Rules and Rule Breaking, Institutions and Institutional Change

Abhinay Muthoo
University of Warwick

Kenneth A. Shepsle
Harvard University

March 2017

Abstract: The present paper begins with the well-traveled notion of institutions as "the rules of the game" and seeks to give some conceptual concreteness to the idea of rules in order to understand what it means to *break* the rules. While rules certainly channel and constrain behavior, and in principle require some actions and prohibit others, deviations from these requirements and prohibitions do occur. In addition, rules may be "interpreted," and the rules, themselves, do change – what humans devise they may revise. With all this conceptual movement, are the rules of the game too wobbly a foundation for the study of institutions?

Note to Ostrom Workshop Colloquium Series readers: This is, indeed, a "work in progress," inasmuch as potential blind alleys and false starts, as well as only the earliest version of a potential way forward, are in evidence in the present draft. The topic, moreover, is somewhat more abstract and foundational than is your normal fare. While I would welcome technical reactions and constructive advice, I am especially interested in the practical matter about how to wrap one's mind around "the rules of the game."

I. Why Is Rule Breaking Cool?

The study of institutions, institutional rules, and institutional change are subjects of great interest in contemporary economics, political science, and political economy. All of these fields have moved a considerable way away from highly individualistic, atomistic formulations as are found in social choice theory, general equilibrium theory, and behavioral approaches. Without abandoning the premise of instrumental, strategic, utility maximizing behavior, the newer studies appreciate the glue connecting individuals to one another as choreographed or scripted by institutional arrangements. An institution is conceived of as a body of rules, norms, and practices that describes "how things are done around here." As we shall see, it is often thought of as "the rules of the game," a view made famous by the late Douglass North. As he put it, "[Institutions] are perfectly analogous to the rules of the game in a competitive team sport. That is, they consist of formal written rules as well as typically unwritten codes of conduct that underlie and supplement formal rules" (North 1990: 4). They are, he contends (North 1990: 3), "…the humanly devised constraints that shape human interaction."

We say more on this and other understandings of institutions, but this is but a preface to our main purpose. We are intrigued not so much by the rule-following frame of an institution, but rather by the prospects for opportunism in institutional contexts, in other words for breaking rules. In this introductory section we articulate two reasons why we think rule breaking is "cool."

First, despite the role played by rules in constraining and coordinating individual behavior, rules are broken, sometimes in hugely consequential ways. Rule breaking is a real empirical phenomenon, perhaps even an empirical regularity. Consider these three examples[1]:

Speakers Brand and Reed and violations of procedure. In the late 19th century the British House of Commons and the U.S. House of Representatives confronted the persistent problem of minority obstruction. Majorities in the Commons, tolerating minority speech prerogatives, exposed themselves to a variety of obstructionist tactics by Charles Parnell, leader of the Irish MPs. After a long period in which the business of the Commons was severely compromised by the "total war" on the UK declared by Parnell, it reached a head in 1881. Its highly respected Speaker, Sir Henry Brand, nominally and actually an umpire and presiding officer rather than a partisan, concluded that the dignity of the House required decisive action on his part. In one specific instance of obstruction that had lasted all too long, he informed the prime minister that the next morning he would preempt further debate and call for a vote, which he did despite violating parliamentary requirements for cloture. A decade later and an ocean away, U.S. House Speaker Thomas Brackett Reed faced American-style legislative obstruction. The

[1] Each of these examples is developed more fully in Shepsle (2017).

Democratic minority tormented his Republican majority with tactics that denied the presence of a quorum and, even when one was present, annoyed the majority with dilatory motions and speeches that delayed voting. In 1890 the Speaker, firmly believing that the purpose of the majority is *to govern*, first instructed the Clerk to record those present in the chamber but refusing to answer to their names in a quorum call as nevertheless counting toward a quorum and, second, refused recognition to speak or to make a motion to anyone whose purpose he determined was dilatory. Neither the right to define a quorum nor the power to decide whether a person seeking recognition was doing so for legitimate legislative purposes was sanctioned by House rules and precedents. Both Brand and Reed broke the rules, did so knowingly, and, supported by legislative majorities, were not punished for doing so.[2]

Sulla and Caesar and the use of illegal force. The Roman general, Sulla, early in the first century BC of the Roman Republic, found himself at loggerheads with the Senate. His ability to prosecute the war against Mithridates, Rome's arch enemy, was severely hampered by shifting coalitions back in the capital that often restricted or opposed his efforts. Ultimately concluding that his frequent returns from the field to straighten out domestic politics were fruitless, he marched an army into Rome (contrary to long-standing tradition that prohibited an army within the city gates), exiled, imprisoned, or slaughtered many thousands of his opponents and their families (including the family of a young Julius Caesar), and proceeded to revise constitutional practices (all in violation of the existing constitution). Some forty years later, another Roman general, Julius Caesar, repeated Sulla's maneuver. Responding to the Senate removing him from his position as Governor of Gaul, and thus to the loss of his immunity from prosecution for charges against him during his prior service as consul, he marched an army across the Rubicon and into Rome, stacked the Senate with his followers, and had it name him dictator (ultimately "dictator for life" that, alas, did not last very long). The practices of the Roman Republic that had survived for five centuries were powerfully damaged, eventually leading to civil war and an end to the republican form.

