

Explaining Institutional Change

N. Emrah Aydinonat & Petri Ylikoski

Preprint date: December 2021

Abstract

In this Chapter, we address the challenge of explaining institutional change, asking whether the much-criticized rational choice perspective can contribute to the understanding of institutional change in political science. We discuss the methodological reasons why rational choice institutionalism (RCI) often assumes that institutional change is exogenous and discontinuous. We then identify and explore the possible pathways along which RCI can be extended to be more useful in understanding institutional change in political science. Finally, we reflect on what RCI theorizing would look like if it started to take endogenous change seriously: by giving up some of its simplifying assumptions, RCI can be a useful tool for analyzing institutional change, but choosing this path has consequences for the generality of the models in RCI as well as for the style of its theorizing.

Keywords: institutions, change, explanation, rational choice, rules-in-equilibrium, endogenous change, endogeneity blindness, economics, political science

1. Introduction

Explaining institutional change,¹ particularly explaining gradual and endogenous change, has long been the Achilles heel of institutional analysis. All major approaches to institutions – historical institutionalism, sociological institutionalism, and rational choice institutionalism – seem to have problems with it. These problems are rooted in the way in which institutions are conceived:

Despite many other differences, nearly all definitions of institutions treat them as relatively enduring features of political and social life (rules, norms, procedures) that structure behavior and that cannot be changed easily or instantaneously. The idea of persistence of some kind is virtually built into the very definition of an institution. (Mahoney & Thelen 2009, p. 4)

The inability to deal with institutional change is a serious challenge for *any* theory of institutions. Moreover, a theory that deals with institutional change only in an *ad hoc* manner or rules out conceivable, and possibly important, forms of institutional change cannot be fully satisfactory as a theory of institutions. Such a theory will have an explanatory handicap. Consider, for example, the

¹ What counts as institutional change is ambiguous. While it could mean a change in the rules (or other elements) of the institution, it could also mean changes in the consequences/functions of the institution. These two do not always match. Furthermore, even small changes in the institution could have dramatic long-term consequences, while sometimes large changes in the institution do not make much difference to the consequences of interest. Here, as elsewhere in the social sciences, the principle that the “size” of the causes and effects do not have to match applies. In any case, in this Chapter, we will have to rely on an intuitive idea of institutional change.

distinction between exogenous and endogenous change. Since most institutional change is (and will be) a result of a combination of exogenous and endogenous factors, approaches that rule out some or all forms of endogenous change will have an explanatory disadvantage. Likewise, consider the distinction between discontinuous and incremental change. Although most theories of institutions see change as a discontinuous phenomenon and overlook incremental developments (Mahoney and Thelen 2009), it is also widely accepted that institutions change incrementally (e.g., North 1990). Leaving out the possibility of incremental change creates an explanatory shortcoming for theories of institutions; not only that such a theory cannot get the cases involving incremental change right, but it cannot shed light on the preconditions of abrupt changes it might be able to explain either. Our point is a theoretical one in that we are not claiming that endogenous or incremental sources of change are more important or more powerful, but that an acceptable theory should be able to deal with them both.

In this Chapter, we address the challenge of explaining institutional change and ask whether the much-criticized rational choice perspective can contribute to the understanding of institutional change in political science. We aim to identify the methodological reasons why rational choice institutionalism (RCI) assumes that institutional change is exogenous and discontinuous. We also ask whether it is possible to overcome these biases. More specifically, we will address the question of why RCI has difficulty conceiving endogenous change and why it has “a tendency to see change mostly in terms of dynamics unleashed by some exogenous shift or shock” (Streeck & Thelen 2005, p. 7). Our diagnosis of *the endogeneity blindness* of RCI will address the general limitations of this approach in explaining institutional change, but also reveal some of the ways in which it can contribute to the understanding of institutional change in political science.

The Chapter will proceed as follows. In Section 2, we will begin by describing the core ideas of the rational-choice approach and showing how they give rise to endogeneity blindness. We will focus on Francesco Guala’s *rules-in-equilibrium approach* (RE) (2016), which is a recent attempt to unify various RCI strands. Nevertheless, our points apply more generally. In Section 3, we will discuss three ways of incorporating endogenous change into RCI: Greif and Laitin’s (2004) introduction of quasi-parameters, Jack Knight’s (1992) bargaining approach, and Farrell and Héritier’s (2003) account of endogenous change in the context of EU institutions. This discussion will help us identify some of the possible ways in which RCI can be extended to be more useful in understanding institutional change in political science. In Section 4, we will explore these pathways by opening up the black box of rule compliance a little bit further, discussing the basic reasons why institutional rules are indeterminate and how this gives rise to important sources of incremental institutional change. In Section 5, we conclude by briefly reflecting on what RCI theorizing would look like if it began to take endogenous change seriously. By giving up the simplifying assumptions underlying endogeneity blindness, game theory can still be a useful tool for analyzing institutional change in political science, but choosing this path has consequences for the generality of the models in RCI as well as for the style of its theorizing.

2. The rules-in-equilibrium approach

There are two prominent ways in which RCI scholars have conceptualized institutions: institutions-as-rules (e.g., North 1990, Ostrom 2005) and institutions-as-equilibria (e.g., Lewis

1969, Schotter 1981, Calvert 1995, Greif 2006; see Greif & Kingston 2011 for a review).² The rules approach sees institutions as rules that govern individual behavior and create behavioral regularities. In this approach, institutions are considered as rules that help individuals solve societal problems or facilitate social interaction. The equilibrium approach, on the other hand, defines institutions as the equilibria of games that represent societal problems. It focuses on the motivations of individuals in choosing particular strategies to solve the coordination and cooperation problems that they face, rather than the rules that govern individual behavior. We will focus here on a recent attempt to unify these approaches, called the rules-in-equilibrium approach (RE) (Guala 2016). However, to understand RE, it is useful to start with the equilibrium approach.

