understanding conditions promoting the evolution of cooperation (McCauley and Moskalenko 2008; Nowak 2006; Richerson et al. 2016). Cooperation requires successful coordination of participating agents as well as a mechanism for reducing free riding (collective action problem, Olson 1965). It is also often important for the stability of cooperative systems to have mechanisms to reduce the likelihood (or consequences) of conflicts between individuals.

Social institutions enable humans to create situations where incentives are aligned with cooperative behaviors and collective action problems can be solved. For example, tax systems provide a means to collect revenues from group members that can be used to fund the creation of public goods and can be backed up by specific sanctions that can be imposed if an individual does not pay their taxes. The work of Ostrom (1990) has shown the kinds of design features that enable groups of people to collectively own natural resources and manage them in a sustainable way. These generally involve institutional rules that govern who is allowed to use the resource, how rules are established, how costs and benefits are shared, how resources are monitored, what sanctions exist for not following the rules, and how disputes are resolved. In this chapter, we will review a variety of different models that consider institutions involved in production, sharing, peer punishment, fairness, leadership, private property, and kinship. We will see how the functions performed by institutions can increase cooperation, coordination, and economic productivity and help mitigate collective action problems and reduce conflicts.

When thinking about the function of institutions, it is important to consider that individuals involved in such systems may not be fully aware of the functional significance of an institution or the role individual actions may play in the institution. In other words, institutions may work by people following particular rules without necessarily knowing why those rules work. In situations that involve complex interactions or

are causally opaque, explicit formal models may be particularly useful in understanding how institutions work. Simulation models of traditional Balinese irrigated rice fields show how rules that are linked to religious practices can help shape when and where particular fields receive water and actually have an important role in regulating the populations of certain pest species that can otherwise decimate the crops (Janssen 2007; Lansing 1991). It is also important to note that human societies are complex, and institutions may not always be optimal or function as intended (see Richerson, this volume). The effectiveness of a particular rule may depend on other features present in a society, which may affect how successful an institution is if it is adopted. Finally, an important aspect of institutions is that in many cases not everyone will benefit equally from specific institutional arrangements. As discussed in this chapter and in several places throughout this volume, key patterns in the long-term evolution of institutions involve changes in degrees of inequality in terms of wealth and power. Models play a key part in understanding whether such inequality is only beneficial for elites at the top of social hierarchies or whether there are also group-level benefits and, in modern societies, whether more inclusive forms of socio-political organization are better for economic productivity (Acemoglu et al. 2008).

3.2 What Social Processes Are Involved in Establishing Institutional Rules?

Another thing we need to consider in modeling institutional evolution is how rules are generated and modified. In an institutional system, individuals within a group construct the rules that affect how individuals are supposed to behave in a given situation and establish the consequences for not behaving in that way. In some cases, the establishment of rules may be fairly formalized, with a specific recognized system by which rules are set and can be changed. We are used to thinking that modern legal and political systems work in this way, but similar processes are at work in institutions in "traditional" or historical societies too (e.g., a chief, king, or other form of leader making judgments or issuing decrees about what behaviors are allowed or prohibited). Which institutional rules get put forward in this manner and whether rules are complied with and persist are also important social processes. For some rules, the social processes involved may be less explicit and may emerge or change through more gradual changes in accepted behaviors or other more normative processes. An important point to consider in thinking about how rules are generated and shaped is who is involved in setting rules. In more egalitarian settings, a majority of the population will have some influence on the nature of the rules that are adopted, while in more despotic scenarios, a more limited number of individuals set the rules that others must follow. As mentioned above, this issue can have implications for the function of institutions and who benefits from institutional arrangements.

Most evolutionary models of cooperation, which were developed within the context of biological systems, do not include the process by which rules can be generated and instead focus on the interactions between individuals who behave in certain ways that are set "genetically." Models of cultural evolution allow for information to be socially transmitted between individuals rather than just inherited from one generation to another via reproduction. While these cultural evolutionary models potentially allow for conventions or conformist behaviors to emerge, they do not explicitly capture the social processes by which rules are generated or agreed upon in the first place. This is particularly true when thinking about more formal institutions. Existing models of institutional evolution have employed a variety of different approaches to modeling social processes. In some models, the rules emerge spontaneously as a type of convention (deviations from which are detrimental for everybody) or by some processes of cooperation and conflict between players/agents (Bowles, Choi, and Hopfensitz 2003; Bowles et al. 2021; Frey and Atkisson 2020; Itao and Kaneko 2020). Some models have considered more explicitly the possibility that rules emerge by agreement (or bargaining) between involved parties (Currie et al. 2021; Gavrilets and Duwal Shrestha 2021; Powers and Lehmann 2013, 2014). In other models, agents have different levels of political power, and the most powerful of them have the most (or even complete) control over the rules of economic interactions (Acemoglu, Egorov, and Sonin 2021; Acemoglu and Robinson 2006, 2008; Acemoglu and Sonin 2012; Gavrilets, Anderson, and Turchin 2010;

Houle et al. 2022; Lawson and Oak 2014). In some cases, institutional changes occur through some other social processes that are not explicitly modeled or specified (Hooper, Kaplan, and Boone 2010; Isakov and Rand 2012; Roithmayr, Isakov, and Rand 2015; Turchin et al. 2013).

3.3 How Does Institutional Change Occur?

Institutional change involves some kind of change in the rules that govern interactions between individuals. This can potentially involve the creation of completely new rules or new roles or can involve adjustments to existing rules, such as changing the costs and benefits associated with different interactions or altering the nature of interactions between individuals. In models of institutional change these processes have to be simplified or abstracted in some way in order for models to be tractable. Often, we will only be able to examine one element of institutional change at a time or will be required to assume certain other elements are in place in a particular society. For example, Currie et al. (2021) modeled change in a peer punishment system by allowing individuals within a group to vote on the value of punishment to be given to any individuals who do not cooperate in a public goods game. In this case the establishment of such a rule or the ability of agents to vote or make decisions, which in reality is also a form of institution, are not examined, although we can abstractly think of a value of zero punishment as being the equivalent of having no system of punishment at all. In another example, Turchin et al. (2013) conceptualized institutions as traits that help societies achieve collective action—the lack of appropriate institutions means that societies will be organized on a smaller scale and will be less effective when competing with other groups that have institutions that allow them to be more coordinated. In this simulation model, individual-level processes were not explicitly considered; rather, groups were modeled in a highly abstract manner as having a vector of 10 binary traits indicating the presence or absence of institutional arrangements that would make them more likely to (1) win battles with others and form a larger "multi-agent" group and (2) stick together with other groups when the "multi-agent" groups are formed. In other existing models, institutions can change when the power of agents shifts due to some exogenous or endogenous events (Acemoglu, Egorov, and Sonin 2021; Acemoglu and Robinson 2006, 2008; Acemoglu and Sonin 2012; Gavrilets, Anderson, and Turchin 2010; Lawson and Oak 2014; Turchin et al. 2013), when the most powerful agents (e.g., the elite) change them to increase their benefits (Gavrilets and Duwal Shrestha 2021; Singh, Wrangham, and Glowacki 2017), to avoid the costs of political reaction from other agents (e.g., non-elite) (Acemoglu, Egorov, and Sonin 2021), or by decentralized stochastic transitions between different states (Bowles et al. 2021). In some models (e.g., Bednar and Page 2018) changes in one institution may affect the performance of another institution due to behavioral spillovers, causing the institution to change in its effect, if not in its form.