LBJ and breaking the seniority norm. When Lyndon Johnson was elected minority leader of the U.S. Senate in 1953, he had been a senator for all of four years. He was perceived at the time as the mere deputy of the Senate Inner Club of senior southerners who dominated the body and its committees. These power brokers were especially dominant on the Senate's "big four" standing committees – Appropriations, Armed Services, Finance, and Foreign Relations. One of the norms of the Senate gave priority to senior legislators in committee assignment requests at the beginning of a new Congress. Southerners, with their

[2] Commons, shortly after Speaker Brand's maneuver, altered its Standing Orders to empower a speaker to control debate in the manner he had seized earlier. Likewise, several weeks after Speaker Reed's move, the House of Representatives adopted what came to be known as the Reed Rules, authorizing a speaker to conduct himself in the manner Reed had done.

seniority advantages, often sought to accumulate assignments to big-four committees. Johnson, seeking to broaden his support within the Senate but also with an eye on a larger national stage, enunciated what, sixty years later, is still known as the Johnson Rule: No senator serving on a big-four committee will be given priority in additional requests to those committees until the requests of more junior legislators (typically northern liberals) are accommodated. This breach of the seniority norm did not go unnoticed, but most southern senators at that time already possessed multiple assignments to big-four committees on the one hand, and accepted LBJ's reassurances on the other hand that this was not an all-out war on the norm.

The three examples – the violation of rules, "traditions," and norms, respectively – reveal that the institutional rules of the game, broadly construed, while typically observed and expected by others to be observed, are breached and, as the illustrations suggest, may involve long-standing, significant practices. Breaking rules, thus, are "cool" because they occur and temper the firm belief in the constraining power of institutions.

A second compelling reason to analyze rule breaking is the virtual absence of such analysis in the literature. The main reason for this is the association, developed more fully below, of an institution with a well-specified game form. The latter describes the flow of strategic interaction depicting, among other things, the identity of the players, when they may move, and what their behavioral options are at each opportunity. The idea of a player doing something other than what the rules specify as permissible is unimaginable – a "black swan." The rules are treated as fixed ex ante in most discussions, leaving no option for rule breaking. In some treatments, where there is the possibility of violating what an institutional arrangement requires or prohibits, the assumption of punishment sufficient to deter rule breaking is implicitly made. In other treatments there is recognition of the possibility of altering, if not violating, rules. Indeed, most bodies of rules explicitly provide for suspending or amending the rules – the "suspension of the rules" practice in the U.S. House of Representatives, the temporary waiving of rules as occurs against otherwise rule-permitting points of order in legislative assemblies generally, and officially recognized revision procedures as in Article V of the U.S. Constitution. However, these are not the same as breaking rules, but rather are instances of changing rules in line with rule-based procedures. There are also discussions in the literature of institutional change more generally. North (1990), for example, suggests that changes in relative prices (e.g., the effect of the Black Death on labor supply) or in preferences (e.g., against dueling) stimulate institutional change. But, again, rule breaking is distinctly different from these.[3]

[3] An exception to the absence of treatment of rule breaking in the literature is Williamson's (1975) discussion of "opportunism with guile," but he does not connect opportunistic behavior to rule breaking explicitly.

In short, rule breakings occur. They range from minor infractions like traveling 65 mph in 55 mph zones; to more extraordinary events like violating parliamentary rules and norms; to the cataclysmic as in the use of overwhelming force and other extra-legal means to violate, indeed overthrow, a constitutional order. In light of the relative neglect this receives in the literature on institutions and the possible exaggeration therefore of the latter's constraining effects, we elaborate on some of the causes and consequences in subsequent sections.

II. An Equivalence

Implicit in most treatments of institutions is an equivalence something like the following:

Institution ≡ **Rules of the Game** + **Physical and Technological Constraints** ≡ **Extensive Form Game**

Institutions are North's "humanly devised constraints" combined with natural constraints, which in turn are equivalent to a game form. Within its structure an institution induces equilibrium play by rational agents – what one of us called a *structure-induced equilibrium* (Shepsle, 1979).[4]

Deconstructing North's idea, which North himself is rather casual about and not always insistent on, an institution is an object, it is designed by humans, and it is constraining. First, it is conceived of as a body of formal rules and informal "understandings" of how strategic interaction is to proceed. No provision is made for the move of Speaker Brand to close debate without observing cloture procedures or of Speaker Reed to count quorums as he pleased regardless of the rules. To the contrary, the rules are assumed to be binding. Rule breaking should not be observed since it is out-of-equilibrium play.