	L	R
L	1, 1	0, 0
R	0, 0	1, 1

Figure 1. A Symmetric Coordination Game

The equilibrium approach begins with a representation of a societal problem in game form. The symmetric coordination game in Figure 1 is helpful in illustrating the main characteristics of this approach. In this game, each player has two options, L and R, which could, for example, represent the side of the road they choose to drive on. The payoff matrix shows that players benefit from coordination and thus they have an incentive to choose what the other player chooses. (L, L) and (R, R) are the Nash equilibria of the game, in which no player can benefit from unilaterally choosing another option. The Nash equilibria of this game are taken to represent alternative coordination conventions that are considered as basic institutions. Naturally, the coordination of two players cannot be conceived of as an institution. One can talk about an institution if a population of players somehow end up choosing the same option repeatedly, expecting everyone else to do the same to solve a recurrent coordination problem (Lewis 1969). For the equilibrium approach, the key characteristics of an institution are the following. First, there are multiple equilibria that represent alternative institutions. Second, the existence of an institution requires concordant mutual expectations (beliefs) by the members of the population (i.e., everyone expects – or, believes – that everyone else will act in a certain way). Third, the institution provides a solution to a recurrent societal problem. Fourth, once this solution emerges no one has an incentive to change her behavior.³

The equilibrium approach to institutions is said to provide the micro-foundations for institutions (Weingast 1996, p. 168). Its main strength over the rules approach is its ability to explain why people are motivated to follow the rules (Greif and Kingston 2011, Guala 2016).

² RCI frequently distinguishes between organisations and institutions (e.g., North 1990, Khalil 1995). For example, North (1990) distinguishes between the players of the game (organisations) and rules of the game (institutions). Others in RCI consider organisations as institutional elements, rather than as institutions: “An institution is a system of rules, beliefs, norms, and organizations that together generate a regularity of (social) behavior” (Greif 2006, p. 30). Nevertheless, this distinction does not deny the institutionality of organisations; rather, it serves to direct the analytical focus on the rules that agents (including organisations) follow or on equilibria that represent alternative institutions.

³ On economics of conventions see Sugden (1986, Chp. 3) and Young (1996).

Although an equilibrium model is unable to tell us exactly how they have discovered these rules, it is able to explain why people continue to follow them even in spite of small deviations, and what must happen (how the incentives must change, how the beliefs must be manipulated) to make them stop following the rules. Thus the equilibrium model will support a functional explanation of the form: "the rules exist because they help people solve a coordination problem." (Guala 2016, p. 31)

The main shortcoming of the equilibrium approach is that it describes individual action "from an external point of view" and does not take into account the human "capacity to represent and to follow rules" that "may help converge on an equilibrium" (Guala 2016, p. 54). Guala (2016) follows other scholars (e.g., Aoki 2001, Greif 2006, Greif and Kingston 2011) in combining the main insights of the rules and equilibrium approaches. Basically, Guala's approach (RE) considers *institutions as rules in equilibrium*, where rules are considered as agents' representations of equilibria and the institution as an equilibrium of a coordination game.⁴ In RE, "an institution may be considered as an equilibrium or as a rule of the game, depending on the perspective that one takes" (Guala 2016, p. 50). "From the point of view of an external observer" an institution corresponds to an equilibrium, but from the perspective of an agent an institutional rule "takes the form of a rule that dictates [...] what to do" in a given situation (Guala 2016, p. 50).

By combining the insights of the rules and equilibrium approaches, RE is able to overcome some of the difficulties faced by both accounts. Unlike the rules account, it can explain why people follow some rules and not others and, unlike the equilibrium account, it can take the agents' point of view into account. Nevertheless, because it makes heavy use of the equilibrium concept, it also inherits some of the problems of the equilibrium approach. First, it does not explain the origin of institutions. RE rather presumes that one of the multiple equilibria, that is, one of the alternative institutions, will be selected. Guala presumes that an external public signal such as a toss of a coin (a correlation device, Aumann 1974) announced by a third party (a "choreographer", Gintis 2009) can coordinate a player's behavior. Guala also assumes that the public signals will be in conditional form, such as "if in England, always drive on the left", and help individuals coordinate, but he does not explain where these public signals come from.

Second, like the equilibrium approach, RE is at its best in explaining the *raison d'être* and persistence of institutions *in general*. It shows that the institution at stake is a *solution* to a general societal problem (i.e., problems that individuals repeatedly face in their encounters with others) and once the solution is attained no one has an incentive to deviate from equilibrium behavior. For example, RE provides a *raison d'être* for money by pointing out that it solves a market coordination problem and can explain its persistence by showing that individuals have no incentive to deviate from the equilibrium behavior once the institution is in place.

Third, although RE does not explicitly try to explain institutional change, the way in which it conceptualizes institutions brings with it a certain way of thinking about change, since if institutions are persistent and self-enforcing structures that solve societal problems then institutional change must be due to external shocks.

⁴ We are simplifying here. Guala employs the solution concept of correlated equilibria, which presumes the existence of a correlation device that facilitates coordination, and institutional rules take the form of conditional rules.

A self-enforcing institution is one in which each player's behavior is a best response. The inescapable conclusion is that changes in self-enforcing institutions must have an exogenous origin. (Greif and Laitin 2004, p. 633)

As RE relies heavily on equilibrium thinking it invites thinking about institutional change as a move from one persistent state to another, or from one equilibrium to another, or in terms of punctuated equilibria (Krasner 1984). Notice that this division between periods of endurance (“normal periods”) and exceptional moments of change (“critical junctures”) is set by a methodological bias rather than empirical facts about historical institutional change. Any explanation of change that can be developed within the RE must be developed on a ‘per need’ basis, giving the explanations an *ad hoc* character. RE takes the self-enforcing nature of the resulting game-theoretical equilibrium as a defining feature of institutions, but it has no account of the stability of an institution over time.⁵

How does RE end up with this peculiar way of analyzing institutions? The key factor is the goal of having a general theory that could account for a *type* of institution, rather than for a particular institution.⁶ The *explanandum* is a highly stylized version of a “fully-formed” generic institution, rather than a concrete historical one. This sets RE aside from other approaches, like historical institutionalism, which has always focused on particular institutions. With the generic *explanandum*, all historical details are seen as contingent features of particular situations that are not of theoretical interest. They are just details that must be filled in when the theory is applied to particular cases. What matters is a theoretical how-possible explanation for a type of institution. This disregard for historical context is further strengthened when an *explanandum* is formulated in terms of the *existence* of the institution, rather than in terms of explaining its particular details. Naturally, with an abstract *explanandum* like this, there is much less change to be observed or to be explained. Thus, explaining change feels much less pressing.

