6

EQUILIBRIUM ANALYSIS OF POLITICAL INSTITUTIONS

Milan W. Svolik[1]

Introduction

At their most political, institutions determine who holds power and how power is exercised. Political institutions thus shape the influence and behavior of their participants and, for that very reason, their adoption and design are the subject of political struggle. Formal political theory—the application of game-theoretic methods to political phenomena—develops this basic insight.[2] Game-theoretically inspired arguments explain the emergence and persistence of political institutions as the equilibrium outcome of strategic interaction among actors who strive to anticipate the implications of the institutional status quo and its conceivable alternatives for both their own and others' welfare.

Formal political theory provides an attractive and general analytical framework for the study of political institutions for several reasons. First, the analysis of political institutions is particularly amenable to formal modeling as institutional rules frequently delineate precisely those elements that are required for a well-defined game-theoretic model: the participants, their available actions and information, and the procedures by which their actions translate into outcomes—like the majority vote, for instance. Depending on whether our central question concerns the consequences of institutions, their emergence and persistence, or their role in equilibrium selection, formal-theoretic analyses of institutions typically proceed in three conceptually distinct ways.[3]

In its most direct approach to the study of institutions, formal political theory takes institutions as given and examines their implications for the behavior of those governed by them. A key feature of this approach is the recognition that political actors understand that their welfare depends on not only their own actions but also the actions of others, who hold potentially conflicting interests. The analysis of the equilibrium consequences of institutions therefore often leads to predictions that are not immediately obvious from the institutions' formal structure, like the sharp prediction of candidate platform convergence and the entrenchment of two parties in majoritarian electoral systems that I discuss in the next section.

By contrast, when we analyze institutions as equilibria, we are interested in not only the political consequences of institutions but also why, given their consequences, key actors have an incentive to establish and maintain those institutions in the first place. In this more analytically complete view of institutions, both the behavior and the institutions that govern it must jointly constitute an equilibrium—they must be self-enforcing. This approach is especially appropriate

for the analysis of foundational institutions and weakly institutionalized settings. When we study the emergence and persistence of democracy, electoral systems, or the rule of law, key conceptual puzzles concern not only these institutions' consequences but also why, in light of their consequences, political actors abide by them in the absence of a higher authority that could enforce their compliance. This approach to the analysis of institutions thus yields insights into not only their potential benefits but also their limitations and the obstacles to institutional reform. In the section on 'institutions as strategic equilibria' (pp. 73–74), I illustrate these points by contrasting two perspectives on the emergence and survival of democracy, which I respectively refer to as the redistributive and the accountability models of self-enforcing democracy.

The third distinct approach to the study of institutions emphasizes their role in focal coordination and equilibrium selection. In many settings, a game-theoretic analysis predicts the existence of multiple equilibria; that is, more than one pattern of behavior is consistent with strategic rationality in a given setting. Schelling's (1960) focal point effect helps us understand how institutionally embedded cultural expectations can play a decisive role in coordinating every actor on one among several equilibria. Hence, rather than an undesirable indeterminacy, this approach to the study of institutions recognizes the multiplicity of equilibria as a fundamental feature of politics and an opportunity to incorporate historically and culturally shaped expectations into the equilibrium analysis of institutions. I illustrate this approach by discussing the focal role of supreme courts and the act of inauguration in the efficient arbitration of political disputes and transfers of power.

The second and conceptually distinct set of benefits that formal political theory brings to the study of institutions stems from the productive discipline that the construction and analysis of transparent political models bring to theory-building. This discipline takes at least two forms. The first is required by the equilibrium concepts that we employ in the analysis of our models. The Nash equilibrium and its refinements posit that in order for a political scenario to be in equilibrium, at a minimum, no actor can have a unilateral incentive to deviate from it. This steady-state requirement is an appealing "negative" criterion for predicting political outcomes.[4] That is, while this requirement may admit multiple outcomes as equilibria, each of which may vary in its plausibility, any behavior that it excludes is definitely implausible: A non-equilibrium prediction effectively expects that at least one actor will act against her own interests. Crucially, the Nash equilibrium and its most frequently employed refinements exist for any well-specified game (Nash 1951). That is, we can employ game-theoretic methods to study essentially *any* political setting. Thus we can think of the analytical discipline required by the Nash equilibrium and related equilibrium concepts as a general methodology for theory-building in political science.[5]

The second form of the productive discipline that formal political theory brings to our study of institutions emerges in the process of model-building. Any game-theoretic analysis begins with an explicit description of the relevant political setting: the players, their preferences, and the actions and information available to them. The conceptual process by which these elements are specified entails abstraction and simplification, and, as a result, a trade-off of descriptive realism in favor of analytical tractability and transparency. Rather than a drawback, formal theory views the analysis of transparent, tractable, even if simple models as a key step toward our understanding of more complex political settings. Its by-product is analytical clarity about assumptions, logic, and empirical implications that allows us to draw connections between conceptually related mechanisms in substantively different contexts. A number of such mechanisms, including costly signaling, cheap talk, ultimatum bargaining, as well as commitment, collective action, and principal–agent problems, has become a part of political scientists' conceptual vocabulary. Thus even those political scientists who do not find particular game-theoretic equilibrium concepts

or mathematical modeling appealing nevertheless benefit from the heuristic transparency that formal political analysis brings to our understanding of politics.

I conclude this chapter by discussing how formal political theory contributes to empirical research and the challenges it encounters therein. The interaction between formal modeling and empirical research is essential for the construction of models that both are tractable and capture key aspects of real-world politics. At the same time, game-theoretically inspired modeling contributes to empirical research by clarifying the strategic origins of frequently encountered problems in causal inference, like endogeneity and selection effects. By providing the microfoundations for these inference problems, formal theoretic research helps guide the search for plausible identification strategies and helps us evaluate their external validity.

