

Climate negotiations, uncertainty, and the prisoners' dilemma trap

Scott Barrett and Astrid Dannenberg
Columbia University

February 2014

From the beginning, the problem of limiting climate change was conceived of as requiring a cap on global emissions, achieved by imposing limits on the emissions of individual countries. This is equivalent to looking at the provision of a public good as requiring contributions from individual players. The approach confronts the prisoners' dilemma directly. The difficulty of this approach comes in overcoming the incentives to free ride. The previous experimental literature has shown that behavior varies in different institutional settings, such as when players can punish free riders. However, such punishments are rarely if ever used to supply global public goods. International institutions work differently. Successful examples, like the International Convention for the Prevention of Pollution from Ships or the Montreal Protocol on Substances that Deplete the Ozone Layer, use technology standards and trade restrictions to transform the prisoners' dilemma into a coordination game.

Coordination games have the theoretical advantage that the collective best outcome can be supported as a Nash equilibrium. However, this equilibrium is risky and competes with a safe but payoff-dominated equilibrium. Moreover, transformation of the prisoners' dilemma often comes at a cost. Technology standards, for example, reduce flexibility. Hence it's not clear that this transformation always is the way to go. Moreover, it's not clear whether players *think* this is the way to go. In this paper we develop a simple model of a public goods meta-game and test it in the experimental lab. In the meta-game, players must decide whether to provide a public good by playing the prisoners' dilemma game or the coordination game. Our experiment involves two sets of treatments. In the first, the coordination alternative is potentially as efficient as the prisoners' dilemma approach. In the second, the coordination approach is potentially less efficient, but may prove the better choice because of its welfare-superior Nash equilibrium. The questions we try to answer are: Which institution works best? Do people pick the best institution? How do they choose between an institution with potentially high payoffs and one with lower payoffs but a strategic advantage? Do they learn with experience?

In our experiment, all groups initially choose to play the prisoners' dilemma, when it is potentially the more efficient approach. About half the groups subsequently try to coordinate, and these groups invariably do better. They contribute more to the public good and they receive higher payoffs. However, the other groups remain trapped in the prisoners' dilemma. They believe they are doing the best they can do, but they are mistaken. The groups that switch do miserably when playing the prisoners' dilemma, and are thus "pushed" into trying the other approach. These groups are also optimistic about the prospects for coordination, and so are "pulled" towards trying the other game. We show that both the push effect and the pull effect are necessary for groups to switch.