Together with the methodological choices of the RE, its highly abstract *explanandum* shapes the characteristics of RE explanations. Although in principle the characteristics, social roles, and the history of particular agents might be represented by a combination of player types and payoffs, these matters are ignored in the analysis for the sake of generality. Similarly, the *agents are usually represented as homogenous*: everybody is assumed to have exactly the same capabilities, resources, beliefs, and incentives. This is not a formal denial of individual differences; they are just not considered theoretically interesting. Moreover, societal problems that are supposed to be the key to understanding institutions are *analyzed as if* they exist *in an institutional vacuum*. However, rational choice models often presuppose, albeit implicitly, the working of certain generic institutions such as markets, property rights, the rule of law, etc.⁷ Also, and more importantly for our purpose, they do not systematically take into account the role of extant institutions in determining institutional outcomes. This is not to say that the importance of existing institutions is denied; rather, in practice they are treated as historically variable contexts that can be ignored in developing a general theory. Finally, RE theorists *take rule compliance for granted*. While in principle it is acknowledged that rules

⁵ One could object to this claim by saying that evolutionary models such as Young's (1998) provide an account of the stability of institutions, but such models account for the stability of institutions at a very general level and are not helpful in explaining the temporal stability of particular institutions.

⁶ For a discussion of Guala's ideas about institution types and tokens, see Aydinonat & Ylikoski (2018).

⁷ See Field (1979, 1981) for a critique of rational choice models of institutions from this point of view.

can be ambiguous, indeterminate, and open to conflicting interpretation, in practice these issues are treated as random disturbances that do not have long-term consequences.

For RE, these *methodological choices and simplifications* are small sacrifices that allow the development of truly general theory. However, in explaining institutional change, as we just argued, these simplifications are far from harmless, making RE incapable of explaining endogenous change. Only by giving up these simplifications and the associated ideas about abstract *explananda* can RCI overcome its current shortcomings in explaining institutional change. The following sections will demonstrate this by showing how giving up these simplifications one by one increases the theory's ability to explain change.

3. Overcoming endogeneity blindness

We have seen that endogeneity blindness is a product of simplifying assumptions that are ultimately justified by the choice of a highly abstract *explanandum*. In what follows we will analyze the resources available to RCI to overcome its endogeneity blindness, based on attempts by some rational-choice theorists to incorporate endogenous change into their theories. We will argue that overcoming endogeneity blindness requires giving up the simplifying assumptions described above. By paying attention to the details ignored by mainstream RCI it is possible to increase its explanatory power with respect to real world institutional phenomena.

We will start with a discussion of Greif and Laitin's (2004) prominent attempt to analyze endogenous change in terms of quasi-parameters. We will then briefly introduce Knight's (1992) bargaining approach that can increase the leverage of RCI. Finally, we will show how Farrell and Héritier (2003) employ the Knightian approach to develop an interesting model of endogenous change.⁸

3.1. Greif and Laitin: quasi-parameters

The basic idea of the Greif-Laitin approach is that an institution influences many things in the society, "such as wealth, identity, ability, knowledge, beliefs, residential distribution, and occupational specialization" (Greif and Laitin, 2004, p. 636) that could change or even undermine the institution in the long run. Normally, most things that institutions can influence are considered as exogenous parameters in the analysis. Greif and Laitin suggest considering them as endogenous when investigating institutional change. They call these parameters *quasi-parameters*, as their status as an exogenous parameter depends on the explanatory question asked. In the Greif-Laitin approach, an institution can cause a change in quasi-parameters, which can in turn either undermine or reinforce the equilibrium underlying that very institution. When more individuals in a wider range of situations find it best to adhere to the institutional rules – i.e., when the equilibrium is self-enforcing in a wider range of parameters – the institution is considered as *self-reinforcing* (Greif and Laitin, 2004, p. 634). When fewer individuals in fewer situations find the rules compelling, the institution is considered to be *self-undermining*. While Greif and Laitin work with the same idea of equilibrium as Guala, the idea of quasi-parameters allows them to conceive an institution as having

⁸ Here, we are focusing on a few selected approaches that employ the rational choice framework. For a comparative review of theories of institutional change, see Kingston & Caballero (2009).

consequences that have a “feedback effect” on itself, thus expanding or reducing the set of situations in which the institution is self-enforcing.

The idea of a quasi-parameter is ingenious. First, it allows one to see the difference between *game-theoretical equilibrium* as a representation of an institution and the *stability of the institution*. For comparison, consider RE. Although the RE approach distinguishes between the theoretical representation of the institution and the institution itself (Guala 2016), it does not operationalize this distinction. In RE analysis, equilibrium equals institutional stability or persistence. Greif and Laitin, on the other hand, identify the proper *explanandum* both for institutional persistence and change with the help of quasi-parameters. Second, in Greif and Laitin’s framework, similar causal mechanisms could explain both stability and change. There is no longer a need to assume that stability and change require different kinds of explanation, as suggested by Guala. Finally, the Greif-Laitin approach allows for incremental change, or at least incremental build-up for a change, since the institutional change occurs “when the self-undermining process reaches a critical level such that past patterns of behavior are no longer self-enforcing” (Greif and Laitin, 2004, p. 634).

Despite these appealing features, the Greif-Laitin approach does not fully address the challenge of explaining endogenous change. It does describe, very abstractly, how endogenous change could happen, but it does not really go very far: the quasi-parameters are just placeholders for whatever indirect institutional consequences might influence the set of situations in which the institution is self-enforcing. The model itself does not provide any insight as to which parameters should be conceived as quasi-parameters *ex ante*. Quasi-parameters must be decided case by case, “based on empirical observations” (Greif and Laitin, 2004, p. 634). The introduction of quasi-parameters helps to save the idealizing assumptions of the game-theoretical approach while bringing in more flexibility for explaining particular cases of institutional change. The crucial question, however, is whether the model actually contributes to explanations of particular cases of institutional change. For example, in Greif and Laitin’s (2004, Greif 2006) historical case studies, the formal models only seem to serve as non-essential parts of the explanatory narratives and they can be discarded without losing much of the explanatory insight as the real explanatory work is done by the context and historical details.