It is important to distinguish between the processes of change that occur within a group that alter institutional arrangements and the broader-scale changes we witness in the diversity of different institutions when comparing between different groups. In some models we may be interested in understanding the individual-level (or micro-level) processes that are occurring within the group that lead to changes in institutional rules. Models in human evolutionary ecology (e.g., Holden, Sear, and Mace 2003; Smith and Choi 2007) focus on how different systems of social organization could emerge if individuals behave adaptively under different social and environmental situations. This approach can be extended to consider institutions more explicitly. In the peer punishment model described above, for example, the individuals are attempting to find values of punishment that will lead to them receiving greater net benefits from the public goods game. An important consideration in modeling situations at this level involves the assumptions that are made about the agents' level of knowledge or the ability of agents to make "correct" decisions. Even if individuals are not able to make decisions that are perfectly optimal (which is almost certainly the case in the vast majority of real-world scenarios), selective processes can still lead the system toward whatever the optimal outcome is in the scenario being modeled. For example, if payoffs are linked to reproduction and behavior in the model is inherited from

parents to offspring, then sub-optimal parents will leave fewer offspring. In cultural evolutionary models, processes such as payoff-biased copying will also lead to increases in beneficial behaviors—although other processes may lead to the spread of more maladaptive behaviors (Boyd and Richerson 1985).

In other cases, we may be more interested in understanding how processes play out in terms of a "meta population" of different groups (the "macro level"). This is often the case when considering why some institutions might be more common than others or in considering long-term trends in institutions as we do in this book. Institutions are group-level features, and some may become more common through processes of cultural group selection. This can involve direct between-group conflict (Bowles, Choi, and Hopfensitz 2003; Bowles et al. 2021) such as warfare—institutions that enable groups to defeat other groups will become more common if the defeated group is killed, disbanded, or incorporated into the winning group. Cultural group selection can also occur through copying (selective payoff-biased imitation [Gavrilets and Duwal Shrestha 2021]) or through differential migration (individuals "vote with their feet" [Boyd and Richerson 2009]). The model of Turchin et al. (2013) described above was concerned primarily with understanding how direct competition between groups may have driven the evolution of increasingly larger societies. In this simulation, institutions that enabled societies to be organized on a larger scale were modeled as changing at random but in a way that meant it was always easier to lose these traits than to gain them. In the absence of between-group competition these institutional traits never increased in frequency. Explicit modeling can help make the distinctions between micro-level and macro-level processes of change clearer and allows us to investigate the potential connections between these levels and to understand when different explanations might be complementary or contradictory.

4. How Are Institutions Modeled?

We will now consider how the conceptual issues introduced above can be developed into formal models. Theoretically, institutions have generally been studied using models and methods from game theory, which is the most appropriate tool for studying social interactions. Under this approach, individual entities are involved in situations known as games, where different behaviors or strategies will produce different outcomes, with the payoffs to an individual often depending not only on what they do but also on what other individuals playing the game do as well.

There are several flavors of this theory including classical (Fudenberg and Tirole 1992), evolutionary (Sandholm 2010), mean-field (Gomes and Saúde 2014; Tembine 2017), and quantum (Piotrowski and Sladkowski 2003; Siopsis, Balu, and Solmeyer 2018) game theories. Most of our focus will be on evolutionary game theory models, which are the easiest to adapt to account for cultural evolution and are also more concerned with the dynamics of changes in strategies than classical game theory. Since the literature on cooperation and cultural evolution involving evolutionary game theory is rather large, we will limit our discussion to relatively simple models with transparent conclusions that are most relevant for understanding institutions and institutional evolution. Using standard terminology in this field, we will refer to entities engaged in social interactions (e.g., individuals or groups of individuals) as agents or players. Below we will look at mathematical models that interpret institutions as rules of social interactions that agents accept voluntarily or involuntarily.

A general theoretical framework for modeling institutions using game theory developed by Hurwicz (1996) is to consider two interrelated games: an "economic game" and a "political game" (Powers, van Schaik, and Lehmann 2016). In this setup, the outcome of the economic game reflects the immediate payoffs to the agents and the kind of scenario that individuals might find themselves in in attempting to extract or acquire some kind of resource. The political game involves agents in separate interactions that establish rules that affect the economic game. For example, we can have an economic game that is based on the prisoner's dilemma, which

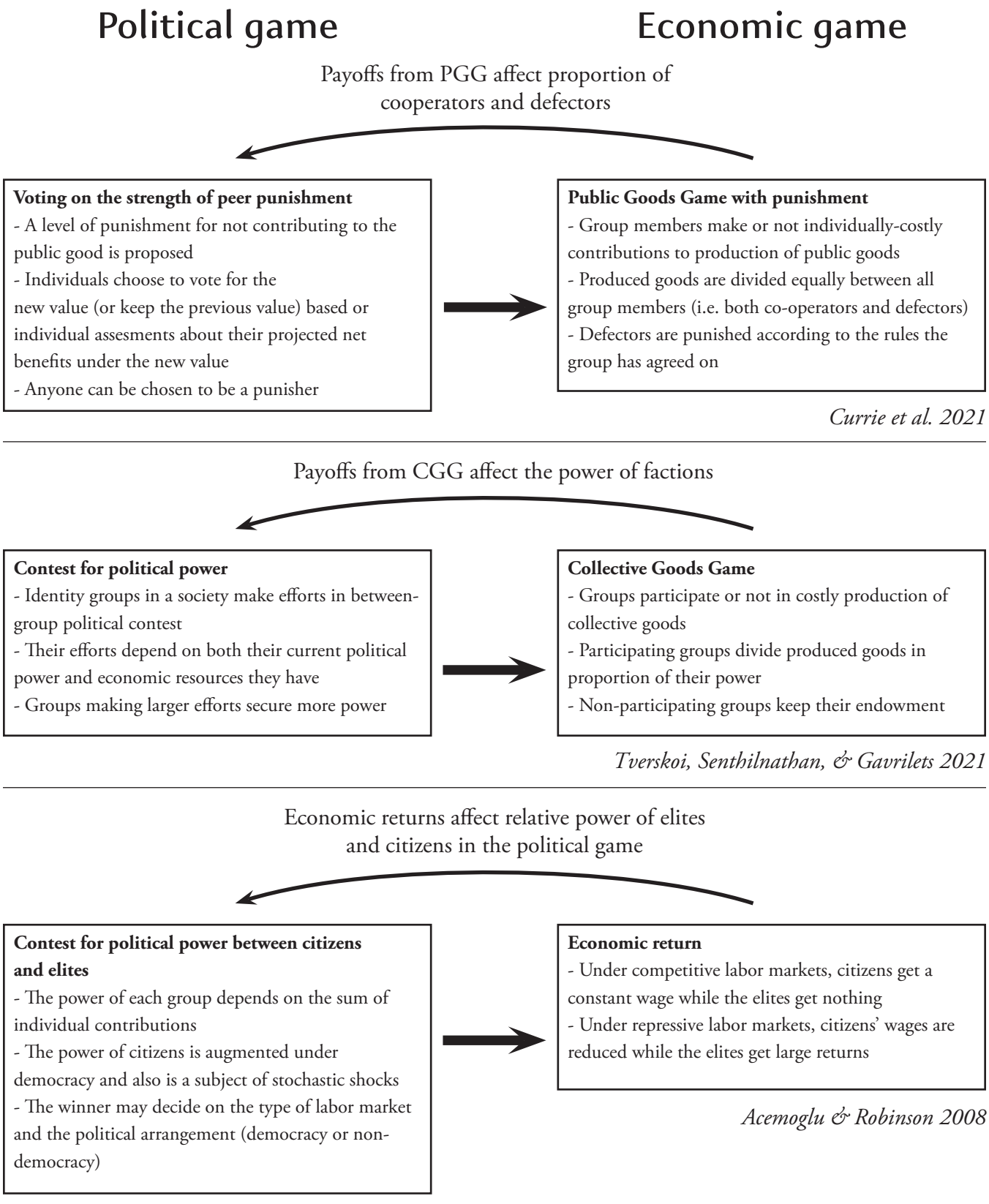

Figure 1. Schematic diagram of the "political game – economic game" framework for modeling institutions proposed by Hurwicz (1996). The approach is extremely flexible and can encompass a wide variety of different economic games and political games. We provide three examples to illustrate this approach based on Currie et al. (2021), Tverskoi, Senthilnathan, and Gavrilets (2021), and Acemoglu and Robinson (2008).

