Cognition 172 (2018) 107–123

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

The determinants of response time in a repeated constant-sum game: A robust Bayesian hierarchical dual-process model

Leonidas Spiliopoulos

Center for Adaptive Rationality, Max Planck Institute for Human Development, 94 Lentzeallee, Berlin 14195, Germany

ARTICLE INFO

Keywords:
Response time
Dual-process (dual-system) models
Win-stay/lose shift heuristic
Pattern-detecting reinforcement learning
Bayesian hierarchical modeling
Repeated constant-sum games
Experimental economics

ABSTRACT

The investigation of response time and behavior has a long tradition in cognitive psychology, particularly for non-strategic decision-making. Recently, experimental economists have also studied response time in strategic interactions, but with an emphasis on either one-shot games or repeated social-dilemmas. I investigate the determinants of response time in a repeated (pure-conflict) game, admitting a unique mixed strategy Nash equilibrium, with fixed partner matching. Response times depend upon the interaction of two decision models embedded in a dual-process framework (Achtziger and Alós-Ferrer, 2014; Alós-Ferrer, 2016). The first decision model is the commonly used win-stay/lose-shift heuristic and the second the pattern-detecting reinforcement learning model in Spiliopoulos (2013b). The former is less complex and can be executed more quickly than the latter. As predicted, conflict between these two models (i.e., each one recommending a different course of action) led to longer response times than cases without conflict. The dual-process framework makes other qualitative response time predictions arising from the interaction between the existence (or not) of conflict and which one of the two decision models the chosen action is consistent with—these were broadly verified by the data. Other determinants of RT were hypothesized on the basis of existing theory and tested empirically. Response times were strongly dependent on the actions chosen by both players in the previous rounds and the resulting outcomes. Specifically, response time was shortest after a win in the previous round where the maximum possible payoff was obtained; response time after losses was significantly longer. Strongly auto-correlated behavior (regardless of its sign) was also associated with longer response times. I conclude that, similar to other tasks, there is a strong coupling in repeated games between behavior and RT, which can be exploited to further our understanding of decision making.

1. Introduction

The cognitive processes underlying decision-making determine both the decisions made and the time taken to arrive at said decisions, referred to as the response time (RT). As a consequence of interest in the cognitive processes underlying decision-making, the discipline of psychology has long appreciated this link between RT and decisions (Donders, 1868). The fact that cognitive processes are largely latent or hidden, necessitates their inference from observables such as choices, RT, and—more recently—neural activity. The discipline of economics on the other hand has historically focused almost exclusively on choices, whilst ignoring the information in RT. An important reason for this difference was that economists were typically interested in the predictive power of their models, but were less concerned with whether the model was an accurate representation of the true cognitive processes. However, experimental economists during the last decade have also begun analyzing RT both in individual and strategic decision-making. Examples of early work include Kocher and Sutter (2006) and Rubinstein (2007) respectively, see Spiliopoulos and Ortmann (2017) for a literature review and critical discussion of the existing studies in experimental economics utilizing RT.

In this manuscript I analyze response time in a strategic situation first investigated in Spiliopoulos (2013b); a *repeated* 2×2 constant-sum game with a unique mixed strategy Nash equilibrium (MSNE) and fixed partner matching. The experimental data collected in this study is ideal for investigating RT in repeated games as it includes many observations per individual (300 rounds in total) and significant within-subject variation in strategic behavior by pitting subjects against three very different computer algorithms (CAs) for 100 rounds each. To the best of my knowledge there are no other published studies that examine RT in a strictly competitive repeated game where social preferences are irrelevant. The closest work is a working paper by Gill and Prowse (2017), which I will return to below, after summarizing other published works.

E-mail address: spiliopoulos@mpib-berlin.mpg.de.

https://doi.org/10.1016/j.cognition.2017.11.006
Received 8 July 2016; Received in revised form 14 November 2017; Accepted 19 November 2017

The existing RT literature has focused on whether cooperation or self-serving behavior is instinctive (on the basis of faster response times), initially using one-shot games, but increasingly also using repeated social dilemmas (e.g., Evans, Dillon, & Rand, 2015; Krajbich, Bartling, Hare, & Fehr, 2015; Piovesan & Wengström, 2009). In a repeated Prisoner's Dilemma, Collins (1977) linked RT to personality types. Repressors who tend to avoid anxiety arising from conflict by making a decision quickly, and sensitizers who instead take longer to choose due to excessive rumination over the conflict. However, this paper does not contain any analysis based on the repetitions of the game or the history of play. Other research in repeated social dilemmas finds that the recent history of play can modulate the relationship between RT and behavior. A large proportion of subjects in repeated public good games are conditional cooperators, who by definition base their behavior on the recent choices of the other players (Fischbacher, Gächter, & Fehr, 2001). Similarly, Lotito, Migheli, and Ortona (2013) found that subjects responded more quickly to cooperation rather than defection in the previous round. Nishi, Christakis, Evans, O'Malley, and Rand (2016) argue that in repeated social dilemma games, response time is determined not by whether cooperation or defection is instinctive, but rather whether the chosen behavior is reciprocal given the choices of others. Interestingly, they find that a player's own prior behavior (cooperation or defection) also influences the response time. In short, the literature on social dilemmas consistently finds that RT is affected by the history of play in the prior round. I will test empirically whether this holds for games where social preferences are irrelevant, which is still an open matter.