3.2. Knight: institutions as unintended consequences of conflict

Social scientific common sense suggests that power asymmetries between agents are important determinants of institutional change. While this insight often motivates empirical RCI analyses, it is not well reflected in theoretical work. As we have seen, the institutional analysis is done in an institutional vacuum without taking the extant institutions into account and assuming that the agents are homogenous. However, the influence of existing institutions on resource allocation and bargaining power is crucial for understanding why the institution has the rules it has.

In *Institutions and Social Conflict* (1992), Jack Knight presents a version of RCI that takes power asymmetries between agents as its analytical basis. While Knight does not offer new ideas about endogenous change, his approach introduces ideas that are useful for analyzing endogenous processes. He argues that understanding conflict over the expected benefits from alternative institutions provides a key to understanding institutions. His account shows how distributional consequences of institutions and conflict over institutional benefits can be integrated into an equilibrium approach to institutions.

According to Knight,

the main goal of those who develop institutional rules is to gain strategic advantage vis-a-vis other actors, and therefore, the substantive content of those rules should generally reflect distributional concerns. (Knight 1992, p. 40).

Rather than assuming that institutions are efficient in facilitating coordination or cooperation, Knight suggests regarding them as by-products of substantive conflicts over distributional outcomes. This makes the power asymmetries between the agents analytically salient. The basic idea is that in a bargaining situation an individual who has more power (e.g., in terms of access to resources such as wealth or holding an influential institutional position) usually has an advantage over individuals who lack power because a lengthy bargaining process would be more costly and risky for the latter. Those who have resources can wait, while those who do not will feel pressure to settle the situation irrespective of the distributional consequences because they cannot afford to wait for a resolution or take the risk of failure. In Knight's analysis, bargaining power acts as an equilibrium selection "device", selecting the equilibrium that favors the powerful. Note that the same considerations are also important for understanding institutional change: for example, if the formal institutional rules are ambiguous or in need of revision, the power asymmetries can determine the course of institutional change. In Knight's account, the preferences and resources of the powerful do not directly determine the content of the rules, but the focus on the asymmetries in power and the interests of the agents are important points of analytical focus.

Knight's approach is not formally in conflict with RE. While RE focuses on the common interests of the parties, Knight employs an analytical perspective that focuses on the heterogeneity of the agents and asymmetries of power. As Knight argues, the emphasis

... on the theoretical primacy of distribution does not deny the importance of gains from coordination or trade. The main point here is that such gains cannot serve as the basis for a social explanation; rather, these benefits are merely a by-product of the pursuit of individual gain (Knight 1992, p. 38).

Those who have more resources would have more leverage and hence an ability to influence the distributional consequences of the resulting institution. Furthermore, as Knight recognizes that the resources available to the participants and the rules of the bargaining are products of already existing institutions, it is much easier to avoid the temptations of the state of nature theorizing.

One of the advantages of this approach is that it allows more powerful explanations in that it can provide more general theoretical explanations as well as more precise (Kuorikoski & Ylikoski 2010) explanations for particular cases. Rather than focusing on the explanation of the existence of an institution, it addresses details of the institution: why does it have the rules it has? The explanatory contrast is not the absence of an institution, but an institution with different characteristics such as rules and distributional consequences. This is an appealing feature when explaining some real-world institutions rather than wondering about the existence of institutions in general. Another advantage is the clear distinction between the sources of an institution – i.e., the struggles over how benefits are distributed – and its consequences (Farrell 2009: 18), which makes the approach much less prone to flirting with functionalism. Rather, it makes systematic theorizing in terms of social mechanisms more salient (Knight 1992; Hedström & Ylikoski 2010). Overall, the emphasis on the differential capacities of agents and the unintended consequences of institutions helps to

connect game-theoretical theorizing with parallel research done by historical and sociological institutionalists.

Nevertheless, Knight (1992) does not provide a full toolset for analyzing endogenous change. A closer look shows that his approach makes it difficult to explain endogenous change because of the great emphasis on the endurance of institutions.⁹ Moreover, Knight follows the common RCI practice of taking rule-compliance for granted. As a result, he downplays the role and extent of institutional change. Nonetheless, Knight provides an important perspective change for RCI analysis of institutions and paves the way to models about endogenous change. An example of such a model is presented by Henry Farrell and Adrienne Héritier (Farrell and Héritier 2003, 2007; Héritier 2007), which we discuss next.

3.3. Farrell and Héritier

Farrell and Héritier (2003) add to Knight's approach the assumption that formal institutional rules are incomplete contracts. They reject the simplification that all institutional features are fixed once the institution is in place. Instead, they assume that the formal rules are subject to renegotiation in the course of their daily application. In particular, they assume that informal rules govern the application of the formal rules. In the context of EU institutions, they claim that Treaties established in EU Intergovernmental Conferences are incomplete contracts that leave room for *bargaining* among the main organizational actors within the EU legislative process: The European Parliament, the Council, and the Commission. In the RCI analysis of Farrell and Héritier, these actors seek to maximize their own legislative competence by bargaining over how ambiguities in the treaty texts ought to be interpreted and applied to the legislative process. The basic idea is that bargaining over the appropriate interpretation of ambiguous formal institutional rules has distributional consequences and that informal institutions will influence this bargaining process.

A crucial element of this model is the distinction between agents who design the formal rules and other agents: those who are affected by the formal rules and those who are involved in the implementation of the formal rules. Not all agents can participate in the negotiation about the formal rules, but they may be involved in the implementation of the formal rules and influence their interpretation, and possibly, the future reformulations of the formal rules. For example, the EU member states have control of the treaty process, but the Parliament, the Council, and the Commission may take over when the treaties are implemented. This may involve the creation of informal institutions that structure actors' relations within the legislative process. These informal institutions may accommodate existing formal rules by specifying or complementing them, but they may also transform them in a way that shifts the institutional structure and its distributional consequences (Farrell and Héritier 2003, 2007; Héritier 2007).