Second, an institution is humanly devised, a view consistent either with a rules regime imposed on some by others or one chosen by the individuals actually governed by them. It is inconsistent with rules divinely supplied (e.g., Ten Commandments) or with emergent practices (though North nevertheless allows for the latter in the form of norms, taboos, customs, conventions, codes of behavior, or traditions).[5] For North, the canonical case is the self-governing group – constitution drafters, legislators, or shareholders/stakeholders, for example, who design constitutional, legislative, and organizational arrangements, respectively.

---

[4] Calvert (1993, 1995 a,b) and Schotter (1981) do not regard an institution as an object within which equilibrium play takes place. Rather, they treat persistent and regular optimal behavior in a primal environment as an institution. "Institution," in their view, is a *label* given to these regularities.

[5] One of the earliest to take a somewhat mystical view of institutions as emergent is Sait (1938) who likens them to coral reefs that emerge and change in inexplicable ways without apparent human agency. ("Who is the author of Roman Law?" he asks rhetorically.) North adopts some of Sait's macro-historical perspective, treating the manorial system, slavery, and the common law as institutions, despite the opacity of explicit human agency but without denying it as Sait would.

Third, institutional rules impose binding constraints. To travel from point A to point B, there are only so many permissible paths. Even if other paths exist, they are contrary to the rules. The ties bind. Left unresolved are instances of soft constraints that constrain but are not constraints in the strict mathematical sense; they are more like targets or, softer still, suggestions. Speed limits limit speed, but "speed creep" nevertheless occurs without institutional repercussions. Entry into the European Monetary Union in 1992 required that candidate countries meet explicit constraints on debt and deficits, but countries were admitted if they came "close." Whether constraints are soft or hard, the issue of exactly *how* they constrain is ambiguous. If behaving in a manner consistent with constraints is narrowly incentive compatible, then rationality is all that is required for constraints to be respected; indeed, constraints are superfluous. If, on the other hand, constraints are strictly binding, then the question remains of why individual agents comply with them. Expectations of some sort of what (bad things) will happen if violations occur are required.

Equating an institution with formal rules and practices conceived of as humanly devised and naturally occurring constraints (the first equivalence in the equation above) is a conceptual formulation. The second equivalence, characterizing an institution as an extensive game form, gives it analytical power. The rules specify (i) who may make decisions; (ii) when; (iii) what a designated agent may, must, or must not do when given an opportunity to move; (iv) the impact and timing of exogenous shocks (nature's moves); (v) what agents know; and (vi) when they know it. With this explicit treatment we also know what it means to break the rules – for example, moving out of turn, taking a forbidden action, or failing to take a required action.

III. Two Interpretations of the Extensive Form

Rubinstein (1991: 910) distinguishes two interpretations of the extensive form in game theory, one associated with Nash and another influenced by Aumann:

- According to the classical view attributed to Nash, a game form "…represents an exact and full description of the physical rules of a given situation."
- According to the revision attributed to Aumann, "…a good model in game theory has to be realistic in the sense that it provides a model for the *perception* of real life social phenomena. It should incorporate a description of the relevant factors involved, as perceived by the decision makers. These need not necessarily represent the physical rules of the world." (italics in original)

From the classical perspective rule breaking is not possible, since an accurate characterization of a circumstance takes account of all the things an agent might do at his or her opportunity to make a choice.

There may be a common understanding among the players that some of those things are "rule following" while others are "rule breaking," but these are just labels. Whether an individual chooses a rule-following action or a rule-breaking action depends not on these labels but rather on the consequences that flow from his or her choice, i.e., the subgame selected. A speaker can count a quorum however he or she pleases, whether or not "allowed" by the formal rules (however labeled). Consequences flow from this choice – the speaker's action may be supported by a decisive coalition; it may be rejected; the speaker may be forced to resign; he or she might be assassinated. The speaker takes these possibilities into account when making his or her choice. We might, in a colloquial fashion, refer to an action as breaking or following the rules. But this is no more than a linguistic convention.

Aumann's revised perspective in effect takes the "full" game tree and prunes it (see below). What remains are the "rules of the game." Again, rule breaking is not possible since the only actions available to an agent at his or her turn are the actions permitted by the rules. A speaker calling for a vote when the rules governing cloture have not been satisfied is unimaginable and cannot be accounted for. The speaker is playing a different game, one that allows for actions that the original game had excluded.