There is also a rich existing literature on games that admit only a unique MSNE—such as the one in this study—focusing on how well human behavior approximates or converges with learning to the Nash equilibrium prediction in terms of (a) marginal mixing probabilities and (b) independently distributed action choices across rounds.[1] Studies focus on mathematical models of learning (e.g., Camerer & Ho, 1999; Cheung & Friedman, 1997; Roth & Erev, 1995) including belief-elicitation (e.g., Nyarko & Schotter, 2002; Rutström & Wilcox, 2009) and on randomization across rounds and learning models of historical patterns of play (e.g., Ioannou & Romero, 2014; Rapoport & Budescu, 1992; Scroggin, 2007; Spiliopoulos, 2012, 2013a). Studies also differ in the means of data collection, such as experimental studies in the lab (e.g., Brown & Rosenthal, 1990; O'Neill, 1987; Shachat, 2002; Shachat & Swarthout, 2004, 2012; Shachat, Swarthout, & Wei, 2015) and field studies of human behavior at different levels of task-relevant expertise (e.g., Chiappori, Levitt, & Groseclose, 2002; Palacios-Huerta, 2003; Palacios-Huerta & Volij, 2008; Walker & Wooders, 2001; Wooders, 2010).

By using a game where social preferences are irrelevant, I will be focusing solely on the strategic determinants of response time as motivated by pure self-interest. I advance existing studies in strategic decision-making, that are usually restricted to simply comparing mean RTs across the whole subject pool for different actions, by explicitly modeling (a) the full distribution of RT, (b) individual heterogeneity, (c) the dependence of RT on other important covariates or determinants, and (d) by rigorously dealing with the issue of outliers. The closest study is a working paper by Gill and Prowse (2017) who investigate a repeated *p*-beauty game for ten rounds with three players. They find a dependence of RT on the outcome of the prior period. Specifically, response time was the longest when a player won in the previous round with a choice between the choices of the other two players. The shortest RT was recorded for choices immediately after a loss where the player's choice was higher than those of both of the other players.

The following determinants are hypothesized to mediate response time in repeated games. The first is whether two behavioral decision models with empirically documented success in predicting action choices, also systematically affect RT. While these two models may be viewed as competitors (either one or the other is in use), another possibility is that they both interact, see for example dual-system models (Kahneman, 2011). In dual-system models, System 1 is typically considered to be faster, easier to execute and is cued by easily retrievable or accessible information. System 2 is typically described as deliberative, in the sense that it is slower than System 1 because it uses more information, and typically requires more complex information integration. In early work, the two systems were often not well-defined in terms of the processes corresponding to each, and their interaction was not formally specified. This afforded a great deal of flexibility in explaining data, inviting critiques—see for example Keren and Schul (2009), Ortmann (2008), Rustichini (2008) and for a general discussion the special issue in *Journal of Economic Psychology* (Alós-Ferrer & Strack, 2014).

Newer work is addressing these issues by clearly specifying and constraining the set of possible processes in each system and their interaction, thereby making clear, falsifiable predictions. For example, dual-self models (Fudenberg & Levine, 2006, 2012; Fudenberg, Levine, & Maniadis, 2014) are predicated on the existence of a short-run self that cares only about short-run payoffs and a long-run self that cares about future payoffs. The two selves interact strategically. The long-run self seeks to influence the utility function of the short-run self, but incurs a self-control cost to do so; however, the final decision is controlled by the short-run self. Achtziger and Alós-Ferrer (2014) formalize a dual-process framework that assumes the existence of two processes and a central executive system that controls which process is used to decide on a course of action. Alós-Ferrer (in press) extends this framework by adding drift-diffusion justifications for each of the two processes. This new framework, the dual-process diffusion model (DPDM), is the one that I will adopt in this study. Apart from some assumptions regarding the relative characteristics of the two processes, this framework remains relatively agnostic about the specific processes. However, the framework has enough structure to make clear *qualitative* predictions about response time and its relationship to conflict or alignment between the two processes, the probabilities of selection of the two processes and errors in the decision-making process.

I operationalize this framework by postulating two specific processes on the basis of their empirical performance in explaining action choices of players in this particular experiment, and other studies with similar games. The first is a simple, yet effective, decision rule, the win-stay/lose-shift (WSLS) heuristic, which prescribes playing the same action as the previous round if it led to a "win", or switching actions if it led to a "loss". The second is the structural pattern-detecting reinforcement (PDRe) learning model introduced in Spiliopoulos (2013b), which is more complex than the WSLS heuristic, both in terms of the information required and its integration. These two models are presented formally in Section 3 along with the specification of the DPDM, in which they are both embedded. In a similar spirit, Worthy and Maddox (2014) present empirical evidence of the concurrent use of a WSLS heuristic and a standard (non-pattern detecting) reinforcement learning algorithm. Their work is related to this manuscript, but there are important differences beyond the technical details about the specification of the interaction of the processes, and the processes themselves. Their findings were validated in the realm of non-strategic decision tasks and they did not study the implications of the interaction of these two processes for response time, only for actions.

Beyond the DPDM hypothesis, I will also test whether the immediate history of play (the actions chosen by both players in the immediate prior round), the degree of auto-correlation in subjects' actions, the recent trend in realized payoffs, and experience also significantly influence RT. For ease of exposition, in the next section I

[1] Take for example the famous game of Rock-Paper-Scissors that admits a unique MSNE where a player must choose each with equal probability, 1/3. If play across rounds is not independently distributed, e.g., a player is more likely to follow a Rock action with another Rock action, then an opponent could exploit this by conditioning on previous actions.