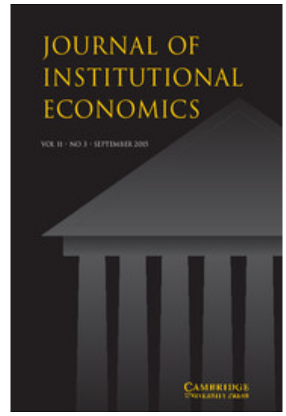

Institutions, rules and equilibria: a commentary

KEN BINMORE

Journal of Institutional Economics / Volume 11 / Issue 03 / September 2015, pp 493 - 496
DOI: 10.1017/S1744137414000599, Published online: 30 December 2014

Journal of Institutional Economics (2015), 11: 3, 493–496

doi:10.1017/S1744137414000599
First published online 30 December 2014

Institutions, rules and equilibria: a commentary

KEN BINMORE*
Philosophy Department, Bristol University, Bristol, BS8 1TN, UK

Abstract: This brief note is a commentary on Hendriks and Guala's (2014) unification of the institutional theories of Lewis, North, and Searle. It argues that the equilibrium theory of Lewis is fundamental and that the kind of equilibrium best suited in this role remains the orthodox notion of Nash.

In their very successful theory of institutions, rules, and equilibria, Hindriks and Guala (2014) unify three approaches to the study of institutions: an equilibrium theory attributed to Lewis (1969); the rule-based theory of North (1990); and Searle's (1995) constitutive-rules theory.

The synthesis of the first two theories into what they call the rules-in-equilibrium approach is not new (Greif and Kingston, 2011) but they offer a fresh and instructive exposition. Their main purpose is to reconcile the constitutive-rules theory with the rules-in-equilibrium approach. A translation of the language of the former into the language of the latter is proposed, which seems very reasonable to me. However, a better assessment of the relative status of the three theories would have been helpful. Everybody knows the verse:

Big fleas have little fleas
Upon their backs to bite 'em
And little fleas have lesser fleas
And so, ad infinitum.

To understand this flea ecology one needs to take account of the fleas of all sizes, but the role of the smaller fleas cannot be understood without appreciating that they ride upon the back of larger fleas. I see equilibrium as the big flea in Hindriks and Guala's ecology. On its back rides North's theory of rules. If Hindriks and Guala are right, Searle's theory about the grammar of rules is either a companion or a subordinate flea to North's.

I gratefully acknowledge support from the European Research Council Seventh Framework Program (FP7/2007–2013), ERC Grant agreement no. 295449.
*Email: k.binmore@ucl.ac.uk

Such a prioritization matters because an institution needs to be in some kind of equilibrium to survive, and survival is fundamental. Rules are necessary but subsidiary (Binmore, 2010). As Hindriks and Guala explain, rules are needed because players unknowingly operating an equilibrium in a game of which they are unaware need rules to follow if an equilibrium is to be sustained.

The players will not usually understand why the institutional rules that govern their behavior hold together. Otherwise they would not invent rationalizations of the type that Searle seeks to capture. But social analysts need to look deeper if we are to escape from the kind of work that continues to proliferate in those parts of political economy and constitutional economics in which large conclusions are drawn from simple mechanism-design models that have been put together with little or no thought as to why anybody should obey the rules that supposedly constrain a citizen's behavior. In a constitutional setting, there is nobody outside the system to enforce the rules. The only viable rules are therefore those that enforce themselves because they call for equilibrium play in an underlying game of life whose rules are enforced by Nature herself. In brief, the answer to the classical question of who guards the guardians is that they must guard each other.

How do we model the game of life? For the sake of simplicity, Hindriks and Guala confine their attention to simple coordination games but the underlying game appropriate to the study of political institutions will usually be at least as complex as a repeated game (Guala, 2012). How are such games to be studied? There are two approaches: rational and evolutionary. I think Lewis (1969) led us astray by focusing on the rationality option,[1] which led him to the indefensible claim that workable conventions need to be common knowledge (Binmore, 2008).[2] It is evolutionary game theory that matters when studying institutions (Young, 2001). To argue otherwise is to neglect the role of history – to side with Thomas Paine rather than Edmund Burke.