Thus, it seems to us, neither of these two interpretations of an institution as a game form provides a place for rule breaking, only optimal behavior. Yet we believe ordinary language and understandings require a distinction between rule following and rule breaking. To wrap our minds around this issue, we start with the game form approach, taking an abstract and foundational direction. In a second pass, we suggest a more substantive game-theoretic formulation.

IV. A Game Form Approach

Model. A game form represents, as in the classical view, a "full description" of a given situation. Define *pruning* and *rule breaking* as follows:

> Definition: Consider an extensive form game Γ in the Nash view. An extensive form Γ' is said to be a *pruning* of Γ if it is the result of removing branches from the nodes of Γ. Let $\Pi(\Gamma)$ be the set of all prunings of Γ, i.e, set of extensive forms derived from Γ by pruning.[6]
>
> Definition: For any $\Gamma' \in \Pi(\Gamma)$, let $B_{ij}(\Gamma')$ be the set of pruned branches from the set of opportunities that would otherwise be available to player i at node j of Γ. Player i is said to *break the rules of Γ'* at node j if he or she chooses an action from $B_{ij}(\Gamma')$.

[6] Clearly, $\emptyset, \Gamma \in \Pi(\Gamma)$, where the former is the extensive form with all branches pruned and the latter is the one with no branches pruned.

A pruning may be thought of as the removal of alternatives from the action set of players. Any agent who chooses an action available in this pruned version Γ' of the full-description situation Γ is said to be following the rules, whereas an agent who chooses one of the pruned actions is breaking the rules.

Reminiscent of the grammar of Crawford and Ostrom (1995), note that removing all but one action at node j in Γ is equivalent to requiring i to take that action; in order to be in compliance with the rules, he or she *must* do it. $B_{ij}(\Gamma')$, on the other hand, is the set of actions he or she *must not* do. If there are multiple actions at node j remaining after a pruning, they comprise those he or she *may* do. This gives us an operational definition of an institution:

Definition: An *institution* is a pruning, Γ', of a game, Γ, i.e., $\Gamma' \in \Pi(\Gamma)$

Referring back to our introductory examples, the rules and precedents of the U.S. House in 1890 provided a single acceptable way to count a quorum, namely to ascertain the number of legislators responding to their name in a calling of the roll and determining whether that number exceeds half the total number of elected legislators. It was physically possible for Speaker Reed to devise an alternative mode, as he did, but the choice he made was an element of $B_{ij}(\Gamma')$. Speaker Reed was a rule breaker, failing to do what he must and doing instead what he must not. Likewise, Sulla and Caesar might have complied with long-standing practice by bivouacking their armies on the hills outside Rome's city wall, but they chose to violate the practice. Finally, LBJ might have allowed the seniority practice regarding appointments to the big-four Senate committees to stand, maintaining them as senior preserves, but he deviated.

From any original Γ we can only speculate on the particular $\Gamma' \in \Pi(\Gamma)$ that arises. Schotter (1981) and Calvert (1995a) imagine a "primal environment" in which institutions become embedded. Relevant parametrizations of the primal environment include: who has the guns or votes or money or other forms of valuable human, financial, or physical capital; demographics (population size, age and gender distribution); physical resource endowments (land, minerals, climate, access to water and navigable rivers); and so on. Alternative prunings reflect various opportunities to exploit possibilities provided by the primal environment through cooperation, coordination, and communication. Some particular $\Gamma' \in \Pi(\Gamma)$ arises, or is chosen, or is imposed, and continues in place so long as no decisive coalition is inclined to and capable of replacing it with some $\Gamma'' \in \Pi(\Gamma)$. Comparative statics on the parameters of the primal environment may provide insights and intuitions about the durability of Γ' or the likelihood of some Γ'' replacing it. This is one plausible story. (We will address the issue of institutional change later in the paper.)

Alternatives to Rule Breaking**.** We note that rule breaking is not the only alternative to rule following. Imagine the temporal sequence given in Figure 1. At date 0 rules are in place in a primal environment, perhaps optimized to fit that environment for a group, a dominant individual or subgroup, or perhaps just carried over from date -1. The design reflects ex ante uncertainties that are not resolved until date 1. On date 2 the rules are applied to the specific state of the world that is now common knowledge. Even if optimized according to some criteria at date 0, the rules are optimal only in an ex ante sense. Once uncertainties are resolved the rules may not be ideal for the specific state that has arisen ex post. At date 3 the players are aware of a known set of primal parameters, a set of rules Γ' (a pruning of Γ), and a set of prohibited actions ($B_{ij}(\Gamma')$). Because the state in which the strategic interaction transpires is now