Equilibrium consequences of institutions

When we study the equilibrium consequences of institutions, we take a particular institution as given and examine how the incentives created by that institution shape the behavior of its participants.[6] This approach to the analysis of institutions is one of the earliest in historical terms—as in Condorcet's eighteenth century investigation of voting cycles under majority rule. It is also conceptually antecedent, as we think that the primary driver of institutional reforms (or the lack thereof) is their anticipated consequences. As an illustration of this approach, consider the game-theoretic analysis of two distinctive consequences of majoritarian electoral systems: platform convergence and Duverger's law.

A canonical model conceives of democracy as a political system in which two candidates compete for the votes of a large number of voters whose policy preferences vary along a single dimension. First, each of the two candidates proposes a policy that he will later implement; then each voter votes for the candidate whose proposed policy is closest to his or her preferred policy. Unlike the voters, the candidates do not care about policy but only about winning office. The candidate who obtains the larger share of the vote wins the election (a tie is decided by the flip of a coin) and implements the policy that he proposed.

This simple model, frequently labeled Hotelling's or the Downsian model of electoral competition after its early proponents (Hotelling 1929; Downs 1957), yields sharp predictions about the nature of electoral politics. In the unique equilibrium, both candidates propose the policy favored by the median voter and therefore tie. The intuition behind this result builds on the prominent role that the median voter plays in majoritarian decision-making.[7] Because the median voter evenly divides the electorate, any candidate who adopts a platform other than the one favored by the median voter could be defeated by a competitor who positions himself closer to the median. Only a candidate who proposes the median voter's favorite policy cannot be defeated in this way: The best that his or her competitor can do is to propose the same median policy and therefore tie the election.[8]

While the Hotelling–Downs model of electoral competition takes the number of candidates as given, the formal analysis of coordination problems faced by voters in majoritarian electoral systems derives Maurice Duverger's (1959) conjecture that these electoral rules result in the competition of only two candidates or parties as an equilibrium (Riker 1982; Cox 1997). In an electoral system with single-member districts and a single vote per voter (SMD), voters who favor candidates whose perceived popularity ranks at the third or worse place have an incentive to desert these candidates in favor of one of the top two contenders out of fear that their vote will be wasted. In turn, we should observe only two serious contenders in any district. This intuition extends to electoral systems with more than two candidates per district: Cox's (1997) $M + 1$ law posits that there will be only $M + 1$ serious contenders in a district that elects M candidates (see also Shepsle 1991).

The prediction that SMD electoral rules will result in a contest between only two competitive candidates can thus be viewed as a justification for the assumption of two-candidate competition in the Hotelling–Downs model and as a statement about the relatively high barriers to entry in SMD systems (Myerson 1999b). Meanwhile, the Hotelling–Downs model of electoral competition clarifies why two-candidate majoritarian elections result in the convergence of candidate platforms to the median voter's preferred policy, and in turn, close election outcomes. Crucially, observe that SMD rules do not expressly prohibit more than two candidates from competing for office, nor do they prevent voters from supporting candidates ranked third or worse. Similarly, the intuition behind platform convergence in the Hotelling–Downs model does not depend on any specific distribution of voters; in particular, we are not assuming that most voters are located close to the median. Rather, it is the anticipation of how SMD rules shape the strategic interaction among voters and candidates—the analysis of the equilibrium consequences of this institution—that accounts for the sharp predictions about the number of candidates and the nature of their platforms.

Institutions as strategic equilibria

According to Douglas North's oft-cited definition, institutions are the "humanly devised constraints that structure political, economic and social interactions" (North 1991: 97). Thus unlike geographical or resource constraints, institutions can be devised and changed by the very actors whose behavior they are supposed to bind. Our explanations approach this feature of institutions with political realism when we conceive of the emergence and persistence of political institutions as the equilibrium outcome of strategic interaction. Analytically, this amounts to a more coherent view of institutions than the analysis of the equilibrium consequences of institutions alone as it presupposes that the agents who decide to adopt or maintain an institution understand the equilibrium consequences of the status quo and its conceivable alternatives.[9]

As an illustration of this approach, consider two competing perspectives on the emergence and survival of democracy, which I respectively refer to as the redistributive and accountability models of democratization. Building on the implications of the Hotelling–Downs model of electoral competition for redistributive politics (Meltzer and Richards 1981), the democratization models of Acemoglu and Robinson (2001, 2005) and Boix (2003) emphasize that the key political consequence of a transition from dictatorship to democracy is the redistribution of income from the rich to the poor. That is, while redistribution is limited under dictatorship because power is held by the rich, electoral competition under democracy must heed the preferences of the (much poorer) median voter and thus results in significant income redistribution. A majority of the population thus prefers democracy to dictatorship and may threaten a revolution with more radical redistributive consequences if its demands are not met. According to this perspective, stable democracy emerges at moderate levels of economic inequality when a significant fraction of the population demands it but the rich do not fear it so much as to find repressing the redistributive demands of the poor under a dictatorship more attractive.[10]

Contrast this reasoning to the accountability view of democracy. According to this perspective, the major difference between dictatorship and democracy is that, in the latter, electoral competition affords the public the opportunity to hold politicians accountable for their performance in office. In turn, democratic elections generate incentives for politicians to promote the general welfare (Barro 1973; Ferejohn 1986), whereas authoritarian elites cater primarily to a much narrower "selectorate" (Bueno de Mesquita *et al.* 2003; Besley and Kudamatsu 2007; Boix and Svolik 2013). Accordingly, in Lizzeri and Persico's (2004) model of franchise extension, elites

favor democracy when the status quo favors only a small fraction of the elite—the landed classes—and the extension of the franchise promises public-good oriented policies that are key for the welfare of the growing commercial and urban classes. Meanwhile, democracy survives according to the accountability perspective when it indeed delivers policies that are better than those under a dictatorship because only then can voters be realistically expected to defend democracy against politicians or groups with authoritarian ambitions (Weingast 1997; Fearon 2011). In the models of Myerson (2006), Bidner and Francois (2013), and Svolik (2013), elections may fail to deliver accountability when voters come to believe that most politicians are self-serving and any attempt to discriminate among them based on their performance is therefore a waste of time. When espoused by a sizeable fraction of the electorate, such expectations are self-fulfilling: They fuel a mutually reinforcing cycle of voter apathy and poor government performance. Hence, according to this perspective, the key threat to democratic stability is not the redistributive conflict between the rich and the poor but rather the failure of elections as an instrument of accountability.