Farrell and Héritier analyze the situation in RCI terms. The informal rules that guide the daily application of ambiguous formal rules are products of bargaining. Two main sources of bargaining power are the agent's formal institutional position (determined by extant institutional rules) and resource availability, which determine what time horizon the agent has and how sensitive it is to

⁹ Knight is aware of the difficulty of explaining endogenous change, especially in the case of spontaneous change: "... institutional change is not easily accomplished" (p. 145), "requires repeating the emergence process" and is "made even more difficult by the fact that unlike the case of initial emergence when expectations are not fixed on a particular rule, change requires persuading social actors to alter an expectation that is now fixed" (p. 147).

failure. These factors influence the ability of agents to shape the outcome of the negotiation and thus the informal rules governing the application of the formal rule. This constitutes the first phase of institutional change since the formal institutional rules might change as the informal institutions that guide their interpretation change.

Since the second phase consists of the interaction between informal and formal institutions, the informal rules *recursively* affect the next bargaining round in the revision of formal rules. The designing actors (such as EU member states) have to take into consideration the *de facto* rule change created and this will influence how formal rules are reformulated (e.g., how EU member states formulate the next intergovernmental treaty). More generally, if the designing agents find the results of the “interstitial” institutional change beneficial, they might formalize these *de facto* rules. In other situations, they might try to abolish the interstitially bargained informal rules or close the observed loopholes. However, this might not be possible when they disagree, and the results of the interstitial bargaining will remain. In any case, the *de facto* informal rules are something that designing agents have to take into account in negotiating the new formal rules. (Farrell and Héritier 2003, 2007; Héritier 2007; see also Jupille 2004 for an account of procedural politics that involves a similar idea about endogenous change.)

In the Farrell- Héritier model, institutional change is not only partly endogenous but also continuous. It is not analyzed in terms of off-and-on formal redesigns driven by independently given preferences, but a more continuous process of bargaining over informal practices that have feedback on future negotiations about the formal rules. The model is based on the idea of a feedback loop between the formal and informal rules that guide their application. Because of the heterogeneity of agents and the role of existing institutions, the institutional outcomes are under constant negotiation and never simply reflect the interests of the participants.

This ingenious model is enabled by giving up some typical RCI simplifications about institutions. First, the Farrell- Héritier model inherits all the benefits of Knight’s approach. The institutions are analyzed from the beginning in terms of conflicts about distributional consequences rather than cozy coordination arrangements between equal parties. However, they go further and give up some other RCI simplifications. The second important move is to pay more attention to the differences between agents. Not only are there power differences between agents, but they also have different roles: not all agents participate in the negotiations about formal rules, but they might still influence the way they are applied. This makes formal game-theoretical modeling of the situation much harder in that we have not only the game of setting the rules, but also the game of interpreting the rules, and the interactions between the two. Furthermore, not only are there multiple interrelated games, but these games also have different players, at least in part.¹⁰

Third, the Farrell- Héritier model does not take rules and rule compliance for granted. The key assumption is that the agreement about formal rules does not uniquely determine how the rules are to be applied or implemented. This idea about “incomplete contracts” opens the door to a continuous process of negotiation about the application of the rules which can give rise to incremental institutional change. Although the Farrell- Héritier model operates with a crude distinction between formal and institutional rules, their model suggests that unpacking the black

¹⁰ See Ross (2008) on how the “game determination” problem relates to the game theoretical accounts of conventions and convention formation.

box of rule compliance makes it possible to open up room for new forms of institutional change. In the next section, we will open this box a little bit further to see how it affects RCI assumptions and how it may enable seeing new forms of endogenous change. To this end, we will highlight some further aspects of rule-following. This will show that paying more attention to how the institutions – and rules – work paves the way for more sophisticated accounts of institutional change (and stability).

4. Taking rule-following seriously

Most RCI models assume homogeneous agents and focus on simple institutions, such as coordination conventions, in an institutional vacuum. Although these models are useful for certain explanatory tasks, they fall short of explaining institutional change. They also have limited use in analyzing the complex institutions that political scientists are interested in. In what follows, we will explore how RCI could be extended to be useful for explaining change in political institutions. So far, our discussion has shown that relaxing the typical RCI assumptions such as the homogeneity of agents, the absence of pre-existing institutions, and unproblematic rule compliance enables a better understanding of institutional change. Following the path pioneered by Farrell and Héritier, let's discuss what taking into account power asymmetries, diverse institutional roles, rule ambiguity and rule compliance might mean for institutional theorizing.

RCI models typically do not take into account the fact that the agents may have diverse institutional and social roles. The Farrell- Héritier model suggests that the differing roles of agents require more attention. Individuals may have roles as rule makers, rule enforcers, or only as rule takers (Mahoney & Thelen 2010). Some individuals might also take multiple roles. Once the importance and diversity of roles is accepted, attention to historical dynamics becomes more important because the roles of individuals, as well as the roles themselves, might change in time. Moreover, again as Farrell and Héritier suggest, because individuals in various roles might rank institutional outcomes differently it would not be possible to understand the bargaining processes that bring about institutional change if one ignores this. For example, many agents whose compliance is crucial for the institution are often “forced” into obedience by the expectation of punishment or the current unfeasibility of more favorable institutional rules (Moe 2005). The “losers” will try to improve their situation when opportunities arise, and the “winners” will counter these moves to continue to utilize their position to their advantage (Pierson 2016). Looking at how individuals holding different positions bargain over institutional outcomes opens up the internal dynamics of the institution which are important for many kinds of endogenous change.

	Typical RCI assumptions	Extended RCI
Agents	Homogenous No institutional/social roles No asymmetries in power (e.g., with respect to access to resources)	Heterogeneous Diverse institutional/social roles Power asymmetries, in terms of a. Resources b. Roles (rule makers vs. rule takers)
Institutional environment	No pre-existing institutions	Existing institutions and practices determine roles, power asymmetries, etc. a. Interaction between formal and informal rules/practices b. Conflict between rules
Rules	No ambiguity, no room for interpretation No bargaining over meaning Universal rule compliance	Ambiguous rules, open to interpretation Bargaining over meaning / interpretation No universal rule compliance

Table 1. *RCI and its possible extensions for political science.* The table shows how RCI can be extended to be useful for understanding institutional change in political science (Extended RCI) by relaxing some of the typical assumptions in RCI.

Second, the interaction between formal and informal rules requires more attention. Both formal and informal rules function against the background of a complex set of practices. Focusing on the formal rules, one can easily see that the application of any formalized rule is built upon an implicit set of practices that are not part of the formalized rule itself. It should also be noted that calling these practices “informal rules” and acknowledging their importance is not sufficient to understand how they operate and interact (cf. Turner 1994). A proper analysis of how rules work *as a part of these practices* is essential for understanding institutional change because the change in these practices can change the meaning and consequences of the formal rules. Paying attention to how rules work also opens the possibility of analyzing the role of material technologies underlying institutions. This is going to be an increasingly important dimension of institutions in the future.