in its basic setup will favor agents who defect rather than cooperate. The political game, however, may involve a system whereby it is decided that agents who defect receive a certain level of punishment. If the punishment level decided in the political game is sufficiently large, it will create conditions that favor cooperation as the net payoff is greater. Administering the punishments carries a cost, however, so the punishment level cannot be too high; otherwise, it decreases the net payoff to the co-operators. This approach is extremely flexible and can be used in a huge variety of different economic game setups or different scenarios for the political game (i.e., punishment is just one example; see figure 1 for three such examples). Indeed, this approach can be used to consider the issues we highlighted earlier such as different agents having different levels of influence in the economic game or different social processes used to set the rules in the political game (e.g., explicit majority-rule voting, deliberation until a consensus is reached). The two-stage approach itself can be extended to include a "constitutional" game that establishes how the rules themselves might be set (Hurwicz 1996; Ostrom 2009). Some of the models of institutions discussed below use this approach explicitly, but in most others one of these two games is only implicit or both games are merged into one.

In the following sections we discuss different models relevant to the evolution of institutions in more detail. We organize our discussion of different models in an order that reflects the themes and overall organization of this book. We start by examining models that have been applied to understanding the kinds of institutions that are fundamental to human social organization and are found to some extent in most, if not all, societies. We then examine models that are relevant to understanding historical transitions toward more complex forms of socio-political organization, e.g., formal leadership, chiefdoms, and states and empires. We then finish by considering models of institutional change in the kind of state-level societies that represent the defining form of organization in the modern world and recent past. Our aim here is not to conduct a comprehensive survey of all models that are relevant to understanding institutional evolution and organizational complexity, but rather to provide a somewhat representative sample of the different kinds of models that have been developed and the different approaches that are taken. At the end of each section, we summarize the main features of those models that have been surveyed, which helps identify commonalities and differences between models and also highlights what the models have not examined.

4.1 Small-Scale Societies and Foundational Human Institutions

Sharing/redistribution. Sharing and cooperation are widespread in small-scale societies (Gurven and Jaeggi 2013; Jaeggi and Gurven 2018; Marlowe 2004; see also chapters 4 and 5, this volume). Bowles, Choi, and Hopfensitz (2003) considered a population of individuals subdivided into groups. Group members have two strategies: engage in a costly cooperation or free ride. There is an institution of redistribution in place: a certain tax is imposed on group members that is used to (partially) compensate cooperators for the costs they pay. Groups differ in the amount of "tax" imposed on group members, with higher taxes leading to more cooperation. The institution of redistribution then evolves by group selection resulting from direct between-group conflict, leading to the spread of cooperative high-tax groups at the expense of non-cooperative low-tax groups. An additional component of the Bowles, Choi, and Hopfensitz (2003) model was an institution of assortative within-group interactions leading to cooperators more likely interacting with other cooperators. (Note that assortativity is a powerful mechanism of maintaining cooperation and reducing the rider-rider problem [Fletcher and Doebeli 2009].) Overall, the model shows that evolution of the institutions of sharing and assortativity can occur by group selection, leading to the evolution of individually-costly, group-beneficial cooperative behaviors.

Fairness. Within the context of dyadic interactions between players, each with two possible actions, there are four possible outcomes of a game. Depending on the rules of the game specified by the corresponding payoff matrix, the two players can have their own strict ranking over these outcomes. If one considers the set of all possible such games, there are 144 unique ways that the two agents can assign their own strict rankings over

the four outcomes. The corresponding rules (i.e., ordinal payoff matrices) can then be arranged in a structured network that allows one to study the coevolution of individual actions and the rules of the game. Frey and Atkisson (2020) assumed that both players prefer games that are stable (i.e., have a unique Nash equilibrium), predictable (i.e., have an equilibrium in pure strategies), and efficient (i.e., ensure highest payoff to the focal player). They also postulated that the player with a higher payoff can change the rules of the game. Frey and Atkisson (2020) then demonstrated that under these self-interested dynamics, in about 50 percent of cases the system evolves toward one of "win-win" games in which two players share the same top-ranked outcome and thus view it as fair. However, increasing the number of players quickly decreases the proportion of attractors leading to fair outcomes. Frey and Atkisson (2020) argued that institutional evolution can be a mechanism for encouraging the spontaneous emergence of cooperation among small groups of inherently selfish agents. The approach taken in this model does not clearly distinguish between an "economic game" and a "political game." However, it is possible to envisage a model that produces similar dynamics but distinctly separates these elements. This may be more informative for scenarios where the economic game itself cannot be changed but can instead be impacted through institutional rules.

Private property. Bowles and Choi (2013) and Bowles et al. (2021) argued that the emergence and spread of farming around 12,000 years ago would not have been possible without a simultaneous emergence of the institution of private property, which allowed farmers to keep their products. They built several models in support of their argument. In the models, the agents choose a technology to use (farming or gathering) and a strategy in dyadic interactions of which there are three possible types. The "sharer" concedes half of the product to the other or the whole product if the other claims it. The "bourgeois" claims the entire product if it is in their possession; if not, they act like a sharer. A "civic" shares when they meet either sharers or other civics, but tries to punish anyone claiming the whole product. In Bowles and Choi (2013), individuals updated their strategies by selective payoff-biased imitation. In Bowles et al. (2021), they played best-response to the previous distribution of the strategies in their group or, with a small probability, used intentional idiosyncratic play, i.e., chose a strategy that would give them a higher payoff if sufficiently many other group members chose it. The latter was introduced to capture things like outrage, a quest for personal dignity in opposition to injustice, and other motives not directly related to material payoffs. The authors also allowed for assortativity in the way interacting pairs are formed. In their models, coevolution of technology and convention occurs by stochastic transitions between different equilibria within a population and between-group conflict (i.e., cultural group selection) spreading an innovation across the whole system. Bowles and Choi (2013) used archaeological and climate data to validate their model and support its conclusions.

Peer punishment of free riders. Powers and Lehmann (2013) argued that the system of monitoring and punishment of free riders does not have to be imposed on the group by leaders but rather can evolve by consensus among group members. In their model, group members vote on the amount of jointly produced resources that is allocated to monitoring and punishment. A similar model was studied by Currie et al. (2021), who assumed that group members vote by consensus on a value for a fine for being a free rider. Individuals are randomly selected to be punishers who also pay the cost of punishing. Individual preferences evolve by payoff-biased reproduction. The results suggest that spontaneous emergence and stable maintenance of the institution of peer punishment is indeed possible. However, it should be noted that this model made some strong assumptions about the ability of agents to assess which values of punishment would be most beneficial to them.

Marriage patterns. In many human societies, the rules of social relationships with others, such as cooperation, rivalry, or marriage, are often determined by the clans the parties belong to. These rules also differ between different clans. To explain the origin and diversity of these rules, Itao and Kaneko (2020) built an agent-based mathematical model. The agents in their model are lineages that differ in two culturally transmitted traits: one that defines the relatedness of lineages and another that defines their mating preferences.