The analysis of institutions as equilibria is especially valuable for our understanding of the emergence and persistence of foundational institutions, like democracy, electoral systems, or the rule of law, and our understanding of weakly institutionalized settings, including international, post-conflict, or authoritarian politics.[11] In these settings, a key conceptual puzzle behind the emergence or persistence of institutions stems from the absence of a higher, independent authority with the power to enforce the compliance of key actors with institutional rules when doing so is against their interests.[12] Thus in their analysis of the emergence and persistence of democracy, both the redistributive and accountability perspectives start by outlining the redistributive and accountability consequences of dictatorship and democracy for key actors: the rich versus the poor and the elites versus the public, respectively. Democracy is an equilibrium—it is self-enforcing (Przeworski 1991)—when no key political actor has an incentive to undermine it in light of its anticipated consequences, whether it be the rich by coopting the military in order to stage a coup or an incumbent by refusing to step down after losing an election.

But in other contexts, it may be more productive to (at least initially) bracket the question of the origins of institutions and focus exclusively on their equilibrium consequences. For instance, while the origins of single-member district electoral systems may be traced back to the proto-democratic institutions of early modern Europe and their colonial dissemination, contemporary persistence of SMD electoral systems may be rooted in the entrenched two-party systems that SMD rules foster (Cox 1997). A shift to an alternative electoral formula may therefore be feasible only when new parties that emerge due to the extension of the franchise or an economic transformation threaten to fragment established parties (Boix 1999). Thus when Cox (1997) studies the coordination dilemmas that electoral systems create for voters and parties, he can reasonably bracket the question of the origin of electoral systems on analytical grounds: An initial analysis may be most tractable if it focuses on the consequences of electoral systems alone. Such a "partial equilibrium" analysis is also warranted on substantive grounds: Electoral rules change only at critical junctures and can therefore be considered both binding and fixed, at least in the short run.

Institutions as focal points

In many political settings, the Nash equilibrium and related solution concepts predict that more than one pattern of behavior is consistent with rational, strategic reasoning—that there are multiple equilibria. In some cases, the indeterminacy associated with a multiplicity of equilibria may be the consequence of an underspecified model and thus only a technical artefact.[13] But in many

instances, the multiplicity of equilibria is a fundamental feature of politics. In these cases, a major function of institutions can be to focally coordinate their participants on one among the multiple equilibria (Schelling 1960).

As an illustration, consider the focal properties of two institutions, the supreme court and the act of inauguration. The scenario in which the supreme court acts as an arbiter of last resort is only one equilibrium in a broader game in which a legal dispute could also be decided by a costly contest whose outcome depends on the contending parties' brute force. When a supreme court arbitrates a disputed election, for instance, its ruling might be seen as just one among the many opinions typically given by various actors, from the candidates to media pundits to foreign observers. But in a functioning constitutional order, the court's opinion becomes a self-fulfilling prophecy by virtue of its prominent status in the constitutional hierarchy.[14] The court's verdict is self-enforcing because it coordinates the expectations of the multitude of actors on whose consent the proper functioning of any constitutional order depends. Hence a losing candidate who would contemplate ignoring the court's ruling must anticipate opposition from not only the declared winner but also those actors. Its focalness endows the court with the power to preclude costly disputes in spite of the fact that, in strictly material terms, the court's opinion was just that—an opinion.

Similarly, the symbolic act of inauguration marks the transfer of power from one leader to the next by focally coordinating expectations about who heads the executive among the large number of actors that constitutes the hierarchy of any government. One metric of the focal power of this symbolic act is the difference between a departing leader's order on the day before and on the day after the inauguration of his successor. By contrast, in political systems that lack the institution of inauguration or where its relevance is in doubt, genuine transfers of power must be accompanied by a demonstration of force or a costly political conflict. Thus in contemporary Russia, for instance, it has become apparent that the effective head of the Russian government is neither the president of the Russian Federation nor its prime minister. Rather it is Vladimir Putin, regardless of the official post that he holds. As a consequence, Putin's potential successor cannot expect to genuinely assume power by the mere symbolic act of inauguration. He will have to wrestle power away from Putin publicly and personally—by exiling, imprisoning, or killing him.[15]

Schelling's (1960) focal point effect thus helps us understand how institutions can become the tools of efficient arbitration of political disputes and transfers of power. These two examples illustrate a broader Humean view of foundational political institutions (Hume 1748 [1987]): Institutions are conventions (rather than contracts) whose key role is to establish shared expectations that coordinate the citizenry on one among multiple (and not equally efficient) self-enforcing ways to organize political interactions.[16] The central role of many constitutional provisions, in this view, is to focally coordinate everyone's expectations on clearly defined limits on the government's authority (Weingast 1997).[17]

Because focalness refers to environmental factors that may be conspicuous but do not directly affect the participants' welfare, focal coordination allows for the incorporation of cultural, psychological, and historically contingent factors into the rational choice analysis of institutions.[18] According to this approach to the analysis of institutions, many of the symbols and ceremonial acts that accompany real-world institutions are seen as cultural artifacts that serve to reproduce a shared understanding of the focal role that specific institutions play in equilibrium selection (Chwe 2001). But this does not imply that cultural or symbolic factors override strategic concerns. Only when multiple patterns of behavior are consistent with strategic reasoning can institutionally embedded cultural expectations play a decisive, focal role in equilibrium selection.