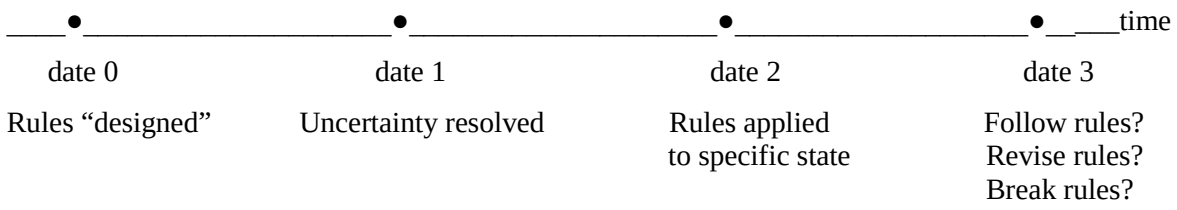

Figure 1

known, the players are also informed of their payoffs for any particular play of Γ'. Each individual must decide whether to follow the rules, to break the rules by taking a proscribed action in $B_{ij}(\Gamma')$, or to engage collectively in a third possibility – revising the rules (in accord with rules-based revision procedures). The latter involves a new design activity at date 3 that transforms Γ' into some Γ" □ Π(Γ) in a manner permitted by Γ'. In a legislative or committee context, for example, this new design activity may be a "motion to suspend the rules" (a one-off activity followed by a return to Γ') or a "motion to amend the rules" (in effect, a permanent move to Γ"). In a constitutional context it may be one of several mechanisms prescribed by Γ' to amend the extensive form. Revising the rules, then, is an alternative to following or breaking the rules.

Why Does Rule Breaking Occur? Putting revision to one side, though we shall return to the topic, we provide several reasons to expect rule breaking at least some of the time. First, equilibrium play under Γ' has distributional consequences. Some prosper while others fare less well. Either category of agent, but especially the latter, may aspire to do better (see Shepsle 2003). On the equilibrium path there is, by definition, no unilateral move within the rules to alter the distributional consequences. The failure to achieve distributional aspirations when playing according to the rules, then, may lead people astray. As North (1990: 5) remarked in terms of winning and losing, individuals will attempt "…to win the

game…by fair means and sometimes by foul means." Sulla devoted much time and effort, going back and forth from the field of battle to the Forum and its back rooms in Rome, in vain attempts to assemble a coalition supporting his war against Rome's external enemies. He tried within the rules and failed. Disadvantaged under existing rules, he moved outside them.

Inefficiency or the dysfunctionality of rules provide other grounds for rule breaking. The resolution of uncertainty at date 1 (see Figure 1), and the consequent identity of the circumstances and payoffs, may indicate that alternative rules in $\Pi(\Gamma)$ are more attractive (to whom?) than Γ'. A revision of the rules may be unavailable, leaving only unpleasant prospects under the existing rules regime or breaking from it. The Articles of Confederation in the newly established United States of America in the 1780s required supermajorities, indeed oftentimes unanimity, for actions and revisions. The prevalence of gridlock under this arrangement, especially as it applied to national security matters, led some to organize a convention in Philadelphia to scrap this scheme and create another *de novo*.

Breaking rules to secure distributional benefits, achieve efficiency gains, or overcome gridlock is often a prelude to rules reform. As noted earlier, part of the understanding between Speaker Brand and Prime Minister Gladstone was that after the Speaker's one-time violation of the rules, the government would proceed to revise the Standing Orders of Commons to permit a speaker more flexibility in controlling parliamentary debate. Similarly, after Speaker Reed's outside-the-box moves, he mobilized his partisan majority to revise U.S. House rules to enable its speaker to count quorums and control the floor more effectively. And both Sulla and Caesar engaged in constitutional revision after their respective power plays. We cannot claim that breaking rules is either necessary or sufficient for rules reform. Some revisions arise without being triggered by violations, and some violations fail to trigger reform.

We should not underestimate the complex calculation entailed in a rule-breaking decision. In principle, the choice of an action in $B_{ij}(\Gamma')$, rather than one available at the j^{th} node of Γ', takes the process to a subgame of Γ that had not been available in Γ'. This is the consequence of i breaking the rules at node j of Γ'. If information is complete, then all players are aware of this. However, it is stronger, often implausibly stronger, to assume players are completely informed not only about the game they are playing, Γ', but also about various subgames they might play if someone were to deviate in the play of Γ'. The deviating player surely has expectations of what is to follow her deviation, compared to her payoffs on the equilibrium path of Γ'. But we might be safe to assume that such expectations are less firmly grounded than those associated with playing by the rules. Hence, it is appropriate to think of rule breakers as institutional entrepreneurs, risk-takers, or outside-the-box thinkers, comfortable with rocking the boat, cutting corners, exploiting soft constraints, or proceeding into *terra incognita*. To give some concreteness to these claims, just imagine the calculations of those who chose to break away from the British empire or

the European Union, to secede from the United States in 1861, or some future secessionists to attempt to create independent homelands in Scotland, the Basque Country, or Catalonia.