The lineage's growth rate increases with the number of closely related lineages (with whom the focal lineage can cooperate) and decreases with the number of lineages with similar mating preferences (with whom the focal lineage competes for mates). Each lineage splits into two when its population size doubles and is eliminated when its size goes to zero. Itao and Kaneko (2020) showed that as a result of cultural evolution, lineages become clustered in the two-dimensional space of their traits, leading to the emergence of clans with the incest taboo. Depending on parameter values, the model predicts the emergence of generalized exchange of mates (if cooperation is strongly needed) or restricted exchange (when the mating conflict is strict) between lineages. They argue that their model may explain the geographical distribution of kinship structures in Indigenous societies.

In summary, in all models discussed in this section, either the emergence of institutions is not considered explicitly or they emerge by voluntary agreements between agents. Institutions evolve by cultural group selection (voting by feet, payoff-biased imitation, or between-group conflict), stochastic transitions (Bowles et al. 2021), and self-interested-design (Frey and Atkisson 2020). Institutional traits are transmitted culturally, but otherwise effects of cultural differences are not explored explicitly. Most models used a combination of analytical derivations and agent-based simulations.

4.2 Political Evolution: The Emergence of Leadership and Complex Societies

For most of humanity's existence it is thought that people lived in small egalitarian bands or villages that were politically autonomous. While leadership can be found in such societies, and indeed forms of leadership are found in non-human species, these roles were often ephemeral and linked to certain specific activities or were otherwise based on age or experience. More formal and centralized modes of leadership emerged over the course of human history, enabling the emergence of larger and larger politically coordinated groups (polities). Models can shed light on the conditions under which leadership might be favored and the institutional dynamics involved in the evolution of complex polities such as chiefdoms, states, and empires.

Leadership. Leaders can coordinate the actions of group members, making their efforts more efficient, helping groups reach consensus more easily, monitoring and punishing free riders, rewarding contributors, and fostering pro-social norms and values (Garfield, Hubbard, and Hagen 2019; Garfield, Syme, and Hagen 2020; Glowacki and von Rueden 2015; Perret, Hart, and Powers 2020; Smith et al. 2016). Synthesis of biological and social-science data (Smith et al. 2016) supported by some modeling (Gavrilets, Auerbach, and van Vugt 2016; Perry et al. 2018; Smith et al. 2016) shows that leaders and followers emerge naturally as a result of heterogeneity in preferences, motivation, personality, physical characteristics, information available, and other features affecting individual performance in different activities. Several papers studied how the institution of leadership evolves in collective actions.

Hooper, Kaplan, and Boone (2010) modeled competition of groups involved in production of public goods. They assumed that groups can be leaderless or with leaders who monitor and punish free riders and also collect tax for their effort. Leaders differ in the tax level they impose on group members while regular group members can choose which group to join. The institution of leadership here evolves by "voting by feet," which is one of the mechanisms of cultural group selection (Richerson et al. 2016). Hooper, Kaplan, and Boone (2010) showed that the institution of leaders receiving a share of group productivity can evolve when "the cost of monitoring and sanctioning fellow group members is high enough that individuals would be unwilling to enforce cooperation on their own, but not so high that the gains from cooperation cannot cover the leader's enforcement costs" (p. 642).

Isakov and Rand (2012), Roithmayr, Isakov, and Rand (2015), and Gavrilets and Duwal Shrestha (2021) considered similar models but postulated that the institution of leadership is already in place and instead focused on the evolution of the level of monitoring and punishment of free riders administered by the leaders. Isakov and Rand (2012) and Roithmayr, Isakov, and Rand (2015) assumed that institutions evolved by

random innovation and selective payoff-biased imitation with leaders copying other more successful leaders. Gavrilets and Duwal Shrestha (2021) compared this mechanism with self-interested design (Singh, Wrangham, and Glowacki 2017), in which leaders attempt to maximize their own payoff by predicting behavior of regular group members (Perry and Gavrilets 2020; Perry et al. 2018). Their conclusion was that although under some conditions selective payoff-biased imitation and self-interested design can be equally efficient, the latter is more powerful.

In Powers and Lehmann (2014), leaders differed in the amount of tax they imposed on subordinates and focused their effort on coordinating group members to increase production of the group (rather than monitor and punish free riders). They showed "that voluntary leadership without coercion can evolve in small groups, when leaders help to solve coordination problems related to resource production" (p. 1), such as, for example, building and maintaining an irrigation system.

A qualitative change in forms of socio-political organization happened when local-level units ("villages") began aggregating into larger and more complex hierarchically structured political units, leading over time to the emergence of chiefdoms, states, and empires. Historically, this process first took place in Mesopotamia, East Asia, South America, and Mesoamerica, followed by secondary developments elsewhere (Service 1978). Two models considered the emergence and evolution of the institution of subordination and tribute/tax flow involved in the establishment and functioning of these entities.

Chiefdoms. The process of aggregation of villages first led to the appearance of simple chiefdoms (Steponaitis 1978, 1981; Wright 1984), in which one village controlled (and received tribute from) several subordinate villages. More complex polities were characterized by greater numbers of subordinate levels, with complex chiefdoms, paramount chiefdoms, and state societies typically defined as those polities with two, three, and four or more administrative levels above the local or primary community, respectively (Anderson 1994; Flannery 1972; Stevenson-Hinde and Zunz 1978; Wright 1984; Wright and Johnson 1975).

Gavrilets, Anderson, and Turchin (2010) studied the processes of formation and evolution of chiefdoms using spatially explicit agent-based simulations. In their model, each agent is a local community (e.g., village), and each local community is a part of a polity that has a hierarchical structure. Specifically, each community in a polity except for the one at the top of the hierarchy (the "chief community") has one superior community and may have up to several subordinate communities. Villages possess a certain baseline amount of resources that can be interpreted as a measure of the settlement's catchment size (Steponaitis 1981). Villages also receive additional resources as a tribute from their subordinate communities and send a part of their resources to their superior community. Polities are formed as a result of conflicts, with the losers paying tribute to the winners. Polities can break as a result of succession of subordinate communities due to rebellion or the death of the paramount chief. The model also accounted for "scalar stress," that is, a decrease in the ability of leaders to process information and maintain efficient control over subordinates as their number (herein, the number of subordinate villages) increased (Johnson 1982). Their model was an attempt to formalize Carneiro's (1970, 1981) argument about the importance of warfare and circumscription (environmental, due to the resource concentration, or social, due to the presence of other human groups nearby) for the appearance of chiefdoms.

A general prediction of their model is continuous stochastic cycling in which the growth of individual polities in size, wealth/power, and complexity is interrupted by their quick collapse. This prediction is well in line with archaeological and historical evidence on "chiefly cycles," when centers of power and authority shifted from one location to another over the landscape (Anderson 1994; Leach 1954; Wright 1984). The model dynamics are mostly controlled by two parameters, one that scales the relative advantage of wealthier polities in between- and within-polity conflicts, and the other that relates to the chief's expected time in power. The model predicts that stability of large and complex polities is strongly promoted if there exists a well-defined and accepted means of succession and if control mechanisms are internally specialized.

Large states and empires. Turchin et al. (2013) modeled the origin and spread of empires. They built on a model in Gavrilets, Anderson, and Turchin (2010) but with three important additions. The first was that the model was implemented within a realistic landscape of the Afroeurasian landmass (divided into a grid of 100×100-km squares), accounting for restrictions on movement due to terrain ruggedness and bodies of water. The second was that it included a process for the diffusion of military technologies that arose on the interface between the Eurasian steppe belt and agrarian societies living next to it. The third is that it explicitly accounted for culturally transmitted ultra-sociality norms and institutions that simplified within-polity cooperation and thus increased economic and military power. (An example of an ultrasocial norm is generalized trust [Turchin 2013], while examples of ultrasocial institutions are professional bureaucracies [Mann 1986], formal education, and universalizing religions.)

