Feedback manipulation and learning in games

Feedback Manipulation and Learning in Games

by Aidas Masiliūnas

Summary

Chapters in this thesis investigate how manipulations that affect the type of feedback but not the set of equilibria influence the learning process and the long-run outcomes of a game. Chapter 2 reports the results of an experiment in which the reliability of feedback is manipulated by reducing the variability in opponent's actions and in the payoff function, leading to dramatically improved convergence rates and reduced behavioural variation in a contest game. A theoretical model presented in Chapter 3 specifies conditions under which sophisticated players would be willing to manipulate the feedback observed by myopic players by taking a suboptimal action in a repeated coordination game. An experiment in Chapter 4 varies the cost of action disclosure and shows that higher costs discourage action disclosure and lead to incomplete feedback that prevents transitions to an efficient equilibrium.

Chapter 2 uses a Tullock contest as an example of a game in which game theory fails to make accurate predictions: Nash equilibrium is chosen only 7% of the time while dominated strategies are chosen more than half of the time. A common view is that the low explanatory power of a Nash equilibrium results from the failure of assumptions about risk-neutral and self-interested behavior. However, deviations from theoretical predictions may be a result of complexities that make optimization difficult for boundedly rational players. Understanding the source of the discrepancy between choices and theoretical predictions would point to the correct way to address this problem. If the discrepancy arises because of non-standard preferences, the problem of low explanatory power could be solved by adding such preferences to the theory. But if the discrepancy results from bounded rationality, it may be desirable to introduce interventions that would enable players to learn to behave in an optimal way.

One factor that may limit the opportunities to learn in contests is noisy feedback arising from a stochastic payoff function. If there were no stochastic elements in the payoff function, dominated strategies would always provide low earnings and given enough time players should learn to avoid such strategies. In standard contests, however, payoffs are stochastic due to strategic uncertainty—changes in the actions chosen by other players—and due to non-strategic uncertainty generated by a randomization device. In a 2x2 design we reduce both types of uncertainty: strategic uncertainty is lowered by matching players to computers who play the same action for a certain number of rounds and uncertainty about the prize allocation is lowered by paying the expected value instead of playing a lottery. We find that when no uncertainty is present, the frequency of dominated actions decreases dramatically and the median response is almost always equal to the theoretical prediction. When either type of

uncertainty is present, choices are very different from the theoretical prediction and dominated strategies are chosen more than half of the time.

Feedback may be manipulated by the experimenter, but it also depends on the actions taken by other players. If feedback affects behavior, a strategic player may therefore attempt to alter the future behavior of other group members by manipulating their observed feedback. A certain action could therefore be chosen just for the information it conveys to other group members rather than for the immediate profit it generates. **Chapter 3** and **chapter 4** use theory and experiments to test whether such strategic motives could explain deviations from an inefficient equilibrium in N-person critical mass games. In these games non-equilibrium outcomes are observed much less frequently than in contests, as typically all groups rapidly coordinate on a common action. However, a social dilemma arises if play converges to an inefficient equilibrium: any unilateral deviation from an inefficient equilibrium would decrease deviant's immediate utility, even though a Pareto improvement could be achieved by collective action. Inefficient equilibrium persists because staying in it is optimal given the belief that others will stay too. If beliefs are the core of the problem, players who anticipate the belief formation process may maximize their earnings by deviating from an inefficient equilibrium because a decrease in immediate payoffs is more than compensated by the benefits of efficient coordination in the future. Chapter 3 shows that under certain conditions self-interested strategic players would be willing to deviate from an inefficient equilibrium. Chapter 4 tests whether players in an experiment are actually motivated by such strategic considerations and whether transitions to an efficient equilibrium are more frequent when strategic teaching is made easier by reducing the cost of action disclosure.

Chapter 3 proposes a solution concept that combines a learning model based on fictitious play and the concept of a Nash equilibrium by assuming that some players are myopic and update beliefs based on observed history while others are sophisticated and correctly anticipate the actions of all other players. The proposed solution concept is a combination of sophisticated player strategies that are optimal given the learning process of the myopic players and the strategies chosen by other sophisticated players. A combination of learning from history and strategic reasoning produces action paths in which sophisticated players find it optimal to use strategic teaching, and the existence of such action paths depends both on the history observed by myopic players and on the length of the reasoning horizon of sophisticated players.

Chapter 4 tests whether deviations from an inefficient equilibrium are motivated by strategic considerations and whether transitions to an efficient equilibrium are more likely if information about one's action can be provided at a low cost. In an experiment players could disclose their action to other group members by paying a cost that was varied across treatments. Players who only care about immediate payoffs would not be willing to pay the cost because disclosure provides no immediate benefits, but strategic players may do so if they expect that a disclosed action increases the chances of a transition to the efficient equilibrium.

Data shows that many players are willing to pay to reveal their actions, especially when the costs are low, but only after choosing to deviate from an inefficient equilibrium. When disclosure costs are low, players classified as more farsighted more often deviate from an inefficient equilibrium and disclose this action, providing further support for the strategic teaching hypothesis. Higher disclosure costs reduce the tendency to reveal actions, increasing strategic uncertainty and making transitions to an efficient equilibrium much less likely. In fact, no group moves to the efficient equilibrium when action disclosure costs are high, but half of the groups do so when the costs are low. Belief learning seems to be the most likely explanation for this treatment difference: stated beliefs generally move in accordance to the predictions of weighted fictitious play, but players who do not disclose their actions are perceived in a similar way to those who choose the inefficient action. Lack of feedback resulting from high action disclosure costs seems to introduce frictions into the learning process, reducing the frequency of transitions to the efficient equilibrium.