

David M. Woodruff  
Harvard University  
dmwoodr@fas.harvard.edu

**Understanding Rules and Institutions: Possibilities and Limits of Game Theory**

“Roundtable: Ethnography Meets Rational Choice: The Case of David Laitin, For Example”  
Annual Meeting of the American Political Science Association  
Washington, DC, September 1-4, 2005

Is rational choice theory compatible with, and useful to, ethnography, which I’ll take to be the interpretation of meaningful action? For an affirmative answer, one might look to Fearon and Laitin’s famous 1996 article on “ethnic peace.” Their argument ran as follows: one way peace between two ethnic groups can be preserved is if each group punishes its own members for bad behavior toward the other group. Such “in-group policing” is effective, they suggest, because people usually have better intelligence about the doings of members of their own ethnic group. Thus members of an ethnic group are in a position to reliably punish just those of their co-ethnics who have behaved badly in inter-ethnic interactions. By contrast, to the extent that people have a hard time identifying poorly behaved individuals who are not members of their ethnic group, they will only be capable of indiscriminate punishment of all the transgressors’ co-ethnics. Such punishment may also deter bad behavior, but is more likely to lead to a spiral of violence. Fearon and Laitin capture these two possibilities in the form of two distinct equilibrium strategies in a repeated prisoner’s dilemma game involving both intra-ethnic and inter-ethnic interactions.

To illustrate the real-world relevance of this argument, Fearon and Laitin provocatively mobilize an anecdote from the *locus classicus* of ethnography, Geertz’s essay on “Thick Description.”

Geertz relates that in early colonial Morocco, a marauding band of Berbers attacked the home of a Jewish trader in the Maghrib named Cohen. He survived

but his guests were killed and his goods stolen. Cohen could get no help from the French authorities, but he belonged to a *mezrag*, or trade-pact system, and he went to his insurance broker, a tribal sheikh, to demand the assistance due. The sheikh knew precisely who had Cohen's merchandise, accompanied him in a climb up the Atlas directly to the shepherd of the thief's tribe, and took control of the entire herd. The tribal warriors soon returned, saw what had transpired, and prepared to attack. But then they saw Cohen and his insurance agent, a palaver began, and Cohen peacefully regained his goods at the precise insured value. [Cohen was given sheep meant to correspond to "four or five times" his loss (Geertz 1973, 8).] Note that 'on the equilibrium path' this institutional innovation of tribal 'information brokers' would make mutually beneficial trade relationships between Jews and Berbers possible, despite problems of opportunism due to a low density of social network relations. And, in the case Geertz relates, the institution also prevented spiraling, here understood as a total breakdown of trading and relations between Jews and Berbers (Fearon and Laitin 1996, 728).

Thus, Fearon and Laitin see this as a clear example of the "in-group policing" mechanism; they imply that the sheikh's action was motivated by incentives resembling those described in their model. However, let's consider the other, extremely brief, account of the *mezrag* in operation provided in Geertz's essay:

To make a trade pact in Morocco, you have to do certain things in certain ways (among others cut, while chanting Quranic Arabic, the throat of a lamb before the assembled, undeformed, adult male members of your tribe) and to be possessed of certain psychological characteristics (among others, a desire for distant things). But a trade pact is neither the throat cutting nor the desire, though it [the pact] is real enough, as seven kinsmen of our Marmusha sheikh discovered when, on an earlier occasion, they were executed by him following the theft of one mangy, essentially valueless sheepskin from Cohen (Geertz 1973, 12).

So in the two cases we see (1) the sheikh imposing material (sheep-denominated) damages for a large theft and two murders and (2) the sheikh executing seven people for a trivial theft. It seems plausible that in case (2) something other than the protection of Cohen's property was at stake—perhaps the sheikh had contextual reasons of his own to show an iron fist to his tribesmen. And there's no reason that something else couldn't be at stake in case (1), as well—perhaps it was contextually important for the sheikh to intimidate the tribe in question (he did take a group of

armed men with him when he went to capture the sheep). There is nowhere near enough evidence to sustain either of these interpretations—but there's not enough to impugn them, either.