Third, the indeterminacy of institutional rules – the foundation for the idea of incomplete contracts in Farrell and Héritier’s analysis – is another key element of institutional change. Taking this into account, however, will pose a challenge for the rational-choice design perspective. It is not realistic to assume that the agents who are setting the rules for an institution can fully anticipate all future scenarios where the rules would be put into use. There are at least two reasons for this. (1) Taking into account a huge number of imaginable future situations would be very costly and time-consuming. (2) There is no reason to assume that any agent can imagine all possible current and future scenarios where the rules might be put to use. Thus, even in the best circumstances, formal rules have to be considered as “incomplete contracts”. There is also a strategic aspect of the ambiguity of the rules in that many formal rules are compromises that are *only possible because of the ambiguity or imprecision of the rules*. Removing ambiguity is often costly and time-consuming, and hence might not be possible or preferable. Moreover, for some agents, it could be strategically advantageous to leave the rules abstract and ambiguous as this might give them some advantage

in later negotiations about their interpretation. As Thelen and Conran put it: “... institution building is almost always a matter of political compromise. Institutions and rules are often left deeply ambiguous by the coalitions of (often conflicting) interests that preside over their founding” (2016, p. 57).

However, the indeterminacy and ambiguity of rules do not imply that we should set aside the basic insights of RCI. As Thelen and Conran observe: “... institutions instantiate power, they are contested. Losers in one round do not necessarily disappear but rather survive and find ways not just to circumvent and subvert rules, but to occupy and redeploy institutions not of their own making” (2016, p. 57). The rational-choice perspective is useful in understanding this process, but only if the existence of an institution is regarded as a continuous process that involves ambiguities and negotiations between heterogeneous agents with diverse roles and interests. This implies that change and stability are two sides of the same coin, rather than things that require completely different explanations. The life of institutions does not consist of serene equilibriums that are sometimes disrupted by external shocks, but of an ongoing struggle with coalitional dynamics that are vulnerable to smaller and larger shifts.¹¹

Fourth, the conflicts between rules of different institutions as well as rule conflicts within institutions can be important sources of institutional change. Rules may overlap and conflict with each other. As such conflict requires a resolution, some of these rules either need to be changed or given a new interpretation, at least with respect to their domain of application. Either way, these conflicts can be important sources of institutional change. The conflicts between institutions are especially interesting because they might expand the set of agents that have a stake in the institution and resources to influence it. This could also mean a change in the definition of the “social problem” that the institution addresses. This point about ambiguity and possible overlap of institutional boundaries is general, since the institutional rules of any given society cannot be conceived as a coherent and fixed bundle of rules as envisioned by many RCI accounts. The real-world societies live with a hodgepodge of complex rules and institutions of various origins and purposes.¹² The conflicts between their rules are an important source of institutional change. However, the typical methodological practice of analyzing institutions in isolation – or even in a vacuum – makes it hard for RCI to recognize this source of endogenous change.

Finally, rule cognition – the way in which people learn and apply the rules – demands more attention. This is most clearly seen in the case of informal rules. As the agents learn the rules from a limited set of observed examples and even more limited cases of positive and negative feedback (sanctions), there is no guarantee that they will learn exactly the same rules and that they will

¹¹ There is a further dimension of rule-following that we will not discuss in detail in this Chapter. A typical RE discussion only sees two possibilities: following the rule or not following it. In real social life, the agents usually have more options. Not following the rule can have multiple meanings: it might mean breaking the rule in a specific situation (while generally following the rule), or it might mean rejection, that is, rejecting its legitimacy as a policy. Sometimes there is also the third option of avoiding the rule. Finally, and most importantly, even following the rule might be guided by strategic employment of its ambiguity in a manner that an outsider might call biased.

¹² There is surprisingly little discussion about the demarcation between individual institutions in the literature. It is easy to agree with North’s (1990, p. 3) definition: “Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction”, but this definition talks about a complete set of institutions rather than an individual institution. Philosophical debates in social ontology suffer from the same problem. From the ontological perspective, this issue might seem secondary, but it is a real issue when one is attempting to explain real world institutions.

interpret these rules in exactly the same way. While there might be a broad agreement about typical applications of a rule, agents' interpretations of the rule might diverge with respect to peripheral cases and novel situations. That is, diverse interpretations of a rule can co-exist. This variance in learned rules can be a source of future incremental changes in meaning, and hence a source of institutional change. This diversity might only become apparent with interpretative conflict among individuals or organizations. Biological speciation and the origins of languages and dialects are good models for thinking about change like this because small populations, local interactions, and variation-creating learning mechanisms can give rise to significant long-term change both within and between populations.¹³ Change could in fact be a random-walk process that does not have uniform direction or generic explanation, but this does not imply that it cannot be explained. While highly abstract RE models cannot capture *institutional drift* like this¹⁴ it is clear that these endogenous processes can still be analyzed in terms of causal mechanisms that incorporate some elements of RCI.

Taken together, these observations point to a rich set of conceptual resources for analyzing endogenous change, and institutional change more generally. Although our discussion has been brief, we hope that it shows that one does not have to give up RCI completely to accommodate these sources of institutional change.

5. Conclusion

The starting point of this Chapter was that any acceptable theory of institutions should be able to handle institutional change. We argued that RCI, particularly RE, suffers from *endogeneity blindness*, which is caused by the excessively abstract *explananda* and simplifying assumptions that accompany it. These assumptions lead to missed explanatory opportunities as they push endogenous change out of the reach of the theory. We have seen that RCI typically aims to provide generic explanations of the existence or persistence of an institution, rather than its attributes. It addresses highly stylized descriptions of types of institutions, rather than the historically variable attributes of particular institutions. The underlying belief is that addressing highly abstract *explananda* gives the theory generality, and thus explanatory power. However, the generality of this approach does not necessarily imply general explanatory power as there are very few real-world cases where the theory would apply directly, and it has given up the most important heuristic ideas that would help it adapt to concrete historical institutions. What in fact happens is that rational-choice theorists end up pursuing functionalist explanations that can *only* address the persistence or *raison d'être* of institutions. Highly limited *explananda* and incongruence with the causal mechanisms pursued by other social scientists limit the explanatory value of most RCI models.