The model predictions were tested against a large dataset documenting the spatio-temporal distribution of historical large-scale societies in Afroeurasia between 1,500 BCE and 1,500 CE. The model-predicted pattern of spread of large-scale societies was very similar to the observed one. Overall, the model explained 65 percent of variance in the data. These results supported theories that emphasized the role of institutions in state-building and suggested a possible explanation as to why a long history of statehood is positively correlated with political stability, institutional quality, and income per capita (see Flitton 2022).

In summary, in the models studied the institution of leadership emerges as a result of between-individual heterogeneity and evolves by cultural group selection or self-interested design. In the model of chiefdoms, the rules were set by more powerful agents (Gavrilets, Duenez-Guzman, and Vose 2008), while in the Turchin et al. (2013) model of empires, the main evolutionary force was cultural group selection (in the form of conflict between groups and the relative advantage of larger groups that are held together by novel institutions). In all models, the relevant traits were transmitted culturally, but otherwise effects of cultural differences were not explored.

4.3 Political and Economic Institutions in Modern Societies

Models discussed in this section consider institutions regulating economic and political relationships in state-level societies. The former focus on the type of goods produced by individual agents or the society as a whole. The latter determine the share of goods going to different factions of the society and also how economic rules are changed.

Bowles et al. (2021) modeled the emergence of the national state as a transition between two types of society, each consisting of commoners producing goods and elites collecting tax (see also Acemoglu and Robinson 2006; Besley and Persson 2009). In the "weak state," commoners are unproductive and tax-evading while elites keep the taxes for themselves without providing public goods. In the "strong state," commoners are highly productive and tax compliant while elites devote tax revenue to the production of public goods. In their model, the transition from the weak state to the strong state occurs stochastically if enough commoners innovate with new types of production, increasing their output despite the absence of public goods from the state, which then induces the elite to best-respond by producing public goods.

A number of papers have modeled transitions between different types of societies driven by competition between elites and commoners (or rich and poor). In Acemoglu and Robinson's (2008) model, the society can produce two types of public goods, one of which is preferred by citizens and the other by elites; only one type can be produced at any time period. There are two possible political regimes: democracy and non-democracy. The group with greater political power determines the current type of public goods produced and the political institutions at the next time moment. The power of elites is defined by their investment into political contest. For commoners, besides their investment into political contest, there is an additional source of power in democracy (because they are more numerous) as well as a stochastic component describing their ability to

occasionally solve the collective action problem. Acemoglu and Robinson (2008) use their model to study characteristics of stochastic transitions between different types of society driven by these fluctuations.

Besley and Persson (2008) modeled competitions of two groups, with the winner making decisions on taxes and spending to maximize its own group benefits. The turnover of groups in power is stochastic. In Persson and Tabellini's (2009) model, there are two regimes, democracy and autocracy, with different productivities. Opportunities to change the regime arise stochastically and when this happens, the probability that the state will remain (or become) a democracy depends on the number of citizens willing to defend it. Citizens' decisions reflect a cost-benefit analysis. Acemoglu and Sonin (2012) considered two types of individuals, elites and the middle class, and three types of states: absolutist monarchy, constitutional monarchy, and democracy. In absolutist monarchy, the elite decides which regime will prevail tomorrow. In both constitutional monarchy and democracy, the middle class decides the next period's regime. They show that a current social arrangement can be stable due to the instability of alternative arrangements that are preferred by sufficiently powerful groups. They also show that economically beneficial changes may be resisted because of further changes they will engender. Gorodnichenko and Roland (2015) allowed for two classes of citizens, rich and poor, and three types of social arrangements: bad autocracy, good autocracy, and democracy. In each period, citizens are able to overcome the collective action problem and successfully change the regime with a certain probability. Their decisions to participate in collective action depend on expected benefits. Gorodnichenko and Roland (2015) explicitly contrasted collectivistic and individualistic cultures in their propensities to establish different political regimes.

Roland and Xie (2016) also studied the effect of collectivistic and individualistic cultures on citizens' willingness to participate in collective actions. They considered two types of collective actions: a popular uprising to overthrow an existing ruler deemed illegitimate and an institutional innovation leading to the establishment of new political institutions (such as when monarchy is replaced by a republic or autocracy by democracy). Besides material benefits, their model allows for psychological benefits arising from the opportunity to "stand out" and from self-satisfaction with conforming to the social norm of revolting in cases when revolting is "just." Roland and Xie (2016) showed that psychological payoffs can lead to an alleviation of the collective action problem depending on its type and on the culture of the society. These additional payoffs can offset the free rider effect and push players to participate in collective actions.

Bisin and Verdier (2018) considered a society of agents separated into two groups defined in terms of certain characteristics, e.g., cultural traits. The groups have different preferences with respect to certain sets of actions and policies. Bisin and Verdier (2018) modeled institutions as weights given to the preferences of the two groups in the overall utility function that the society as a whole attempts to maximize. The weights can also be interpreted as the relative political power of the first group. Bisin and Verdier (2018) postulated that the institution (i.e., weights) changes from one time step to another to maximize the current social preference by means of future policy choices. They also allowed for relative frequencies of the two groups to change in time as a result of vertical (i.e., by parents) and horizontal (i.e., by peers) social influence on preferences. They proceeded to study the joint evolution of culture and institutions. Their model may display rather complex nonlinear behavior including oscillations, strong sensitivity to initial conditions, and threshold effects, implying a great diversity of resulting institutions. Bisin and Verdier (2018) applied their model to study how culture and institutions affect the sustainability of extractive societies as well as the formation of civic capital and of legal systems protecting property rights.

Acemoglu, Egorov, and Sonin (2021) considered a society composed of multiple factions that differ in their preferences for an economic policy adopted by the society. They modeled an "institution" as a set of weights assigned to different factions in collective decision making. Under their assumptions, there is always just one faction (effective median voter fraction) that has complete control over the current economic policy and future political arrangements. Acemoglu, Egorov, and Sonin (2021) discussed conditions on parameters

of the utility function that would guarantee stability of the current institution or would cause its change. The latter would happen if the faction in power prefers to change the current institution to increase its economic benefits or to reduce the likelihood that future shocks would cause it to lose power. They also considered how the initial social institution influences future changes in the society, how stability of institutions can be increased, and how social mobility and cultural change affect these processes.

Lawson and Oak (2014) developed a rather different dynamic approach for modeling evolution of a society composed of multiple factions differing in power. They assumed that each faction possesses a certain amount of resources and at each time step can either cooperate or not in the redistribution of resources in proportion to the power of the participating factions. To model the dynamics of power, they adapted the replicator equation from evolutionary biology (Hofbauer and Sigmund 1998). Under this equation, power begets more power. An interesting outcome of their model concerns non-equilibrium dynamics in which growing inequality among cooperating factions leads to an avalanche of defection, followed by a decrease in inequality that in turn creates conditions for cooperation. Then inequality starts growing again and the cycle repeats itself. Lawson and Oak (2014) discussed possible applications of their model to explain the growth and collapse of historical societies and the role of inequality in these processes (Turchin 2003).