What kind of equilibrium? The orthodox answer is Nash equilibrium – the weakest notion which ensures that the outcome be the end-product of an evolutionary process (usually cultural in an institutional context). Hindriks and Guala muddy the waters a bit here by using Maynard Smith's (1982) Hawk–Dove game as an example. Like most biologists, Maynard Smith talks about evolutionary stable strategies (ESS) rather than Nash equilibria. This is fine in symmetric games with only two strategies like the Hawk–Dove game, but simple evolutionary processes can easily converge on Nash equilibria that are not ESS

1 The credit for the theory of conventions arguably belongs in any case largely to David Hume (1739) and Tom Schelling (1960).

2 If so, it would not be conventional to speak French in France! Incidentally, the dispute over whether Lewis (1969) or Aumann (1976) properly deserves credit for the concept of common knowledge has recently been resolved in favor of the sociologist Friedell (1969).

in games with more than two strategies (Hofbauer and Sigmund, 1998). Worse still, a repeated game lacks any pure ESS at all.[3]

In rational game theory, it is sometimes proposed that Nash equilibria be replaced by correlated equilibria. A correlated equilibrium of a game G is just a Nash equilibrium of a new game H constructed from G by appending an opening move in which Nature sends correlated signals to each player on which they can condition their choice of strategy in G. In endorsing the correlated equilibrium concept, Hendriks and Guala use Maynard Smith's example of the bourgeois Hawk–Dove Game that begins with Nature telling each player who is the 'owner' of the resource in dispute in the original Hawk–Dove game.

I have devoted a good part of my life to studying fairness as a coordinating device (Binmore, 2005). I therefore agree very much with the emphasis Hendriks and Guala give to correlation, but I feel that correlation devices are so important that their structure needs to be laid on the table rather than hidden from view in the abstract definition of an equilibrium concept. That is to say, I think we need to talk about Nash equilibria of the game H rather than correlated equilibria of G. One is then forced to be completely explicit about the coordinating device that opens the game H. The danger in restricting attention to correlated equilibria of G in rational game theory is that one may inadvertently invent ways for the players to get along together that are hopelessly impractical or historically impossible.[4] In evolutionary game theory, the situation is worse since evolutionary processes in H can be much more flexible than in G because of the possibility of evolutionary drift off the equilibrium path (Samuelson, 1997).

References

Aumann, R. (1976), 'Agreeing to Disagree', *Annals of Statistics*, **4**: 1236–1239.

Axelrod, R. (1984), *The Evolution of Cooperation*, New York: Basic Books.

Binmore, K. (2005), *Natural Justice*, New York: Oxford University Press.

Binmore, K. (2008), 'Do Conventions Need to be Common Knowledge?', *Topoi*, **27**: 17–27.

Binmore, K. (2010), 'Game Theory and Institutions', *Journal of Comparative Economics*, **38**: 245–252.

Brandenburger, A. (2014), *The Language of Game Theory: Putting Epistemics into the Mathematics of Games*, Singapore: World Scientific Publishing.

Friedell, M. (1969), 'On the structure of shared awareness', *Behavioral Science*, **14**: 28–39.

Greif, A. and C. Kingston (2011), 'Institutions: Rules or Equilibria?', in N. Schofield and G. Caballero (eds.), *Political Economy of Institutions, Democracy and Voting*, Berlin: Springer.

Guala, F. (2012), 'Reciprocity: Weak or Strong? What Punishment Experiments do (and do not) Demonstrate', *Behavioral and Brain Sciences*, **35**: 1–15.

3 Axelrod's (1984) 'proof' that tit-for-tat is an ESS is not correct.

4 The set of all correlated equilibria can be large even when G is quite simple – essentially the same as the result of successively deleting dominated strategies (Brandenburger, 2014).

Hindriks, F. and F. Guala (2014), 'Institutions, Rules, and Equilibria: A Unified Theory', *Journal of Institutional Economics*, Published online: 16 October 2014, DOI: 10.1017/S174413744000496.

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

Hume, D. (1739/1978), *A Treatise of Human Nature*, Oxford: Clarendon Press.

Lewis, D. (1969), *Conventions: A Philosophical Study*, Cambridge, MA: Harvard University Press.

Maynard Smith, J. (1982), *Evolution and the Theory of Games*, Cambridge: Cambridge University Press.

North, D. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.

Samuelson, L. (1997), *Evolutionary Games and Equilibrium Selection*, Cambridge, MA: MIT Press.

Schelling, T. (1960), *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.

Searle, J. (1995), *The Construction of Social Reality*, New York: The Free Press.

Young, P. (2001), *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton: Princeton University Press.