The heuristic value of formal models

The first step of a game-theoretic analysis of politics is to explicitly describe the relevant political setting: the players, their preferences, and each player's available actions and information. Political institutions are especially amenable to this initial step since their infrastructure often mirrors many of these elements. Nonetheless, the make-up and operation of most real-world institutions are too complex to be modeled in their entirety. In turn, the conceptual process by which these elements are specified entails abstraction and simplification, steps that require a trade-off of descriptive realism in favor of analytical tractability and transparency.

Rather than a drawback, game-theoretically inspired theory-building views the analysis of simple, tractable models as a necessary and productive first step toward the understanding of more complex environments. The abstraction and simplification involved in formal theory-building foster a *productive discipline* that is essential to any kind of theory-building, both mathematical and less formal.[19] The specification of the elements of a well-defined game amounts to a statement of the key forces that the analyst believes operate in the political setting under study. Meanwhile, the resulting analytical tractability and transparency facilitate the reproducibility of theoretical arguments (just like data and code sharing facilitate replication in empirical political science) and make it easier to draw connections between analytically related mechanisms in substantively different fields. Thus even those political scientists who are skeptical about the plausibility of game-theoretic equilibrium concepts will benefit from the productive discipline required by the methodology of model-based theorizing and the ensuing heuristic transparency.

As an example of a productive departure from descriptive accuracy, consider the ultimatum bargaining model. According to this model, one of two players first proposes how to divide a positive quantity (money in the simplest application) and then the other player either accepts or rejects the proposal. If the proposal is accepted, each player gets his or her share of the proposed division; if it is rejected, both players get nothing. This simplest model of sequential bargaining results in a sharp equilibrium prediction: Because the second player is effectively choosing between accepting the first player's proposal or getting nothing, the first player optimally demands and obtains the entire quantity divided; no rejections occur in equilibrium.[20] Hence the alternate name for this model: "take-it-or-leave-it" bargaining.

This seemingly unrealistic model of bargaining has become one of the key building blocks for the formal analysis of a range of political institutions.[21] In their analysis of budgetary agenda-setting, Romer and Rosenthal (1978) reformulated the ultimatum game by conceiving of the first player as a committee that proposes a policy (e.g. a school board) and of the second player as a collective actor that decides by a majority vote (e.g. a referendum) whether to accept the committee's proposal or whether to retain the status quo. This extension of the ultimatum game became the workhorse model in the study of legislative organization (Denzau and Mackay 1983) and the separation of powers, especially the study of executive–legislative bargaining (Cameron 2000). Meanwhile, an extension of the ultimatum game according to which a rejection of the first player's proposal results in a costly conflict became the leading approach to the study of war (Fearon 1995; Powell 1999). In political applications of the ultimatum game, the initial trade-off of descriptive realism in favor of analytical tractability proved to be essential in facilitating the development of more realistic models in which conflict may occur in equilibrium (typically as a result of private information) and in which the power of key actors is shaped either by the political environment (e.g. the location of the status quo) or by the details of the institutional setting (e.g. open versus closed committee rules, the requirements for a veto override.)

Political applications of the ultimatum game also illustrate the heuristic value of our thorough understanding of this mechanism: The tractability and transparency of these models help us

discern analytical connections between areas as disparate as the committee system in the U.S. House of Representatives, the separation of powers, and international crisis bargaining. A highly incomplete sample of other mechanisms frequently encountered in the formal analysis of institutions includes the classic dilemmas (i.e. the Prisoner's Dilemma, the Stag Hunt, the Game of Chicken, the Battle of the Sexes, the Matching Pennies), the problems of commitment and non-credible threats, collective action problems, principal–agent problems, costly signaling, and cheap talk.[22] As in the case of the ultimatum game, these mechanisms and the intuitions behind them have been productively applied in issue areas well beyond those in which they initially emerged.

Conclusion: the symbiosis of theoretical modeling and empirical testing

I conclude this chapter by discussing how formal analysis of institutions benefits from and contributes to empirical research. While formal political modeling is primarily a theoretical enterprise, its interaction with empirical research is essential for the construction of models that capture key aspects of real-world politics.[23] At the same time, the empirical evaluation of formal models presents unique challenges: Models are at best approximations of reality and good models are deliberate about the many aspects of reality that they ignore and the few that they focus on. In turn, it is easy to dismiss any formal model as failing to capture some aspect of real-world politics. But that would miss the key purpose of modeling: to offer a tractable analysis of some aspect of politics. To paraphrase a famous quotation, since all models are approximations at best, all models are wrong—but some are useful (Box and Draper 1987: 424).

Any kind of theory-building, whether mathematical or less formal, entails abstraction and simplification and hence a departure from descriptive realism. Formal theorists view these aspects of modeling as a productive step toward the construction of models that serve a specific purpose, which is as often analytical or heuristic as it is to yield testable empirical predictions. This approach to modeling as a multi-purpose enterprise is more productive than attempts to construct a single, complex model that would be at the same time analytically tractable, theoretically revealing, and yield a wealth of concrete empirical predictions (Myerson 1992: 64–66). The value of specific models should therefore not be judged by their descriptive details or even by the richness or precision of their empirical predictions, but rather by whether they accomplish their intended purpose.[24]

In turn, when assessing the predictions of formal models, we should focus on the key forces hypothesized to shape equilibrium behavior, test a model's comparative statics rather than point predictions, and evaluate its goodness of fit by the standard of alternative, competing models. Meanwhile, when evaluating the appropriateness of modeling assumptions, we need to differentiate between models whose emphasis is analytical and therefore have only crude empirical implications and models whose emphasis is to highlight their predictions for a specific type of data or identification strategy. Thus while Schelling's (1960, 1966) and Fearon's (1995) models of commitment and bargaining problems in international crises are primarily analytical and heuristic, Signorino's (1999) and Lewis and Schultz's (2003) models highlight the implications of these mechanisms for statistical inference from international conflict data.