Houle et al. (2022) generalized the Lawson and Oak (2014) model in several directions. First, they explicitly considered cooperation of factions in the production of collective goods (rather than redistribution of certain resources). Second, they described the power dynamics using the standard Tullock contest model (Konrad and Kovenock 2009; Tullock 1980) to which they added a parameter explicitly controlling the strength of checks and balances mechanisms restricting the ability of powerful factions to bend the rule of the economic interactions to their own favor. Third, they added non-material components (interpreted specifically as conformity with the majority of the society and allegiance to the state) to the factions' utility function. Similar to Lawson and Oak (2014), their model predicts that growing economic and political inequality tends to lead to the collapse of cooperation between factions that were initially seeking to cooperate. However, certain mechanisms can delay this process, including the decoupling of political and economic power through rule of law and allegiance to the state or dominant faction. Increasing the number of factions in the society decreases stability. Counterintuitively, anti-conformity (a social norm for independent action) can also stabilize a society by preventing initial defections from cascading. Heterogeneity in baseline resources between factions makes the society more stable to negative effects of inequality. The availability of certain material resources that can be acquired by the state without cooperation with other factions (e.g., natural resources) has the opposite effect. Houle et al. (2022) tested predictions of their model using data on horizontal inequality (i.e., economic, social, and political inequality between different identity groups) in modern societies. Using social unrest as a proxy for the breakdown of cooperation in society, they found support for many of the predictions of their theory.

Tverskoi, Senthilnathan, and Gavrilets (2021) built an individual-based extension of the model in Houle et al. (2022). They explicitly considered individual members of factions and modeled their decisions on whether or not to participate in the production of public goods that the faction then can invest in between-group cooperation or keep for itself. Using analytical approximations and agent-based simulations, Tverskoi, Senthilnathan, and Gavrilets (2021) showed that the model exhibits rich behavior characterized by multiple stable equilibria. They also observed non-equilibrium dynamics but under much narrower conditions than Houle et al. (2022). This suggests that societies in which individuals act independently are more stable than those in which actions of individuals are completely synchronized. Tverskoi, Senthilnathan, and Gavrilets (2021) also showed that small groups can be more successful in competition than large groups if the jointly produced goods are rivalrous and the potential benefit of cooperation is relatively small. Otherwise, large groups dominate.

In summary, in the models discussed in this section, more powerful factions benefit by setting institutions to their own advantage, institutions change as a result of political contest between factions also

affected by stochastic fluctuations, and the effects of culture are treated mostly in terms of differences between collectivistic and individualistic societies.

5. Discussion

In this chapter, we have discussed the ways in which formal mathematical modeling of social process can inform the study of the evolution of institutions and organizational complexity. We have argued that a full framework for modeling institutional evolution needs to consider the function or purpose of particular institutions, the social mechanisms by which rules are established and set, and the processes by which institutional rules and structures change over time. These aspects of institutional evolution are distinct yet often interrelated, and there is not necessarily a single "best" way to model institutional evolution. However, we have built on previous work in suggesting that evolutionary game theory involving a multi-stage approach offers a promising means by which these different elements of institutional evolution might be effectively captured and analyzed in a tractable way.

We have provided examples of several models that are relevant to modeling institutional evolution, with a focus on those that are particularly relevant to the themes and organization of this volume. All of these models have given valuable insights and served the intentions for which they were developed. However, few, if any, follow an explicit multi-stage modeling approach or address all the potentially relevant aspects of institutional evolution. For example, Frey and Atkisson (2020) provide a framework in which there is only a single-stage "game," but the game being played by agents can itself change. When thinking about this in terms of real-world systems, this means that either something fundamental about the adaptive problem being considered has changed (i.e., there is a change in nature of the problem being modeled at the functional level) or institutional rules have changed and this is being reflected in the net or overall payoffs. If the latter is the case, then this modeling approach does not explicitly consider either the social mechanisms involved in shaping institutional rules or the interplay between the political game and economic game. Even models such as Currie et al.'s (2021) model of peer punishment, which does explicitly separate the economic game from the political game and allows for punishment to be introduced into a system without punishment, treats the voting process (i.e., the social mechanism that determines the value of punishment) as a given element, but which in reality could take different forms. Similarly, models of pooled, centralized punishment often assume the presence of an institution that is capable of performing the punishment. At a larger scale, Turchin et al.'s (2013) geographically explicit model of empire and state dynamics abstracts away the within-group mechanisms of institutional implementation and change in favor of focusing on the processes of between-group competition and longer-term patterns and processes of change. None of these points are fundamental criticisms of these particular models. Indeed, institutions are complex things, and it is important that we develop simple models. A challenge for future modeling efforts is to extend, combine, or otherwise build on these existing examples.

Thinking about how to model institutional evolution draws attention to the fact that institutions are a particular form of social phenomena that involve considering the features of agents, the various different ways they interact, and how these co-evolve and create emergent phenomena at different levels (Currie et al. 2016; Smaldino 2014). Developing models of institutional evolution further will require an interdisciplinary approach that combines the strengths of different approaches. Models from the social sciences (e.g., political science, economics) tend to focus more on the social mechanisms involved in institutional phenomena but rarely consider longer intergenerational changes in populations or their consequences, while the reverse is often the case for models informed predominantly by evolutionary theory (Currie et al. 2021). Our chapter has not included all the different approaches that are developing models relevant to institutional evolution, and different disciplines, particularly in areas such as sustainability science and social-ecological systems, are

likely to prove very valuable (Ostrom 2009; Polhill et al. 2016; Powers, Ekárt, and Lewis 2018; Schlüter et al. 2019; Weissing and Ostrom 1991).

Modeling approaches in cultural evolution combine elements from social and natural sciences to some extent; however, the traits examined are typically modeled at the individual level without considering some of the structural and emergent properties that institutions involve (Smaldino 2014). A particular area that requires further development is incorporating how other non-institutional aspects of culture affect changes in institutions and vice versa (Currie et al. 2016). By culture we mean specific characteristics and knowledge of a particular group of people, including language, customs, norms, beliefs, religion, shared historical knowledge, cuisine, music and arts, etc., that are passed on through social learning. The existence of certain norms, beliefs, or other cultural practices may shape how institutions change or inform whether particular institutions are adopted because some rules and structures may be easier to understand or may be more palatable to a group and thus increase compliance and effectiveness. On the other hand, the introduction of new rules can create a selective environment that favors certain norms of behavior. In most existing models such as those surveyed above, culture is not modeled explicitly. However, there are models where culture explicitly affects certain model parameters (e.g., Houle et al. 2022) or coevolves with institutions (Bisin and Verdier 2018; Bowles et al. 2021; Güth and Ockenfels 2005). At the same time, many of the models described above incorporate some element of cultural transmission such as imitation. Extending a cultural evolutionary approach is likely to prove valuable (Currie et al. 2021), and, more generally, evolutionary thinking allied with information and approaches from social sciences and humanities disciplines can provide a broad conceptual framework in which institutional evolution can be effectively modeled and better understood.

Acknowledgements. We thank J. Bednar, P. J. Richerson, and J. Wallis for comments and suggestions. S.G. was supported by the John Templeton Foundation, the U. S. Army Research Office grants W911NF-14-1-0637 and W911NF-18-1-0138, the Office of Naval Research grant W911NF17-1-0150, and the Air Force Office of Scientific Research grant FA9550-21-1-0217. T.E.C. was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Project: The cultural evolution ecology of institutions; Grant Agreement 716212).

References

Acemoglu, D., G. Egorov, and K. Sonin. 2021. Institutional Change and Institutional Persistence. In The Handbook of Historical Economics, ed. A. Bisin and G. Federico, 365–90. London: Academic Press.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared. 2008. Income and Democracy. American Economic Review 9:808–42.

Acemoglu, D., and J. A. Robinson. 2006. Economic Origins of Dictatorship and Democracy. Cambridge: Cambridge University Press.

———. 2008. Persistence of Power, Elites, and Institutions. American Economic Review 98:267–93.

Acemoglu, D., and K. Sonin. 2012. Dynamics and Stability of Constitutions, Coalitions, and Clubs. American Economic Review 102:1446–76.

Anderies, J., and M. Janssen. 2016. Sustaining the Commons. Arizona State University, Center for Behavior, Institutions and the Environment, Arizona State University.