Any rational choice argument is built from two pieces: a vision of how people make choices, and a depiction of the circumstances under which those choices are made. The alternate interpretations of Geertz's anecdotes I just presented rely on a presumption of rationally chosen action governed by (broadly understood) incentives (payoffs). With regard to how people make choices, the interpretations do not differ. They differ, instead, in their *depiction of the circumstances of choice*. While Fearon and Laitin relate incentives to repeated situations regulated by the trade pact, the plausible alternative interpretation describes incentives particular to individual situations where the rules of the trade pact happen to be in play. To use Weber's terms, Fearon and Laitin focus on *formally* rational action, that is, on action motivated by bringing particular cases under general rules. The alternative contextual explanations describe *substantively* rational action, driven by a balance of considerations in a particular case.

I wish to make three points, focusing on understanding the intersection of rules (like those of the trade pact) and context. First, as just illustrated, the incentives that shape people's actions in relation to a single rule can be different in different situations. People can do the same thing—obey a rule, or enforce a rule, or violate a rule—for different reasons. The mere articulation of a rule that could allow formally rational action does not ensure that formal rationality governs in practice. My second point is that the potential relevance of case-specific incentives to rules has important empirical implications, affecting how we conceive the processes that allow institutions to create order. Third, even those who accept rational choice as a working assumption should conclude that game-theoretic methods of *describing situations*

have little advantage when there are distinct contextual incentives for conformance to rules in distinct circumstances. Game theory is only powerful when the real-world incentives for conforming to rules are *general* and formal rationality governs. When substantive rationality looms large, game theory is at best not very powerful; at worst, the presumption that game theory will be powerful can obscure some of the key processes by which order gets built.

### Contextual incentives and “cellular” order

On reflection, it’s not very surprising that contextual, or case-specific, incentives can shape attitudes to rules. There are at least two ways this can happen. The first way is straightforward: the costs and benefits of conforming to or violating rules can vary across cases. Imagine Betty has contracted to buy sugar from Steve, with payment on delivery. While the sugar is still in transit, the price of sugar drops radically. How does Betty decide whether to pay or whether to weasel out of the contract? Let’s give a simple picture of her incentives:

<i>Action</i>	<i>Payoff</i>
Fulfill contract	sugar – contracted price
Weasel out of contract, buy sugar at lower price	sugar – market price – costs of weaseling

In other words, she pays when

market price + weaseling costs < contracted price

The incentives in this case are contextual. The benefits of violating the rule (weaseling on the contract) depend on the difference between the contracted price and the market price, which may or may not be outweighed by the legal costs.

Actually, there’s no reason to stop with price swings. There can be an arbitrarily large number of contextual factors that determine the cost of weaseling to Betty (Commons 1957, 65). Does Betty have alternate suppliers? Does Steve have alternate customers? Are Betty and Steve

linked in a kinship or religious network? Are there tax implications to paying or not paying? Does Steve's business have enough working capital to wait out a lawsuit, or not? Is Betty's firm under pressure from other creditors? All of these things will determine incentives relevant to Steve's and Betty's attitudes to the rule in any particular transaction.

In effect, Betty's bill-paying behavior will reflect *multiple* equilibria in *multiple* games, in which the rules of contract are simply a single element. Some of these games will differ only in the payoffs to different actions; others will feature moves, such as shifting suppliers, not always available. There can be an arbitrarily large number of such games, insofar as there is no guarantee that incentive structures and available moves repeat across transactions.

So far we've been considering cases where contextual incentives affect the costs and benefits of conforming to rules. A second, more complex form of interaction between rules and context occurs when a rule makes achievement of some contextual goal more or less expensive. It is important to remember that actions *consistent* with a rule are not always motivated by the general purposes behind the rule's creation. In other words, substantive rather than formal rationality can prompt the invocation of rules. Not long ago I was having lunch with another political scientist in Somerville, at a place that had tables out on the sidewalk. While we were eating, a man sat down at another table. He was middle-aged, light brown in color, and began muttering to himself as he sorted through the unrecognizable contents of two plastic supermarket bags. Within a few minutes the restaurant's cashier emerged to explain to him that tables were only for customers, and he apologized and left. Not much later, another person sat down at another nearby table—there were several available. This man was white, and looked old and tired. He did not mutter to himself or sort through his bags, but just sat for some minutes resting. The cashier did not emerge to run him off. It seems likely that there was some substantive

difference between these two cases that motivated the invocation of the formally rational rule in one case but not in the other. In this light, it would be inappropriate to explain the invocation of the rule solely by the rule's terms. Similarly, when Geertz's sheikh had seven relatives killed as punishment for a trivial theft, the formal rules of the trade pact may well have provided a convenient occasion or pretext for a display of dominance needed for other purposes.