Institutional Change**.** Institutions are rarely static for a number of reasons. First, designers often acknowledge the incompleteness of their knowledge – gaps in the contingencies that might arise. At the constitutional moment they often provide procedures to revise established institutions – constitutional amendment procedures, motions to suspend or amend legislative rules, legislative changes in the rules governing standing in court procedures, etc.

Second, designers may also anticipate the occurrence of rare events, though uncertain of their frequency or magnitude. Institutions are given flexibility to adjust. Contracts and treaties, for example, often possess "escape clauses" that free the parties from contractual obligations and provide for renegotiation. The Roman Republic, to give another example, invented the dictator institution to address unusual threats to order – external attack, domestic insurrections, epidemics, food shortages, etc. The Senate was empowered to appoint a dictator for a fixed term (usually six months) during which constitutional rights and privileges were suspended until the crisis passed. Many Latin American constitutions possess similar constitutional clauses (known as "regimes of exception" – see Loveman 1993).

Third, underlying parameters of the environment in which institutions are embedded undergo evolutionary or punctuated changes. Institutions adjust through revision or reinterpretation. In the United States, for example, the growing revenue requirements of the national government in the early twentieth century led to a constitutional revision freeing the government from the prohibition against taxing incomes. Constitutional reinterpretation played a role in response to the growing complexity of economy and society as industrialization proceeded: the Constitution's commerce clause came to be understood to incorporate an increasingly broad definition of interstate commerce, and thus a more wide-ranging jurisdiction for the U.S. national government at the expense of the states.

Changes in the primal environment generate "demand" for institutional change. Endogenous revision procedures, escape clauses, reinterpretation, and provision for emergency responses provide mechanisms for more or less permanent institutional change. The mechanisms available, however, may be inadequate in any given situation – too slow, too clumsy, too many obstacles – and these are circumstances vulnerable to rule breaking. Speakers Brand and Reed felt unable to rely upon the normal channels for change and determined it best to break the rules. U.S. presidents increasingly attempt to accomplish things off their own bats – by overreaching through their use of executive orders and signing statements – to compensate for ineffective legislative responses resulting from gridlock. Sulla and Caesar

lost patience with the corruption in Rome and took matters into their own hands. In sum, dysfunctional arrangements combined with mechanisms insufficient to address institutional failures make Γ' vulnerable to extra-legal maneuvers.

Discussion**.** The game form approach to rule breaking just outlined is suggestive but hardly complete or conclusive. We have described the set of possible institutions, Π(Γ), that result from pruning a (Calvert-Schotter) primal environment, Γ (see note 4). This partitions the set of physically available actions at each node into permissible and impermissible subsets. Rule breaking entails strategy selections in which some agents choose impermissible actions as part of their strategies, i.e., choices from $B_{ij}(\Gamma')$ for the institution Γ' □ Π(Γ). Because the institution, Γ', is given and commonly known, player strategies are based on the options the institution makes available to each of them. Rule breaking constitutes doing something not anticipated in this game, raising the vexing matter of exactly what an individual strategy actually is.[7]

There are several comparisons worth considering that constitute opportunities for further elaboration. First is the one between rule breaking and incomplete contracts. Incomplete contracts make provision for the unanticipated in a *procedural* sense. When an unexpected event occurs that affects the ability of contracting parties to perform as required, the contract specifies what to do or how to proceed; in effect it puts in place a method for renegotiation. Likewise, in constitutions (thought of as social contracts), unforeseen contingencies are resolved by a known third party, e.g., the Supreme Court, or through a contract revision process, i.e., constitutional amendment or convention. In terms of our earlier concepts, Γ' provides a subgame when an unanticipated move by Nature occurs. Consider this example. If a natural disaster were to occur on election day, making it impossible for citizens to vote, we imagine a process involving the high court would ensue, determining how to conduct the election in light of the unforeseen contingency. Rule breaking, however, is different. Even if there were no unforeseen contingency, an agent might select an action in $B_{ij}(\Gamma')$. Rule breaking constitutes a *violation* of the contract. An incumbent US government or one of the state governments that suspends an election scheduled for the first Tuesday after the first Monday in November of an even-numbered year violates the

[7] What makes thinking about strategies complex is the issue of how to deal with deviations from the rules that are considered ex ante impossible – so-called "black swans." No one anticipates them ex-ante; hence players have not made provision for such eventualities in their own strategies. This is not departing from rationality; it is adopting a perspective that people just can't or don't foresee zero-probability events. The question then arises as to what happens when the black swan turns up. Ex-post, players pick a new game form, but their ex-ante strategy – the one chosen *before* the black swan turned up – is not based on that. It is based on people following the rules. Of course, when we think about the player who actually contemplates breaking a rule, he or she will (but *only* he or she will initially) speculate what the new subgame will look like (and could mis-specify that of course). How to treat zero-probability events is a hard matter analytically and there is not much guidance from the literature of which we are aware.

rules. A government that refuses to step down after losing an election is a rule breaker (Przeworski 1991, 2006).