Although it is customary in RCI to distinguish between explanations of the origins and stability of institutions (e.g., Greif 2006, Guala 2016), there is a sense in which this distinction is misleading.

¹³ These points have been taken seriously by credible cultural evolution theories (e.g., Boyd and Richerson 2005).

¹⁴ Acemoglu and Robinson have something similar in mind when they write: "In the same way that the genes of two isolated populations of organisms will drift apart slowly because of random mutations in the so-called process of evolutionary or genetic drift, two otherwise similar societies will also drift apart institutionally—albeit, again, slowly. Conflict over income and power, and indirectly over institutions, is a constant in all societies. This conflict often has a contingent outcome, even if the playing field over which it transpires is not level. The outcome of this conflict leads to institutional drift. But this is not necessarily a cumulative process." (2012, p. 431)

While the emergence of new institutions is quite rare, the change is a constant feature of most institutions. Hence, it needs to be understood. Furthermore, these three *explananda* (i.e., origin, stability, and change) are not completely independent. For example, the configuration of causes that gave rise to an institution need not be the same as one that sustains it, or drives its change, but the two processes might still have interesting common elements. Of course, while theories and models that address only a subset of institutional *explananda* are not useless, this should not prevent us from acknowledging their limitations. In our case, it must be acknowledged that systematically ignoring change-related *explananda* creates such limitations for a theory of institutions. It is a serious hindrance for a theoretical approach to rule endogenous change out on purely conceptual grounds or otherwise to make it invisible. Moreover, it is possible that the perceived stability of institutions in RCI is an optical illusion created by too abstract a perspective. If this is the case, RE would be even further from capturing how institutions really work.

Our discussion has focused mainly on endogenous change, but our main points apply more generally. Endogenous change is theoretically intriguing, but all institutional change demands attention. We do not claim that endogenous processes are the most important cause of institutional change. This is clearly an empirical question, and most interesting cases of change are probably combinations of both endogenous and exogenous processes (Koning 2016). The point is that it is difficult to justify the explanatory handicaps of a theory that rules out some possible forms of institutional change – especially on non-empirical grounds. Because understanding change is a key element of understanding how institutions stay in existence (or vanish), a general theory of institutions that ignores change is not credible.

Acknowledging the limitations of typical RCI models is also helpful in seeing how to move forward in analyzing the more complex institutions that interest political scientists. One of our main claims is that giving up some typical RCI methodological simplifications makes explanatory progress – understood as an ability to explain how real social institutions work – possible. These simplifications are mostly justified by abstract *explananda*. If they are replaced by more concrete *explananda* that political scientists and other social scientists address, the pathway we have sketched is fully open. RCI can still generate important insights about social institutions. Although RCI – no matter how it is interpreted – is never going to be sufficient for a comprehensive understanding of the workings of institutions, it can be a powerful part of the toolbox, as recent debates show (Mahoney & Thelen 2009; Hall 2009).¹⁵

It must be realized, however, that giving up the simplifications of RCI to gain an explanatory advantage in explaining institutional change has consequences for the style of theorizing. First, to explain change, institutions must be considered as historical individuals, rather than tokens of some abstract types. Institutions are not natural kinds (pace Guala 2016). There is general knowledge to be had about institutions, but it does not arise from starting with highly abstract *explananda*. General knowledge will not be about functionally defined institution types, but about social mechanisms that are parts of institutions (Aydinonat & Ylikoski 2018). This understanding is often the result of painstaking analysis of particular institutions. Thus, “History matters” (North 1990, p. vii). And history here means more than just initial conditions for a general model. An

¹⁵ See Ylikoski and Aydinonat (2014) for our account of the epistemic import of highly idealized models. See Aydinonat (2008) for a discussion of the value and limits of highly idealized game-theoretical models of institutions.

institution is a historically changing entity and we have to understand the dynamics of this change. When institutions are conceived this way, there is no reason to assume that explaining their stability or persistence has any explanatory priority. Explaining persistence is important, but so is accounting for the changes and origin.

Second, the historical nature of the study of institutions has consequences for the role of elaborate formal models. While they will remain useful for more limited purposes, their role is now less central. They are tools for understanding the “logics of situation” the agents encounter, not direct explanatory representations. This is not a completely new idea. For example, when Thomas Schelling was arguing that the highly idealized game-theoretical models that are typical in RCI should not be considered as “instant theory” that could be directly applied to real-world cases, but as a framework for analysis (Schelling 1984, p. 241),¹⁶ he was making a similar point. What is needed is the development of ways to combine historically sensitive analysis of institutions with insights provided by the formal models. As we hope to have shown in this Chapter, it matters what you model: to fully understand how institutions work, we have to address their historical nature – how they originate, how they change, and how they end – not just some theoretically salient aspect of them.

6. References

- Acemoglu, D. & J. Robinson (2012). *Why Nations Fail: The origins of power, prosperity and poverty*. London: Profile Books.
- Aoki, M. (2001). *Toward a Comparative Institutional Analysis* (Vol. 1). Cambridge: MIT Press.
- Aoki, M. (2007). Endogenizing institutions and institutional changes. *Journal of Institutional Economics*, 3(1), 1. <https://doi.org/10.1017/S1744137406000531>
- Aoki, M. (2008). Analysing Institutional Change: Integrating Endogenous and Exogenous Views. In J. Kornai, L. Mátyás, & G. Roland (Eds.), *Institutional Change and Economic Behaviour* (pp. 113–133). London: Palgrave Macmillan UK. <https://doi.org/10.1057/97802305834296>
- Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1), 67–96. [https://doi.org/10.1016/0304-4068\(74\)90037-8](https://doi.org/10.1016/0304-4068(74)90037-8)
- Aydinonat, N. E. (2008). *The Invisible Hand in Economics: How Economists Explain Unintended Social Consequences*. London: Routledge.
- Aydinonat, N. E., & Ylikoski, P. (2018). Three Conceptions of a Theory of Institutions. *Philosophy of the Social Sciences*, 48(6), 550–568. <https://doi.org/10.1177/0048393118798619>
- Boyd, R., & Richerson, P. (2005). *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.
- Brousseau, E., Garrouste, P., & Raynaud, E. (2011). Institutional changes: Alternative theories and consequences for institutional design. *Journal of Economic Behavior & Organization*, 79(1–2), 3–19. <https://doi.org/10.1016/j.jebo.2011.01.024>

¹⁶ See Aydinonat (2008, Chp. 8) for a more extensive discussion of these issues.

