Games and Economic Behavior 112 (2018) 125–138

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

King of the Hill: Giving backward induction its best shot ☆

Martin Dufwenberg [a,b,c], Matt Van Essen [d]

[a] University of Arizona, United States
[b] University of Gothenburg, Sweden
[c] CESifo, Germany
[d] University of Alabama, United States

ARTICLE INFO

Article history:
Received 31 October 2017
Available online 13 August 2018

JEL classification:
C72
C92

Keywords:
Backward induction
Interactive epistemology
Player set cardinality
Experiment

ABSTRACT

We study a class of deceptively similar games, which however have different player sets and backward induction (BI) predictions that vary with their cardinality. The game-theoretic principles involved are compelling as predictions rely on weaker and less controversial epistemic foundations than needed to justify BI more generally. Are the BI predictions empirically relevant for this class of games? We design and report results from a relevant experiment.

1. Introduction

We offer two independent, and arguably equally important, motivations:

MOTIVATION #1

Some classes of games can be meaningfully parameterized by the cardinality of the player set (N), and shown to possess properties that depend in interesting ways on N. For example, the class of N-player Cournot games nicely links the cases of monopoly ($N = 1$) and perfect competition ($N \to \infty$).

We explore a class of N-player games where predictions systematically vary with N in a different and intriguing way. The following problem illustrates:

> Consider N "subjects" in a line, in front of a "king" on a throne. The subject first-in-line must choose whether or not to dethrone the king. If not, the game ends (and all subjects go home). If the subject dethrones the king then he becomes the new king. The subject next-in-line must now choose whether or not to dethrone the new king. If not, the game ends. If the subject dethrones the king then the subject becomes the new king, and the subject next-in-line must choose

☆ Part of the research was done while MD was on faculty at Bocconi University. We are grateful to that institution for providing funds for the experiment. We thank Geir Asheim, Pierpaolo Battigalli, Amnon Rapoport, Peter Norman Sørensen, referees & editors, and the participants at several seminar and conference presentations for helpful comments and discussion.

E-mail addresses: martind@eller.arizona.edu (M. Dufwenberg), mjvanessen@cba.ua.edu (M. Van Essen).

https://doi.org/10.1016/j.geb.2018.07.007
0899-8256/

whether to dethrone, etc. The interaction continues until some subject does not dethrone the sitting king, or until there is no subject in line. The most preferred outcome is to become a king who is not dethroned. Second best is to remain a subject. The worst outcome is to be dethroned. Will the original king be dethroned?

[*Stop and think before reading on!*]

This is an old problem which, however, seems little-known. Brams and Kilgour (1998; footnote 5) describe one version, and a Google search reaches others (often with the players being lions & lambs instead of subjects & kings). One of us learned about it from Jacob Goeree 20+ years ago. Casual empiricism (try it on friends & colleagues!) suggests most people never heard of it, and find it hard to see through the thicket. However, reasoning by backward induction (BI) one realizes that the solution exhibits an odd–even effect. The original king will be dethroned if N is odd, not dethroned if N is even.

We believe there is applied potential. If the king were a sovereign state and the subjects hostile neighbors we might get an example concerning geopolitical stability. For another example, consider voting procedures, where N parties or individuals sequentially reject and propose budgets.[1] We conjecture that parallel problems may also arise in societies with weak property rights (cf. Kaplow and Shavell, 1996; Bar-Gill and Persico, 2016), where agents may take each others' goods. Odd-even effects arise also in behavioral models of intertemporal choice, e.g. $\beta\delta$-models of procrastination (O'Donoghue and Rabin, 1999).[2]

To the best of our knowledge, no one explored the empirical relevance of odd-even effects. Using a variety of king of the hill (KOH) games, that match versions of the story we told, we design a series of lab-experiments to tackle this task.

Motivation #2

Among scholars who worked on the epistemic foundations of game-theoretic solution concepts, BI (in extensive games of perfect information) is a controversial procedure which it takes "strong assumptions about the player's belief-revision policies" to justify.[3] A key objection, first articulated by Kaushik Basu and Phil Reny in the mid-1980s, goes something like this[4]:

Suppose player i deviates from the BI path, and j is asked to move and has to take into account that i will move again. BI, implicitly, calls for j to assume that i will conform with BI in the future. Maintaining that belief is awkward, since j has seen evidence that i is, in fact, not making choices consistent with BI. If j therefore entertains the possibility that i may not conform with BI going forwards, he may have reason to deviate from BI himself.[5] But if this is true, i may have an incentive to deviate from the BI path to start with!

The power of this argument is seen most starkly in centipede games (Rosenthal, 1981) in which all players move multiple times, or chain store paradox games (Selten, 1978) where one player does so. The experimental literature on the empirical relevance of BI is largely centered on centipede games, and the BI solution for selfish players then does not predict particularly well.[6] Since Basu–Reny objections apply, it is natural to wonder whether BI would work better in games in which that were not the case.

We address this issue by designing a series of lab-experiments involving KOH games such that, in the basic versions, each player moves only once. Basu–Reny objections have no bite; if i deviates from the BI path, this offers no presumption regarding subsequent play as i has no further choice.

More...

Apart from the two main motivations already described, we are interested in exploring two additional issues each of which will be reflected using particular treatment variables. First, we consider two different versions of KOH games that differ regarding whether subjects move in sequence (as in the above problem) or simultaneously (as if they surrounded

[1] See Stewart (1999) for a hilarious related analysis: N pirates sequentially propose how to divide their loot, followed by voting whether to accept the proposal or throw the proposer overboard. Conclusions do not exhibit an odd-even effect, but depend starkly on N.

[2] One can show that if (say) Ann must select one of N consecutive days on which to do a boring task, then, for appropriate parameters reflecting her inclination to instant gratification ($\beta < \delta$) and awareness of this ("sophistication"), applying BI regarding the choices of her future selves, she will do the task immediately iff N is odd. We thank Geir Asheim for this example.

[3] The quote is from Battigalli and Siniscalchi (2002, p. 374). For other relevant references, see e.g. Pearce (1984), Reny (1992), Asheim (2002), Perea (2014), Aumann and Arieli (2015), and Battigalli and De Vito (2018).

[4] Examples of references that embrace versions of this line of thinking include Basu (1988), Reny (1988, 1993), Binmore (1987), Ben-Porath (1997), Gul (1997), and Asheim and Dufwenberg (2003) to whom we refer for more commentary and a model which shows how arguably attractive, epistemic assumptions ("common certain belief of full admissible consistency") admit play to leave the BI path.

[5] Up to here, the point was (essentially) made already by Luce and Raiffa (1957, pp. 80–81).

[6] See e.g. McKelvey and Palfrey (1992), Fey et al. (1996), Rapoport et al. (2003), Bornstein et al. (2004), Levitt et al. (2011), and the review Krockow et al. (2016). See also Binmore et al. (2001) who report non-support for BI, in a study that does not employ centipede games, and yet many key comparisons involve players moving multiple times (note e.g. the results mentioned at the top of p. 85).