A second comparison is between rule breaking and lawlessness. Dixit (2004) develops the latter idea which emphasizes the manner in which governance of some sort proceeds in the absence (or inadequate functioning) of official institutions. Noting that "economic activity does not grind to a halt because the government cannot or does not provide an adequate underpinning of law," Dixit(2004: 3) explores the formation of alternative modes of governance – "groups and societies have much to gain if they can create alternative institutions to provide the necessary economic governance." His argument applies as well to conditions of "statelessness" characterizing the Hobbesian state of nature. His main interest is in what fills the vacuum. In terms of our formulation, conditions of "lawlessness" are described by the game Γ in the primal environment where anything physically possible is available to agents. Dixit is interested in the pruning $\Gamma' \in \Pi(\Gamma)$ selected in these settings. Rule breaking, on the other hand, is interested in a different issue. It takes a selected $\Gamma' \in \Pi(\Gamma)$ and examines the circumstances in which players will be tempted to move outside the range of actions it permits.

Each of these comparisons is suggestive and calls for further thought. There are several other points that also require elaboration:

- Rule breaking and the triggering of institutional change: We have hinted at occasions in which rule breaking often dramatically highlights anomalies and limitations in existing rules, and thus provides a spark for rules reform. But rule breaking would appear to be neither necessary – for example, constitutional amendments often are ratified without any prior violation – nor sufficient – legislative rules often remain on the books despite their regular violation (see Shepsle 2017: Chapter 8) – for institutional change.
- Rule breaking under conditions of asymmetric information: Calvert (1993) imagines a repeated prisoners' dilemma situation with private information. In each period, Nature reveals to each player his or her cost of playing cooperatively. If the cost is "too high," then a player defects. But costs are private information and so a "policy" of permitting players to defect – to violate the norm of long-term cooperative play – in high-cost circumstances is vulnerable to moral hazard. To overcome this problem, a player must "apologize" by behaving cooperatively for a fixed number of periods after his defection, while allowing his partner to defect, after which there is a return to the cooperative outcome. Calvert provides conditions under which rule-breaking followed by apology is part of equilibrium play. As he more generally notes, "Even in life's 'repeated' interactions, circumstances differ from one iteration to the next, often in ways in which the

players are asymmetrically informed. As a result, real-world norms and rules fostering social gains should take these variations into account and provide for exceptions and unavoidable violations. Under such institutions, actors actually carry out trade-offs, sometimes obeying rules and sometimes violating them and enduring the consequences." (Calvert 1993: 201)

- Rule breaking and the history of play: Both Speaker Brand and Speaker Reed tolerated minority obstruction for long periods before responding with rule-violating action. None of their predecessors had resorted to procedural abuses. Yet each had the support of decisive coalitions and suffered no retribution for their manipulation of "regular order." Is there something about the history of play that yields more tolerance of rule breaking?
- Empirically, is rule breaking a black swan? Finally, there is the empirical matter of the incidence of rule breaking. We have provided illustrations of rule-breaking occurrences throughout this paper, some quite dramatic, but we know little about their frequency or magnitude. What we hope to correct is a view prevalent in the literature on institutions that they structure and constrain individual interactions. Surely they do this, but not always and thus it would be informative to gauge the degree to which their rules and practices bind.

V. A Second Pass: Modeling Rule Following and Rule Breaking

The extensive form is an awkward mathematical object to analyze. While it sheds light on a rules-based conception of institutions, it doesn't take us very far, at least not yet. So, in this present section we propose an alternative approach, one taking more familiar modeling components on board. We have only just started, so the jury is out as to whether or not it is a way forward.

We have in mind a simple, dynamic model to study when rules remain unchanged through time, or instead when they are altered and, in this latter case, whether by design or by rule breaking. We model a given set of rules as a *distribution of bargaining power* among agents. We do this in the belief that a major impact of rules is on bargaining power which, in turn, determines whether deals are struck and who gets what.

We consider an organization with a finite number $N \geq 3$ agents, and which operates over an infinite number of periods $t = \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$ Each period t is divided into a number of dates. At the start of period t there are rules in place, denoted $\Gamma(t-1)$, that governed the organization's interactions in the immediately preceding period.