The evolution of models of electoral competition illustrates the productive interaction of theoretical modeling and empirical analysis. When viewed in strictly empirical terms, the predictions of the Hotelling–Downs model discussed in the section on "Equilibrium consequences of institutions" (pp. 72–73) are obviously "wrong": Candidates never propose perfectly identical platforms and there are no perfectly tied large-turnout elections on record. Attempts to bring the model's prediction closer to empirical observations led to extensions that capture richer political

settings and, in turn, more realistic predictions. Some of the key extensions include models in which candidates care about policy (Wittman 1973; Calvert 1985), cannot commit to platforms (Osborne and Slivinski 1996; Besley and Coate 1997), compete over multiple dimensions (Roemer 2001), face the threat of entry (Palfrey 1984), differ in platform-specific skills (Krasa and Polborn 2010) or valence characteristics (Ansolabehere and James M. Snyder Jr., 2000; Groseclose 2001; Aragonesa and Palfrey 2002); and models in which voters experience a random shock to their preferences before they vote (Lindbeck and Weibull 1987; Coughlin 1992; Persson and Tabellini 2000), form parties (Roemer 2001), become candidates themselves (Osborne and Slivinski 1996; Besley and Coate 1997), and may choose to abstain (Myerson 1998, 2000).[25]

These extensions help clarify what kind of departures from the initial Hotelling–Downs setting result in equilibrium platform divergence and election outcomes that are not (ex-post) perfectly tied. Yet, at the same time, these extensions also confirm that competitive pressures toward platform convergence are a general feature of majoritarian electoral systems. Thus, once confronted with data, the Hotelling–Downs model proved to be a theoretically productive starting point.

Game-theoretically inspired modeling not only benefits from interaction with empirical research, it also contributes to it by highlighting the challenges to statistical inference and external validity. By emphasizing that most political interactions involve actors with potentially conflicting interests, formal analysis provides the microfoundations for recurring problems in empirical inference, like endogeneity and selection effects.

Consider one major obstacle in the estimation of the causal effects of institutions—their endogeneity. Endogeneity refers to the concern that the presumed consequences of institutions are simultaneously the cause of their emergence and persistence. According to the most skeptical view, institutions are epiphenomenal—their presumed consequences are entirely due to the material conditions that are responsible for their emergence and persistence.[26] This concern about the endogeneity of presumed causes to their effects mirrors the conceptual distinction between the partial analysis of the equilibrium consequences of institutions and the full analysis of institutions as equilibria. In fact, a frequently employed nomenclature distinguishes these as the exogenous and endogenous approaches to the study of institutions (Weingast 1998; Diermeier and Krehbiel 2003).

In turn, the challenge in identifying the consequences of institutions is to isolate the causal effect of an institution from the indirect effects of factors that contribute to its persistence. As our earlier discussion of Duverger's law highlights, by favoring two-candidate competition majoritarian electoral systems generate vested interests in the perpetuation of such electoral rules, and, in turn, further the entrenchment of both the electoral system and two major parties. As a result, the causes of these institutions can be plausibly separated from their consequences only at particularly favorable historical moments. Electoral systems, for instance, may be plausibly exogenous only at the time of their colonial dissemination or when new parties threaten established ones due to unexpected demographic or economic shocks.[27] Models of Duverger's law thus provide the analytical microfoundations for institutional path-dependence and highlight the challenges to empirical inference about the consequences of electoral systems.

Finally, explicit—but not necessarily formal—theory-building can help empirical researchers evaluate the external validity of specific cases and research designs. Most empirical research aspires to claim that its findings are relevant beyond the specific case, context, or data from which it originates. But claims of such external validity then amount to a conceptual justification for why some political mechanism operates similarly across different contexts. Explicit modeling of the interaction between institutions and their participants helps us evaluate which features of an empirical case or data are relevant to the mechanism at work and thus generalizable beyond the specifics of any context.

Notes

1 I would like to thank Jose Cheibub, Bonnie Weir, Matt Winters, and the editors for insightful comments and conversations, and Abdullah Aydogan, Tyler Pack, Hollan Peterson, Ashlea Rundlett, Kelly Senters, and Wei Zhong for editorial and research assistance.

2 For an introduction to formal political theory, see McCarty and Meirowitz (2007), Morrow (1994), and Osborne (2004).

3 For complementary perspectives on formal analysis of institutions, see Weingast (1998) and Shepsle (2006).

4 Myerson (1991: 107–108) refers to this "negative" predictive plausibility of the Nash equilibrium as an "upper solution concept."

5 For a discussion of the value and limits of game-theoretic analysis in political science and related social sciences, see Aumann (1985), Geddes (2003: Ch. 5), Kreps (1990), Morton (1999), Myerson (1992, 1999a, 2009), Powell (1999: Ch. 1), Rubinstein (1991), and Tsebelis (1990: Ch. 2).

6 In an influential paper, Shepsle (1979) refers to this type of institutional analysis as "structure–induced equilibrium."

7 The median voter has exactly one half of the electorate to both his or her left and right when the electorate is ordered along a single policy dimension. Black's Median Voter Theorem (Black 1958) states that the ideal point of the median voter is a Condorcet winner—it is preferred by a majority to any other platform—if preferences are single-peaked. See Myerson (2013) for a recent review of social choice theory.

8 For an introductory formal exposition, see Osborne (2004: 73), Gehlbach (2013: 2–3), and McCarty and Meirowitz (2007: 101–103).