- Calvert, R. R. (1995). Rational actors, equilibrium and social institutions. In J. Knight & I. Sened (Eds.), *Explaining Social Institutions* (pp. 57–94). Ann Arbor, MI: University of Michigan Press.
- DellaPosta, D., Nee, V., & Opper, S. (2017). Endogenous dynamics of institutional change. *Rationality and Society*, 29(1), 5–48. <https://doi.org/10.1177/1043463116633147>
- Farrell, H., (2009). *The Political Economy of Trust*. Cambridge: Cambridge University Press.
- Farrell, H., & Héritier, A. (2003). Formal and Informal Institutions Under Codecision: Continuous Constitution-Building in Europe. *Governance*, 16(4), 577–600. <https://doi.org/10.1111/1468-0491.00229>
- Farrell, H., & Héritier, A. (2004). Interorganizational Negotiation and Intraorganizational Power in Shared Decision Making: Early Agreements Under Codecision and Their Impact on the European Parliament and Council. *Comparative Political Studies*, 37(10), 1184–1212. <https://doi.org/10.1177/0010414004269833>
- Farrell, H., & Héritier, A. (2007). Codecision and institutional change. *West European Politics*, 30(2), 285–300. <https://doi.org/10.1080/01402380701239723>
- Field, A. J. (1979). On the Explanation of Rules Using Rational Choice Models. *Journal of Economic Issues*, 13(1), 49–72. <https://doi.org/10.1080/00213624.1979.11503610>
- Field, A. J. (1981). The problem with neoclassical institutional economics: A critique with special reference to the North/Thomas model of pre-1500 Europe. *Explorations in Economic History*, 18(2), 174–198. [https://doi.org/10.1016/0014-4983\(81\)90025-5](https://doi.org/10.1016/0014-4983(81)90025-5)
- Giddens, A. (1984). *The Constitution of Society*. Cambridge: Polity Press.
- Gintis, Herbert (2009) *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton: Princeton University Press.
- Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge: Cambridge University Press.
- Greif, A., & Laitin, D. D. (2004). A Theory of Endogenous Institutional Change. *American Political Science Review*, 98(4), 633–652. <https://doi.org/10.1017/S0003055404041395>
- Greif, A., & Kingston, C. (2011). Institutions: Rules or Equilibria? In A. Greif & C. Kingston (Eds.), *Political Economy of Institutions, Democracy and Voting* (pp. 13–43). https://doi.org/10.1007/978-3-642-19519-8_2
- Guala, F. (2016). *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton: Princeton University Press.
- Hall, P. A. (2009). Historical Institutionalism in Rationalist and Sociological Perspective In J. Mahoney & K. Thelen (Eds.), *Explaining Institutional Change* (pp. 204–223). Cambridge: Cambridge University Press.
- Héritier, A. (2007). *Explaining institutional change in Europe*. Oxford; New York: Oxford University Press.
- Jupille, J. (2004). *Procedural Politics*. Cambridge: Cambridge University Press.
- Khalil, E. L. (1995). Organizations Versus Institutions. *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft*, 151(3), 445–466.
- Kingston, C., & Caballero, G. (2009). Comparing theories of institutional change. *Journal of Institutional Economics*, 5(2), 151–180. <https://doi.org/10.1017/S1744137409001283>
- Knight, J. (1992). *Institutions and Social Conflict*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511528170>

- Koning, E. A. (2016). The three institutionalisms and institutional dynamics: Understanding endogenous and exogenous change. *Journal of Public Policy*, 36(4), 639–664. <https://doi.org/10.1017/S0143814X15000240>
- Krasner, S. D. (1984). Approaches to the State: Alternative Conceptions and Historical Dynamics. *Comparative Politics* 16(2), 223–246. <https://doi.org/10.2307/421608>
- Lewis, D. (1969). *Convention*. Harvard: Harvard University Press.
- Mahoney, J., & Thelen, K. (2009). A Theory of Gradual Institutional Change. In J. Mahoney & K. Thelen (Eds.), *Explaining Institutional Change* (pp. 1–37). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511806414.003>
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton: Princeton University Press.
- Pierson, P. (2016). Power in Historical Institutionalism. In O. Fioretos, T. G. Falleti, & A. Sheingate (Eds.), *The Oxford Handbook of Historical Institutionalism* (pp. 124–141). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199662814.013.7>
- Rixen, T., & Viola, L. A. (2015). Putting path dependence in its place: Toward a Taxonomy of institutional change. *Journal of Theoretical Politics*, 27(2), 301–323. <https://doi.org/10.1177/0951629814531667>
- Ross, Don. (2008). Classical Game Theory, Socialization and the Rationalization of Conventions. *Topoi* 27: 57-72. <http://doi.or/10.1007/s11245-008-9028-1>
- Schelling, T. C. (1984). *Choice and Consequence*. Cambridge, Mass.: Harvard University Press.
- Schotter, A. (1981). *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.
- Streeck, W., & Thelen, K. (2005). Introduction: Institutional Change in Advanced Political Economies. In W. Streeck & K. Thelen (Eds.), *Beyond Continuity: Institutional Change in Advanced Political Economies* (pp. 1–39). Oxford: Oxford University Press.
- Sugden, R. (1986). The economics of rights, co-operation and welfare. Oxford: Basil Blackwell.
- Turner, S. P. (1994). *The Social Theory of Practices*. Cambridge: Polity Press.
- Weingast, B. R. (2002). Rational-Choice Institutionalism. In I. Katznelson & H. V. Milner (Eds.), *Political Science: The State of the Discipline* (pp. 660–692). New York: W.W. Norton & Company.
- Young, H. P. (1996) Economics of convention. *Journal of Economic Perspectives*, 10 (2): 105–122.
- Young, H. P. (1998). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton: Princeton University Press.