Date 1: Rules Maintained or Changed. Through a commonly known procedure, the rules for the current period are discussed and agreed upon, which we denote $\Gamma(t)$. If it turns out that $\Gamma(t) = \Gamma(t-1)$, then

we say the rules have been *maintained*. But if $\Gamma(t) \neq \Gamma(t\text{-}1)$, then we say that the rules have been *changed by design*.

Date 2: Rule Breaking? Given the outcome at date 1, any agent may attempt to *break* the agreed upon rules. If the agent succeeds, then the de facto rules under which business is conducted in this period will be different from that determined at date 1, say $\Gamma'(t) \neq \Gamma(t)$. Matters move to the next date, whether $\Gamma(t)$ or $\Gamma'(t)$ is now in place.

Date 3: Efforts/Investments. The agents simultaneously choose individual efforts, $e^i_t \geq 0$ ($i = 1, 2, \ldots, N$), where $C_i(e)$ is i's smooth, strictly increasing and convex, cost-of-effort function. The total surplus to be divided among the N agents is $V(e_t)$, where e_t is the period t vector of effort levels. We assume that V is strictly increasing and concave.

Date 4: Bargaining over Surplus Division. The N agents bargain over the surplus $V(e_t)$ according to the rules in place, $\Gamma(t)$ or $\Gamma'(t)$.

We have not developed this model further yet, but several brief points may be made. First, "in the beginning" there are rules. We say that at date 1 they are "discussed and agreed upon," but it really isn't necessary for either of these things to take place. The rules may be imposed by an agent external to the institution, or they may be carried over automatically from the past (as is the myth about the US Senate as a "continuing body," having never in its history adjourned but only recessed). Second, if someone determines at date 2 to break the rules and succeeds, then this implies that from the breach onward it is as if the organization were playing according to rules that differ from those chosen at date 1. Rule breaking, in other words, is construed as imposing different rules on the organization, and doing so in a manner different from the process by which rules are ordinarily selected. Third, we allow for rule breaking to occur *before* agent investment decisions are made, the thought being that a rules regime in effect – whether $\Gamma(t)$ or $\Gamma'(t)$ – will shape agent investment calculations. But it is plausible to allow for rule breaking *after* agent investment, so we might consider that modeling option. Finally, we have collapsed the activities of the organization into a bargaining framework. The idea here is that at the end of the day organizational politics is all about agents creating surplus and then jostling for advantage in claiming portions of it. When push comes to shove, the rules selected/imposed at date 1, or their violation at later dates, will have a bearing on who gets what.

Note to Ostrom Workshop Colloquium Series readers: We have not progressed beyond this point, which means any thoughts or advice you have will have considerable impact on how (whether?) we proceed.

References

Calvert, Randall. 1993. "Communication in Institutions: Efficiency in a Repeated Prisoner's Dilemma with Hidden Information." In Barnett, William A., Melvin J. Hinich, and Norman J. Schofield, *Political Economy: Institutions, Competition, and Representation* New York: Cambridge University Press. Pp. 197-223.

_____. 1995a. "Rational Actors, Equilibrium, and Social Institutions." In Knight, Jack and Itai Sened, *Explaining Social Institutions*. Ann Arbor: University of Michigan Press. Pp. 57-95.

_____. 1995b. "The Rational Choice Theory of Social Institutions: Cooperation, Coordination, and Communication." In Banks, Jeffrey S. and Eric A. Hanushek, *Modern Political Economy*: *Old Topics, New Directions*. New York: Cambridge University Press. Pp. 216-269.

Crawford, Sue E.S. and Elinor Ostrom. 1995. "A Grammar of Institutions." *American Political Science Review* 89: 582-600.

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

Loveman, Bryan. 1993. *The Constitutions of Tyranny: Regimes of Exception in Spanish America.* Pittsburgh, Pa: University of Pittsburgh Press.

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

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

_____. 2006. "Self-enforcing Democracy." In Donald Wittman and Barry Weingast (eds.), *Oxford Handbook of Political Economy*. New York: Oxford University Press.

Rubinstein, Ariel. 1991. "Comments on the Interpretation of Game Theory." *Econometrica* 4: 909-924.

Sait, Edward McChesney. 1938. *Political Institutions: A Preface*. New York: Appleton-Century.

Schotter, Andrew. 1981. *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.

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

_____. 2003. "Losers in Politics (and How They Sometimes Become Winners): William Riker's Heresthetic." *Perspectives on Politics* 1: 307-315.

_____. 2017. *Rule Breaking and Political Imagination*. Chicago: University of Chicago Press.

Williamson, Oliver. 1975. *Markets and Hierarchies*. New York: Free Press.