9 For early appeals to treat institutions as equilibria, see Riker (1980), Shepsle (1986), and Calvert (1995a)

10 This prediction is based on the models in Acemoglu and Robinson (2001, 2005); Boix (2003) predicts that democracy and inequality should be negatively correlated.

11 On electoral systems, see Boix (1999); on the rule of law, see Weingast (1997) and Hafer (2006); on international institutions, see Koremenos, Lipson, and Snidal (2004); on authoritarian institutions, see Gandhi (2008) and Boix and Svolik (2013).

12 The relevance of this conceptual point to foundational political questions is eloquently captured in James Madison's "In framing a government which is to be administered by men over men, the great difficulty lies in this: You must first enable the government to control the governed; and in the next place, oblige it to control itself" (Madison 1788 [2010]).

13 The global game reformulation of the Stag Hunt game, for instance, yields a unique equilibrium, whereas the original Stag Hunt has multiple; on global games, see Carlsson and van Damme (1993) and Morris and Shin (2003).

14 On the role of law in focal coordination, see McAdams (2000).

15 Similarly, the last Mexican caudillo Plutarco Elías Calles officially stepped down from the Mexican presidency in 1928 but nevertheless overshadowed his three successors, inspiring the expression "the president lives in the presidential palace, but the man who gives orders lives across the street" (Krauze 1997: 430). The genuineness of his political demise during Lázaro Cárdenas's presidency had to be marked by a forced exile to the United States.

16 See especially the discussions by Calvert (1995b), Hardin (1989), and Myerson (2004, 2009).

17 For related arguments, see de Figueiredo and Weingast (2005) and Tucker (2007).

18 Greif (2006: Ch. 9) discusses the focal role of cultural beliefs in medieval institutional development; Elster (2004: 176–177) suggests that focalness played a substantial role in the forming of a consensus about what "just" compensation was for forced laborers during World War II; Elster, Offe, and Preuss (1998: 62) propose that past constitutional choices become natural focal points in constitutional design during transitions to democracy.

19 Put differently, non-formal theories implicitly assume most of the elements required for a well-defined game model but they are typically less transparent about them; see Epstein (2008).

20 For an introductory exposition of this model, see McCarty and Meirowitz (2007: 176–177) and Osborne (2004: 181–182).

21 In fact, when ordinary people play this game, they are rarely as self-regarding as the basic ultimatum bargaining model assumes (Camerer and Thaler 1995).

22 On the role of commitment problems in the analysis of institutions, see North and Weingast (1989), Greif, Milgrom, and Weingast (1994) and Acemoglu and Robinson (2005); on the role of institutions in resolving collective action problems, see Olson (1965), Hardin (1982), Ostrom (1990), and Aldrich

(1995); on principal–agent treatments of electoral competition and bureaucracy, see Ferejohn (1986) and Epstein and O'Halloran (1999); for institutional applications of costly signaling and cheap talk, see Banks (1991), Gilligan and Krehbiel (1987), and Krishna and Morgan (2001). For a rigorous introduction and more comprehensive review of these topics, see Austen-Smith and Banks (1999, 2005), Gehlbach (2013), Mueller (2003), and Persson and Tabellini (2000).

23 Myerson (1992: 64–66) refers to the iterated interaction between theoretical modeling and empirical evidence as the "modeling dialogue."

24 On empirical evaluation of formal models, see especially Myerson (1992), Morton (1999), Granato, Lo, and Wong (2010), Clarke and Primo (2012), and Lorentzen, Fravel, and Paine (2013). For skeptical perspectives on whether game-theoretic models should be empirically evaluated in the first place, see Aumann (1985) and Rubinstein (2006).

25 For excellent exposition and review of these developments, see Osborne (1995), Persson and Tabellini (2000), Grofman (2004), and Austen-Smith and Banks (2005: Chs. 7–9).

26 For a discussion of this skeptical view, see Przeworski (2009) and Pepinsky (forthcoming).

27 For an influential attempt to overcome institutional endogeneity via an instrumental variable approach, see Acemoglu, Johnson, and Robinson (2001).

References

Acemoglu, Daron and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91: 938–963.

—. 2005. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Aldrich, John H. 1995. *Why Parties? The Origin and Tansformation of Party Politics in America*. Chicago, IL: University of Chicago Press.

Ansolabehere, Stephen and James M. Snyder Jr. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice* 103 (3–4): 327–336.

Aragonesa, Enriqueta and Thomas R. Palfrey. 2002. "Mixed Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory* 103 (1): 131–161.

Aumann, Robert J. 1985. "What Is Game Theory Trying to Accomplish?" In *Frontiers of Economics*, ed. Kenneth J. Arrow and Seppo Honkapohja. London: Basil Blackwell.

Austen-Smith, David and Jeffrey S. Banks. 1999. *Positive Political Theory I*. Ann Arbor: University of Michigan Press.

—. 2005. *Positive Political Theory II*. Ann Arbor: University of Michigan Press.

Banks, Jeffrey S. 1991. *Signalling Games in Political Science*. New York: Routledge.

Barro, Robert. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14: 19–42.

Besley, Timothy and Stephen Coate. 1997. "An Economic Model of Representative Democracy." *Quarterly Journal of Economics* 112 (1): 85–114.

Besley, Timothy and Masayuki Kudamatsu. 2007. "Making Autocracy Work." *STICERD Development Economics Discussion Paper no. 48*.

Bidner, Chris and Patrick Francois. 2013. "The Emergence of Political Accountability." *Quarterly Journal of Economics* 128 (3): 1397–1448.

Black, Duncan. 1958. *Theory of Committees and Elections*. Cambridge: Cambridge University Press.

Boix, Carles. 1999. "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review* 93 (3): 609–624.

—. 2003. *Democracy and Redistribution*. Cambridge: Cambridge University Press.