Anderson, D. G. 1994. The Savannah River Chiefdoms: Political Change in the Late Prehistoric Southeast. Tuscaloosa: University of Alabama Press.

Bednar, J., and S. E. Page. 2018. When Order Affects Performance: Culture, Behavioral Spillovers, and Institutional Path Dependence. American Political Science Review 112:82–98.

Besley, T., and T. Persson. 2008. Wars and State Capacity. Journal of the European Economic Association 6(2–3): 522–30.

———. 2009. The Origin of State Capacity: Property Rights, Taxation, and Policy. American Economic Review 99:1218–44.

Bisin, A., and T. Verdier. 2018. On the Joint Evolution of Culture and Institutions. Journal of Economic Literature. https://www.econ.nyu.edu/user/bisina/ciaff.pdf.

Bowles, S., and J.-K. Choi. 2013. Coevolution of Farming and Early Holocene. Proceedings of the National Academy of Sciences USA 110:8830–35.

Bowles, S., J.-K. Choi, and A. Hopfensitz. 2003. The Co-evolution of Individual Behaviors and Social Institutions. Journal of Theoretical Biology 223:135–47.

Bowles, S., J.-K. Choi, S.-G. Hwang, and S. Haidu. 2021. How Institutions and Culture Change. In The Handbook of Historical Economics, ed. A. Bisin and G. Federico, 391–434. London: Academic Press.

Boyd, R., and P. J. Richerson. 1985. Culture and the Evolutionary Process. Chicago: University of Chicago Press.

———. 2009. Voting With Your Feet: Payoff Biased Migration and the Evolution of Group Beneficial Behavior. Journal of Theoretical Biology 257:331–39.

Carneiro, R. L. 1970. A Theory of the Origin of the State. Science 169:733–38.

———. 1981. The Chiefdom: Precursor of State. In The Transition to Statehood in the New World, ed. G. D. Jones and R. R. Kautz, 37–79. Cambridge: Cambridge University Press.

Currie, T. E., M. Campenni, A. Flitton, T. Njagi, E. Ontiri, C. Perret, and L. Walker. 2021. The Cultural Evolution and Ecology of Institutions. Philosophical Transactions of the Royal Society London B

376:20200047.

Currie, T. E., P. Turchin, J. Bednar, P. J. Richerson, G. Schwesinger, S. Steinmo, R. Wacziarg, and J. Wallis. 2016. Evolution of Institutions and Organizations. In Complexity and Evolution: A New Synthesis for Economics, ed. D. S. Wilson and A. Kirman. Cambridge: MIT Press, Ernst Strüngmann Forum.

Flannery, K. V. 1972. The Cultural Evolution of Civilizations. Annual Review of Ecology and Systematics 3:399–426.

Fletcher, J. A., and M. Doebeli. 2009. A Simple and General Explanation for the Evolution of Altruism. Proceedings of the Royal Society London B, 276.

Flitton A. 2022. Assessing Different Historical Pathways in the Cultural Evolution of Economic Development. Evolution and Human Behavior 43:71–82.

Frey, S., and K. Atkisson. 2020. A Dynamic Over Games Drives Selfish Agents to Win–Win Outcomes. Proceedings of the Royal Society London B 287:20202630.

Fudenberg, D., and J. Tirole. 1992. Game Theory. Cambridge: MIT Press.

Garfield, Z. H., R. L. Hubbard, and E. H. Hagen. 2019. Evolutionary Models of Leadership: Tests and Synthesis. Human Nature 30:23–58.

Garfield, Z. H., K. L. Syme, and E. H. Hagen. 2020. Universal and Variable Leadership Dimensions Across Human Societies. Evolution and Human Behavior 41 (5): 397–414.

Gavrilets, S., D. G. Anderson, and P. Turchin. 2010. Cycling in the Complexity of Early Societies. Cliodynamics: The Journal of Theoretical and Mathematical History 1:5536t55r.

Gavrilets, S., J. Auerbach, and M. van Vugt. 2016. Convergence to Consensus in Heterogeneous Groups and the Emergence of Informal Leadership. Scientific Reports 6:29704.

Gavrilets, S., E. A. Duenez-Guzman, and M. D. Vose. 2008. Dynamics of Coalition Formation and the Egalitarian Revolution. PLOS One 3:e3293.

Gavrilets, S., and M. Duwal Shrestha. 2021. Evolving Institutions for Collective Action by Selective Imitation and Self-interested Design. Evolution and Human Behavior 42:1–11.

Glowacki, L., and C. von Rueden. 2015. Leadership Solves Collective Action Problems in Small-scale Societies. Philosophical Transactions of the Royal Society London B 370:20150010.

Gomes, D. A., and J. Saúde. 2014. Mean Field Games Models - A Brief Survey. Dynamic Games and Applications 4:110–154.

Gorodnichenko, Y., and G. Roland. 2015. Culture, Institutions and Democratization. NBER Working Paper, No. 21117. Issued April 2015, revised May 2020.

Gurven, M., and A. V. Jaeggi. 2013. Food Sharing. In Emerging Trends in the Social and Behavioral Sciences, ed. R. Scott and S. Kosslyn. John Wiley & Sons.

Gu¨th, W., and A. Ockenfels. 2005. The Coevolution of Morality and Legal Institutions: An Indirect Evolutionary Approach. Journal of Institutional Economics 1:155–74.

Hofbauer, J., and K. Sigmund. 1998. Evolutionary Games and Population Dynamics. Cambridge: Cambridge University Press.

Holden, C., R. Sear, and R. Mace. 2003. Daughter-biased Parental Investment Leads to the Evolution of Matriliny. Evolution and Human Behavior 24:99–112.

Hooper, P. L., H. S. Kaplan, and J. L. Boone. 2010. A Theory of Leadership in Human Cooperative Groups. Journal of Theoretical Biology 265 (4): 633–46.

Houle, C., D. J. Ruck, R. A. Bentley, and S. Gavrilets. 2022. Inequality Between Identity Groups and Social Unrest. Royal Society Interface 19:20210725.

Hurwicz, L. 1996. Institutions as Families of Game Forms. Japanese Economic Review 47:113–32.

Isakov, A., and D. Rand. 2012. The Evolution of Coercive Institutional Punishment. Dynamic Games and Applications 2:97–109.

Itao, K., and K. Kaneko. 2020. Evolution of Kinship Structures Driven by Marriage Tie and Competition. Proceedings of the National Academy of Sciences 117 (5): 2378–84.

Jaeggi, A. V., and M. Gurven. 2018. Food-sharing Models. In The International Encyclopedia of Anthropology, ed. H. Callan. John Wiley & Sons.

Janssen, M. A. 2007. Coordination in Irrigation Systems: An Analysis of the Lansing-Kremer Model of Bali. Agricultural Systems 93:170–90.

Johnson, G. A. 1982. Organizational Structure and Scalar Stress. In Theory and Explanation in Archaeology, ed. C. Renfrew, M. Rowlands, and B. A. Segraves, 389–421. New York: Academic Press.

Konrad, K. A., and D. Kovenock. 2009. The Alliance Formation Puzzle and Capacity Constraints. Economic Letters 103:84–86.

Lansing, J. S. 1991. Priests and Programmers: Technologies of Power in the Engineered Landscape of Bali. Princeton, NJ: Princeton University Press.

Lawson, D. J., and N. Oak. 2014. Apparent Strength Conceals Instability in a Model for the Collapse of Historical States. PLOS One 9:e96523.

Leach, E. 1954. Political Systems of Highland Burma: A Study of Kachin Social Structure. Berg.