To summarize, then, a particular rule can intersect with contexts of varied incentives and opportunities. One potential contextual incentive is that the rule provides an economical means of achieving an aim—running off undesirables, intimidating challengers—that would have been sought even if the rule did not obtain. In these sorts of cases, game theory, I think we'd have to admit, isn't very powerful. All it amounts to is a kind of a protocol for writing down the results of ethnographic research in what could be a huge number of different models.

This isn't just a methodological point; it also has important implications for how one understands the way institutions create order. When incentives are contextual, order is *cellular* (cf. Stinchcombe 2001, 84, 97). General adherence to any rule reflects adherence to the rule in a multitude of "cells," each structured by distinct incentives and opportunities that interact with the rule in distinct ways. Things like life histories, relationships, and wealth—and, most importantly, any number of situation-specific goals that can't be brought under such general headings—matter greatly to the effectiveness of the rule. When incentives are contextual there can and often will be social situations where a rule is perfectly effective in guaranteeing order alongside others where the same rule, backed by the same state enforcement capacity, does not lead to order. Let's go back to our sugar-purchase example. Suppose Betty notes that her competitor Bob tried to weasel out of a contract with Simone, but wound up losing a court case. Does she conclude that she has no alternative but to accept delivery of sugar at an above-market

price? Well, it depends. What are Steve's incentives and capacity to wait out a lawsuit? How convincing is the technicality on which Betty plans to rely compared to that which Bob tried? And so on and so forth. Again, the same rule "fulfill contracts" may guarantee order for situations like Bob and Simone's but not like Betty and Steve's. Order obtains in some cells and not in others. Under cellular order, formally rational considerations are always potentially trumped by substantively rational ones.

### **General incentives and "broadcast" order**

It is not my intent to claim that actors encounter all rules in contexts that create idiosyncratic incentives. Some kinds of rules are associated with general incentives, which are the same any time an actor encounters a rule, whatever her other circumstances. Consider driving on a divided high-speed highway like America's Interstates. Rules stipulate that on each side of the highway travel is only permitted in one direction. These rules are violated in a vanishingly small collection of cases. The incentive to travel in the right direction is that failure to do so involves an extremely high risk of a dangerous collision with oncoming drivers. This incentive is essentially general—not getting killed pretty much trumps any other considerations that one might have, such as getting to work faster, etc. Follow the rule, or else. Unlike the case of contextual incentives, with general incentives formal rationality and substantive rationality always both command the same action. .

When incentives are general, game theory is powerful. It provides a compact way of expressing the structure of common choice situations. For instance, one can model the driving rule just discussed quite easily and convincingly as a coordination game. Again, the power of game theory here has empirical, not just methodological, implications. General incentives make order creation a "broadcast" phenomenon structuring the acts of general classes like drivers,

rather than of individuals characterized by idiosyncratic incentives. Drivers' life histories, relationships, social status, wealth, desire to get to work on time, etc., are irrelevant. In effect, application of punishment in the form of crashing is entirely impersonal; so too is the reward of the much higher chance of reaching one's destination safely driving in the right direction.

There are cases of "broadcast" order-creation that are extremely relevant, even central, to comparative politics. For instance, there's clearly a difference between states in which organized violence by non-state actors must be conspiratorial and those where it need not be, and it would be both straightforward and convincing to give a game-theoretic account of this in terms of general incentives. (Unless there are a lot of publicly operating armed bands, any particular one will quickly be liquidated by a reasonably strong state, so getting to lots of such bands is a classic coordination problem that is hard to solve.) Likewise, some aspects of money and property, capitalism's central institutions, rest on broadcast mechanisms of order. Or, to take an example from Professor Laitin's work, the structuring of choices on language can often be understood via coordination models that have a broadcast character. Nevertheless, rules surrounded by contextual incentives—affecting those subject to the rules or those enforcing them—are ubiquitous, and so too is cellular order.

### **Is theoretical modesty an adequate reaction to contextual incentives?**

So far, I haven't challenged at all the idea that game theory could explain adherence to and violation of rules. All I've done is to argue that adherence to some rules can be modeled in a single game while adherence to other rules will require multiple games to explain. I've also argued that this line between compact and sprawling game-theoretical accounts also marks the border between broadcast and cellular order.



But isn't this dichotomy between "individual games for all the (many) contextual incentives relevant to a rule" and "one game describing general incentives to adhere to a rule" overly stark? Isn't there some middle ground on offer? After all, the task of social science would seem to involve transcending case-by-case storytelling (even storytelling employing the idiom of game theory).