Boix, Carles and Milan W. Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions and Power-sharing in Dictatorships." *Journal of Politics* 75 (2): 300–316.

Box, George E. P. and Norman R. Draper. 1987. *Empirical Model-building and Response Surfaces*. New York: Wiley.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Calvert, Randall L. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science* 29 (1): 69–95.

—. 1995a. "Rational Actors, Equilibrium, and Social Institutions." In *Explaining Social Institutions*, ed. Jack Knight and Itai Sened. Ann Arbor: University of Michigan Press.

—. 1995b. "The Rational Choice Theory of Social Institutions: Cooperation, Coordination, and Communication." In *Modern Political Economy: Old Topics, New Directions*, ed. Jeffrey S. Banks and Eric A. Hanushek. Cambridge: Cambridge University Press.

Camerer, Colin F. and Richard H. Thaler. 1995. "Anomalies: Ultimatums, Dictators and Manners." *Journal of Economic Perspectives* 9 (2): 209–219.

Cameron, Charles M. 2000. *Veto Bargaining: Presidents and the Politics of Negative Power*. Cambridge: Cambridge University Press.

Carlsson, Hans and Eric van Damme. 1993. "Global Games and Equilibrium Selection." *Econometrica* 61 (5): 989–1018.

Chwe, Michael. 2001. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton: Princeton University Press.

Clarke, Kevin A. and David M. Primo. 2012. *A Model Discipline: Political Science and the Logic of Representations*. New York: Oxford University Press.

Coughlin, Peter J. 1992. *Probabilistic Voting Theory*. New York: Cambridge University Press.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. New York: Cambridge University Press.

de Figueiredo, Rui and Barry R. Weingast. 2005. "Self-Enforcing Federalism." *Journal of Law, Economics, & Organization* 21 (1): 103–135.

Denzau, Arthur T. and Robert J. Mackay. 1983. "Gatekeeping and Monopoly Power of Committees: An Analysis of Sincere and Sophisticated Behavior." *American Journal of Political Science* 27 (2): 762–79.

Diermeier, Daniel and Keith Krehbiel. 2003. "Institutionalism as a Methodology." *Journal of Theoretical Politics* 15 (2): 123–144.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Duverger, Maurice. 1959. *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.

Elster, Jon. 2004. *Closing the Books: Transitional Justice in Historical Perspective*. Cambridge: Cambridge University Press.

Elster, Jon, Claus Offe, and Ulrich K. Preuss. 1998. *Institutional Design in Post-communist Societies: Rebuilding the Ship at Sea*. Cambridge: Cambridge University Press.

Epstein, David and Sharyn O'Halloran. 1999. *Delegating Powers*. Cambridge: Cambridge University Press.

Epstein, Joshua M. 2008. "Why Model?" *Journal of Artificial Societies and Social Simulation* 11 (4): 1–12.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414.

—. 2011. "Self-enforcing Democracy." *Quarterly Journal of Economics* 126 (4): 1661–1708.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50: 5–26.

Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship*. New York: Cambridge University Press.

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.

Gehlbach, Scott. 2013. *Formal Models of Domestic Politics*. Cambridge: Cambridge University Press.

Gilligan, Thomas and Keith Krehbiel. 1987. "Collective Decision-making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law, Economics, & Organization* 3: 145–193.

Granato, Jim, Melody Lo, and M. C. Sunny Wong. 2010. "A Framework for Unifying Formal and Empirical Analysis." *American Journal of Political Science* 54 (3): 783–797.

Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. New York: Cambridge University Press.

Greif, Avner, Paul Milgrom, and Barry R. Weingast. 1994. "Coordination, Commitment, and Enforcement: The Case of the Merchant Guild." *Journal of Political Economy* 102: 745–776.

Grofman, Bernard. 2004. "Downs and Two-Party Convergence." *Annual Review of Political Science* 7 (1): 25–46.

Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45 (4): 862–886.

Hafer, Catherine. 2006. "On the Origins of Property Rights: Conflict and Production in the State of Nature." *Review of Economic Studies* 73 (1): 119–143.

Hardin, Russell. 1982. *Collective Action*. Baltimore, MD: Johns Hopkins University Press.

—. 1989. "Why a Constitution?" In *The Federalist Papers and the New Institutionalism*, ed. Bernard Grofman and Donald Wittman. New York: Agathon Press.

Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal* 39: 41–57.

Hume, David. 1748 [1987]. "Of the Original Contract." In *Essays, Moral, Political, and Literary*, ed. Eugene F. Miller. Indianapolis: Liberty Classics.

Koremenos, Barbara, Charles Lipson, and Duncan Snidal (eds.). 2004. *The Rational Design of International Institutions*. Cambridge: Cambridge University Press.

Krasa, Stefan and Mattias Polborn. 2010. "Competition between Specialized Candidates." *American Political Science Review* 104 (4): 745–765.

Krauze, Enrique. 1997. *Biography of Power: A History of Modern Mexico, 1810–1996*. New York: Harper.

Kreps, David M. 1990. *Game Theory and Economic Modelling*. New York: Oxford University Press.

Krishna, Vijay and John Morgan. 2001. "Asymmetric Information and Legislative Rules: Some Amendments." *American Political Science Review* 95: 435–452.

Lewis, Jeffrey B. and Kenneth A. Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11 (4): 345–367.

Lindbeck, Assar and Jorgen W. Weibull. 1987. "Balanced-budget Redistribution as the Outcome of Political Competition." *Public Choice* 54 (3): 273–297.

Lizzeri, Alessandro and Nicola Persico. 2004. "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform'." *Quarterly Journal of Economics* 119: 707–765.

Lorentzen, Peter L., M. Taylor Fravel, and Jack Paine. 2013. "Bridging the Gap: Using Qualitative Evidence to Evaluate Formal Models." Unpublished manuscript, University of California, Berkeley.