Mann, M. 1986. The Sources of Social Power. I. A History of Power from the Beginning to A.D. 1760. Cambridge: Cambridge University Press.

Marlowe, F. W. 2004. What Explains Hadza Food Sharing? Research in Economic Anthropology 23:69–88.

McCauley, C., and S. Moskalenko. 2008. Mechanisms of Political Radicalization: Pathways Toward Terrorism. Terrorism and Political Violence 20 (3): 415–33.

North, D. C. 1990. Institutions, Institutional Change and Economic Performance. Cambridge: Cambridge University Press.

Nowak, M. 2006. Evolutionary Dynamics. Cambridge, MA: Harvard University Press.

Olson, M. 1965. Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge, MA: Harvard University Press.

Ostrom, E. 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge: Cambridge University Press.

———. 2009. Understanding Institutional Diversity. Princeton, NJ: Princeton University Press.

Perret, C., E. Hart, and S. T. Powers. 2020. From Disorganized Equality to Efficient Hierarchy: How Group Size Drives the Evolution of Hierarchy in Human Societies. Proceedings of the Royal Society B: Biological Sciences 287:20200693.

Perry, L., M. Duwal Shrestha, M. D. Vose, and S. Gavrilets. 2018. Collective Action Problem in Heterogeneous

Groups with Punishment and Foresight. Journal of Statistical Physics 172:293–312.

Perry, L., and S. Gavrilets. 2020. Foresight in a Game of Leadership. Scientific Reports 10:2251.

Persson, T., and G. Tabellini. 2009. Democratic Capital: The Nexus of Political and Economic Change. American Economic Journal - Macroeconomics 1 (2): 88–126.

Piotrowski, E. W., and J. Sladkowski. 2003. An Invitation to Quantum Game Theory. International Journal of Theoretical Physics 42:1089–99.

Polhill, J. G., T. Filatova, M. Schlüter, and A. Voinov. 2016. Modelling Systemic Change in Coupled Socio-environmental Systems. Environmental Modelling & Software 75:318–32.

Powers, S., and L. Lehmann. 2013. The Co-evolution of Social Institutions, Demography, and Large-scale Human Cooperation. Ecology Letters 16:1356–64.

———. 2014. An Evolutionary Model Explaining the Neolithic Transition from Egalitarianism to Leadership and Despotism. Philosophical Transactions of the Royal Society London B 281:20141349.

Powers, S. T., A. Ekárt, and P. R. Lewis. 2018. Modelling Enduring Institutions: The Complementarity of Evolutionary and Agent-based Approaches. Cognitive Systems Research 52:67–81.

Powers, S. T., C. P. van Schaik, and L. Lehmann. 2016. How Institutions Shaped the Last Major Evolutionary Transition to Large-scale Human Societies. Philosophical Transactions of the Royal Society London B 371:20150098.

Richerson, P., R. Baldini, A. V. Bell, K. Demps, K. Frost, V. Hillis, S. Mathew, E. K. Newton, N. Naar, L. Newson, C. Ross, P. E. Smaldino, T. M. Waring, and M. Zefferman. 2016. Cultural Group Selection Plays an Essential Role in Explaining Human Cooperation: A Sketch of the Evidence. Behavioral and Brain Sciences 39, article number UNSP e30.

Roithmayr, D., A. Isakov, and D. Rand. 2015. Should Law Keep Pace With Society? Relative Update Rates Determine the Co-evolution of Institutional Punishment and Citizen Contributions to Public Goods. Games 6 (2): 124–49.

Roland, G., and Y. Xie. 2016. Culture and Collective Action. In Contemporary Issues in Development Economics, ed. T. Besley, 44–60.

Sandholm, W. H. 2010. Population Games and Evolutionary Dynamics. Cambridge: MIT Press.

Schlüter, M., L. J. Haider, S. J. Lade, E. Lindkvist, R. Martin, K. Orach, N. Wijermans, and C. Folke. 2019. Capturing Emergent Phenomena in Social-ecological systems: An Analytical Framework. Ecology and Society 24.

Servedio, M. R., Y. Brandvain, S. Dhole, C. J. Fitzpatrick, E. E. Goldberg, C. A. Stern, J. V. Cleve, and D. J. Yeh. 2014. Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. PLOS Biology 12:1002017.

Service, E. R. 1978. Classical Theories of the Origin of Government. In Origins of the State. The Anthropology of Political Evolution, ed. R. Cohen and E. R. Service, 21–34. Philadelphia: ISHI.

Singh, M., R. Wrangham, and L. Glowacki. 2017. Self-interest and the Design of Rules. Human Nature 28:457–80.

Siopsis, G., R. Balu, and N. Solmeyer. 2018. Quantum Prisoners' Dilemma Under Enhanced Interrogation. Quantum Information Processing 18: article number 144.

Smaldino, P. E. 2014. The Cultural Evolution of Emergent Group-level Traits. Behavioral and Brain Sciences

37:243–54.

Smith, E. A., and J.-K. Choi. 2007. The Emergence of Inequality in Small-scale Societies: Simple Scenarios and Agent-based Simulations. In The Modelbased Archaeology of Socionatural Systems, ed. T. Kohler and S. van der Leeuw, 105–19. Sante Fe: SAR Press.

Smith, J. E., S. Gavrilets, M. B. Mulder, P. L. Hooper, C. E. Mouden, D. Nettle, C. Hauert, K. Hill, S. Perry, A. E. Pusey, M. van Vugt, and E. A. Smith. 2016. Leadership in Mammalian Societies: Emergence, Distribution, Power, and Payoff. Trends in Ecology and Evolution 31:54–66.

Steponaitis, V. P. 1978. Location Theory and Complex Chiefdoms: A Mississippian Example. In Mississippian Settlement Patterns, ed. B. D. Smith, 417–53. New York: Academic Press.

———. 1981. Settlement Hierarchies and Political Complexity in Nonmarket Societies: The Formative Period of the Valley of Mexico. American Anthropologist 83:320–63.

Stevenson-Hinde, J., and M. Zunz. 1978. Subjective Assessment of Individual Rhesus Monkeys. Primates 19:473–82.

Tembine, H. 2017. Mean-field-type Games. AIMS Mathematics 2:706–35.

Tullock, G. 1980. Efficient Rent Seeking. In Toward a Theory of the Rent-seeking Society, ed. J. M. Buchanan, R. D. Tollison, and G. Tullock, 97–112. College Station: Texas A & M University.

Turchin, P. 2003. Historical Dynamics: Why States Rise and Fall. Princeton, NJ: Princeton University Press.

———. 2013. The Puzzle of Human Ultrasociality: How Did Large-scale Complex Societies Evolve? In Cultural Evolution, ed. P. J. Richerson and M. H. Christiansen, 61–73. Cambridge: MIT Press.

Turchin, P., T. Currie, E. Turner, and S. Gavrilets. 2013. War, Space, and the Evolution of Old World Complex Societies. Proceedings of the National Academy of Sciences USA 109:14069– 74.

Tverskoi, D., A. Senthilnathan, and S. Gavrilets. 2021. The Dynamics of Cooperation, Power, and Inequality in a Group-structured Society. Scientific Reports 18:18670.

Weissing, F., and E. Ostrom. 1991. Irrigation Institutions and the Games Irrigators Play: Rule Enforcement Without Guards. In Game Equilibrium Models II, 188–262. Springer.

Wright, H. T. 1984. Prestate Political Formations. In On the Evolution of Complex Societies. Essays in Honor of Harry Hoijer, ed. W. Sander, H. T. Wright, and R. M. Amams, 41–77. Malibu: Malibu.

Wright, H. T., and G. A. Johnson. 1975. Population, Exchange, and Early State Formation in Southwestern Iran. American Anthropologist 77:267–89.