I don't think the middle ground here is attractive. Cellular and broadcast order simply work too differently, meaning efforts to reconcile the sorts of contextual incentives characteristic of cellular order with the demands of tractability or compact modeling have high analytic costs. Consider the approach to contextual factors Fearon and Laitin take in their paper on ethnic peace. As they emphasize at several points, their "claim is not that the mechanisms we have identified are the only ones that matter, but that they have not been clearly identified before and do explain a part of the empirical puzzle"(Fearon and Laitin 1996, 727). Fearon and Laitin also point out specific aspects of context, such as state-building, that they have neglected.

Modesty of this sort greatly complicates straightforward verification of the argument with ethnographic evidence. Suppose we observe my hypothetical sugar-buyer Betty fulfilling a contract in a situation where the state has committed to punish her if she does not. Would someone offering a "partial" game-theoretic explanation of contract compliance as reflecting state enforcement be justified in concluding that Betty's action supports the argument? No! For it could be that the price differential was low, Steve was in a position to seek alternate customers, reputational costs were high, no more insistent creditors were on the scene, etc. Until all these other potential influences on Betty's decision were measured, her decision to fulfill the contract would say nothing whatsoever about whether enforcement explains "part of the empirical puzzle." Action merely consistent with the proposed partial explanation provides no support for

that explanation, insofar as the action is consistent with a number of other plausible explanations. (Indeed, action inconsistent with the proposed partial explanation does not impugn it, either; it could be that under other conditions enforcement would have been decisive in Betty's decision-making.)

Sorting out these issues requires a careful ethnography that seeks contextually relevant counterhypotheses for what was "off the equilibrium path," i.e., what other alternatives were open to Betty and why she did not choose them. The game-theoretic model might be good preparatory work for such an ethnography, insofar as it offers a plausible possibility for what was off the equilibrium path. But the claim of a partial explanation will collapse into a *de facto* negation of the importance of context if acts consistent with the model are taken as evidence of its operation. This temptation to misapprehend substantive rationality as formal rationality—to mistake cellular for broadcast order—is a significant analytic cost of offering game theory as merely a partial explanation.

## **Conclusion**

I have argued that modeling of institutions as equilibrium strategies in a repeated game is effective, in some circumstances—those characterized by general incentives and broadcast order. Here formal rationality and substantive rationality never work at cross purposes, meaning, among other things, that formal rationality cannot serve as an excuse for decisions made on substantive grounds. However, a danger arises when susceptibility of rules to compact game-theoretic modeling is assumed when incentives are in fact contextual. For if it is cellular order that a rule creates—if the rule determines behavior in some contexts and not in others, or if the rule can be mobilized as a convenient excuse—the correct game-theoretic depiction of the rule's operation would involve multiple games with distinct incentives and different available moves.

Here the desire for a tractable model conflicts with the desire for an accurate one; when tractability wins out, the importance of context is negated and ethnographic precision becomes elusive. Modernization theory was criticized for arguments amounting to the suggestion that all substantive rationality empirically gives way to formal rationality (a generalization, by the way, that would have been positively anathema to Weber!). Unless it takes context seriously, the game-theoretic account of institutions threatens will apply this same mistaken conclusion not just to “modernity,” but to history far more generally.

One final point. The argument presented here has sought to demonstrate that there may be no easy way to transform a rule into a tractable game-theoretic model explaining the rule’s affect on practice. But in many parts of the argument, one could easily replace “conformance to the rule” with any other explanandum of interest to political scientists—civil war, revolution, democracy, authoritarianism, tax incidence, etc. The serious difficulties facing a game-theoretic analysis of rules bespeak even more serious difficulties in applying the approach to these more complex explananda. For this reason, it would be a mistake to build comparative politics around the assumption that tractable game-theoretic models can provide satisfactory answers to our enduring questions.

## References

- Commons, John Rogers. 1957. *Legal Foundations of Capitalism*. Madison, Wis.: University of Wisconsin Press.
- Fearon, James D. and David D. Laitin. 1996. Explaining Interethnic Cooperation. *American Political Science Review* 90, no. 4: 715-735.
- Geertz, Clifford. 1973. Thick Description: Toward an Interpretive Theory of Culture. In *The Interpretation of Cultures; Selected Essays*:1-30. New York: Basic Books.
- Stinchcombe, Arthur L. 2001. *When Formality Works: Authority and Abstraction in Law and Organizations*. Chicago: University of Chicago Press.