McAdams, Richard H. 2000. "A Focal Point Theory of Expressive Law." *Virginia Law Review* 86 (8): 1649–1729.

McCarty, Nolan and Adam Meirowitz. 2007. *Political Game Theory*. New York: Cambridge University Press.

Madison, James. 1788 [2010]. "Federalist No. 51." In *The Federalist Papers*, ed. Michael A. Genovese. New York: Palgrave Macmillan.

Meltzer, Allan H. and Scott F. Richards. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89: 914–927.

Morris, Stephen and Hyun Song Shin. 2003. "Global Games: Theory and Applications." In *Advances in Economics and Econometrics*, ed. Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky. New York: Cambridge University Press.

Morrow, James D. 1994. *Game Theory for Political Scientists*. Princeton: Princeton University Press.

Morton, Rebecca B. 1999. *Methods and Models: A Guide to the Empirical Analysis of Formal Models in Political Science*. New York: Cambridge University Press.

Mueller, Dennis C. 2003. *Public Choice III*. Cambridge: Cambridge University Press.

Myerson, Roger B. 1991. *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press.

—. 1992. "On the Value of Game Theory in Social Science." *Rationality and Society* 4 (1): 62–73.

—. 1998. "Population Uncertainty and Poisson Games." *International Journal of Game Theory* 27 (3): 375–392.

—. 1999a. "Nash Equilibrium and the History of Economic Theory." *Journal of Economic Literature* 37 (3): 1067–1082.

—. 1999b. "Theoretical Comparisons of Electoral Systems." *European Economic Review* 43: 671–697.

—. 2000. "Large Poisson Games." *Journal of Economic Theory* 94: 7–45.

—. 2004. "Justice, Institutions, and Multiple Equilibria." *Chicago Journal of International Law* 5 (1): 91–107.

—. 2006. "Federalism and Incentives for Success of Democracy." *Quarterly Journal of Political Science* 1 (1): 3–23.

—. 2009. "Learning from Schelling's Strategy of Conflict." *Journal of Economic Literature* 47 (4): 1109–1125.

—. 2013. "Fundamentals of Social Choice Theory." *Quarterly Journal of Political Science* 8 (3): 305–337.

Nash, John F. 1951. "Noncooperative Games." *Annals of Mathematics* 54: 289–295.

North, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives* 5: 97–112.

North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49 (4): 803–832.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Osborne, Martin J. 1995. "Spatial Models of Political Competition under Plurality Rule: A Survey of Some Explanations of the Number of Candidates and the Positions They Take." *Canadian Journal of Economics* 28 (2): 261–301.

—. 2004. *An Introduction to Game Theory*. Oxford: Oxford University Press.

Osborne, Martin J. and Al Slivinski. 1996. "A Model of Political Competition with Citizen Candidates." *Quarterly Journal of Economics* 111: 65–96.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *Review of Economic Studies* 51 (1): 139–135.

Pepinsky, Thomas. Forthcoming. "The Institutional Turn in Comparative Authoritarianism." *British Journal of Political Science*.

Persson, Torsten and Guido Tabellini. 2000. *Political Economics*. Cambridge, MA: MIT Press.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton: Princeton University Press.

Przeworski, Adam. 1991. *Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America*. Cambridge: Cambridge University Press.

—. 2009. "Is the Science of Comparative Politics Possible?" In *The Oxford Handbook of Comparative Politics*, ed. Carles Boix and Susan C. Stokes. Oxford: Oxford University Press.

Riker, William H. 1980. "Implications from the Disequilibrium of Majority Rule for the Study of Institutions." *American Political Science Review* 74 (2): 432–446.

—. 1982. "The Two-party System and Duverger's Law: An Essay on the History of Political Science." *American Political Science Review* 76 (4): 752–766.

Roemer, John E. 2001. *Political Competition: Theory and Applications*. Cambridge, MA: Harvard University Press.

Romer, Thomas and Howard Rosenthal. 1978. "Political Resource Allocation, Controlled Agendas, and the Status Quo." *Public Choice* 33: 27–43.

Rubinstein, Ariel. 1991. "Comments on the Interpretation of Game Theory." *Econometrica* 59: 909–924.

—. 2006. "Dilemmas of an Economic Theorist." *Econometrica* 74 (4): 965–883.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

—. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Shepsle, Kenneth A. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science* 23 (1): 27–59.

— 1986. "Institutional Equilibrium and Equilibrium Institutions." In *Political Science: the Science of Politics*, ed. Herbert Weisberg. New York: Agathon.

—. 1991. *Models of Multiparty Electoral Competition*. Chur, Switzerland: Harwood Academic Publishers.

—. 2006. "Rational Choice Institutionalism." In *The Oxford Handbook of Political Institutions*, ed. R. A. W. Rhodes, Sarah Binder, and Bert A. Rockman. New York: Oxford University Press.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93 (2): 279–297.

Svolik, Milan W. 2013. "Learning to Love Democracy: Electoral Accountability and the Success of Democracy." *American Journal of Political Science* 57 (3): 685–702.

Tsebelis, George. 1990. *Nested Games: Rational Choice in Comparative Politics*. Berkeley: University of California Press.

Tucker, Joshua A. 2007. "Enough! Electoral Fraud, Collective Action Problems, and Post-Communist Colored Revolutions." *Perspectives on Politics* 5 (4): 535–551.

Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *American Political Science Review* 91 (June): 245–263.

—. 1998. "Political Institutions: Rational Choice Perspectives." In *A New Handbook of Political Science*, ed. Robert E. Goodin and Hans-Dieter Klingemann. Oxford: Oxford University Press.

Wittman, Donald A. 1973. "Parties as Utility Maximizers." *American Political Science Review* 67 (2): 490–498.