Games and Economic Behavior 109 (2018) 364–381

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Rationalizable partition-confirmed equilibrium with heterogeneous beliefs ☆

Drew Fudenberg [a], Yuichiro Kamada [b,*]

[a] Department of Economics, Massachusetts Institute of Technology, United States
[b] Haas School of Business, University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 12 January 2017
Available online 2 February 2018

JEL classification:
C72

Keywords:
Rationalizability
Extensive-form games
Self-confirming equilibrium
Heterogeneous beliefs
Purification
Random matching

ABSTRACT

Many models of learning in games implicitly or explicitly assume there are many agents in the role of each player. In principle this allows different agents in the same player role to have different beliefs and play differently, and this is known to occur in laboratory experiments. To explore the impact of this heterogeneity, along with the idea that subjects use their information about other players' payoffs, we define rationalizable partition-confirmed equilibrium (RPCE). We provide several examples to highlight the impact of heterogeneous beliefs, and show how mixed strategies can correspond to heterogeneous play in a large population. We also show that every heterogeneous-belief RPCE can be approximated by a RPCE in a model where every agent in a large pool is a separate player.

1. Introduction

Fudenberg and Levine (1993b) and Fudenberg and Kreps (1995) showed how learning from repeated observation of the realized terminal nodes in each play of a game can allow the long run outcome to approximate a self-confirming equilibrium (Fudenberg and Levine, 1993a), in which the strategies used are best responses to possibly incorrect beliefs about play that are not disconfirmed by the players' observations. This paper defines and analyzes a solution concept that makes three modifications to the self-confirming concept, inspired by the following three considerations. First, the set of beliefs that is consistent with the players' observations depends on what they observe when the game is played, and in some cases of interest players do not observe the exact terminal node, but only a coarser *terminal node partition*, such as when bidders in an auction do not observe the losing bids. Second, both in the lab and in the field, there are often many agents in each player role, so that different agents in a given player role can have different beliefs and play differently, and experimental data frequently suggests that subjects' beliefs and play are indeed heterogeneous.[1] Third, experimental data shows that

☆ We are grateful to Masaki Aoyagi, Pierpaolo Battigalli, Eddie Dekel, Ignacio Esponda, Yuhta Ishii, David K. Levine, Stephen Morris, Yusuke Narita, Jérôme Renault, Tomasz Strzalecki, and seminar participants at Harvard University, Princeton University, Tel Aviv University, Toulouse School of Economics, University of California, Berkeley, and University of Tokyo for helpful comments. We also thank the Advisory Editor and the reviewers of this journal whose comments significantly improved the paper's exposition. Jonathan Libgober and Emily Her provided excellent research assistance. NSF grants SES-0646816, SES-0951462, and SES-1258665 provided financial assistance.

* Corresponding author.
E-mail addresses: drew.fudenberg@gmail.com (D. Fudenberg), y.cam.24@gmail.com (Y. Kamada).

[1] For example, Fudenberg and Levine (1997) relate heterogeneous beliefs to data from experiments on the best-shot, centipede, and ultimatum games.

https://doi.org/10.1016/j.geb.2017.12.021
0899-8256/

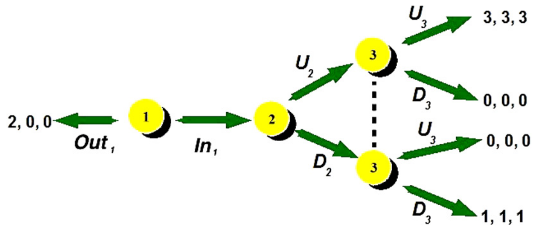

Fig. 1. The broker, seller, and buyer are denoted players 1, 2, and 3, respectively.

subjects play differently when they are informed of opponents' payoff functions than when they are not.[2] To model these three facts, we develop and analyze *heterogeneous rationalizable partition-confirmed equilibrium* (heterogeneous RPCE).

We do not develop an explicit learning theory here, but the model we develop is motivated by the idea that there is a large number of ex-ante identical agents in each player role, who are rematched each period to play an extensive form game and interact anonymously, so that they are strategically myopic and do not try to influence play in future matches. Such random matching is implicit in many learning models, and is explicitly modeled in the Bayesian learning models of Fudenberg and Levine (1993b) and Fudenberg and He (2016).

The long-run implications of learning with random matching depend on what information is revealed at the end of each round of play. In the *information-sharing model*, all agents in the same player role pool their information about what they observe after each round of play, which leads to rationalizable partition-confirmed equilibrium with unitary beliefs, which we studied in Fudenberg and Kamada (2015) (hereafter "FK").[3] In the *personal-information model*, each agent observes and learns only the play in her own match, and no information sharing takes place. This is the treatment most frequently used in game theory experiments. It allows different agents in the same player role to maintain different beliefs, even after many iterations of the game, and even when the agents are identical ex ante.

The large-population learning models described above assume personal information, and so their steady states can have heterogeneous beliefs, which is why Fudenberg and Levine (1993a) defined and analyzed heterogeneous self-confirming equilibrium. Dekel et al. (2004) argue that in Bayesian games it may be appropriate to allow different types of the same player to have different beliefs, and Battigalli et al. (2015) allow heterogeneous beliefs in their extension of self-confirming equilibrium to cases of "model uncertainty." Their model also allows for players to observe a "message" that is generated by a "feedback function" at the end of each play as opposed to the realized terminal node. In Section 2 we say more about the relationship between feedback functions and terminal node partitions.

1.1. Illustrative examples

To motivate our extension of RPCE to heterogeneous beliefs, we will give informal descriptions of RSCE and its extension to unitary RPCE, and then discuss a few examples to show how allowing for heterogeneous beliefs can make a difference. We will return to these examples later after we have introduced our formal definitions.

Roughly speaking, RSCE (Dekel et al., 1999) combines the idea that there is a commonly known equilibrium outcome with that of "rationalizability at reachable nodes," which means that any strategy of player i that some other player j thinks i might be playing maximizes player i's payoff at information sets where player i has not yet been observed to deviate from the equilibrium path, and that each player believes that other players believe this, and so on. As Dekel et al. (1999) show, the combination of these two considerations leads to stronger conclusions than the intersection of the implications of each alone.

Unitary RPCE relaxes RSCE by replacing the assumption of a commonly known distribution on terminal nodes with the assumption of a commonly known partition structure. Both of these concepts are "unitary" in the sense that it is common knowledge that there is only a single strategy and belief that is really held by any agent in the role of player i. However, both solution concepts allow for "hypothetical versions" of each player's strategy and belief, which correspond not to what that player is doing but to things that other players might reasonably think she is doing. The heterogeneous RPCE studied in this paper allows for different agents in the role of a given player to have different beliefs, and moreover for agents to believe that other agents in the same player role have different beliefs than they do.

We now informally discuss some of the implications of heterogeneous beliefs. In those examples we claim that some outcome is inconsistent with unitary beliefs but is consistent with heterogeneous beliefs, given our background assumption that all players have a common belief in rationality and in confirmed beliefs.

Perhaps the most immediate implication is that it allows the play of one or more players to be strictly mixed in cases where unitary beliefs require the outcome to be a single terminal node, as in the game in Fig. 1.

[2] See for example Prasnikar and Roth (1992).

[3] FK extends an earlier literature on equilibrium concepts that combine restrictions based on the agents' observations with restrictions based in their knowledge of opponents' payoffs, including Rubinstein and Wolinsky (1994), Battigalli and Guaitoli (1997), Dekel et al. (1999), and Esponda (2013).