
CHRISTOPH HAUERT, University of British Columbia
Sanctioning institutions for governing the commons

Cooperation represents a key organizing principle in genetic and cultural evolution. Yet cooperation is a conundrum because cooperators make a sacrifice to benefit others. Despite the fact that groups of cooperators outperform groups of defectors, Darwinian selection or utilitarian principles based on rational choice favors defectors. Nevertheless, cooperation is ubiquitous in biological and social systems. Indeed, cooperation can be stabilized by punishing defectors. Punishment is also ubiquitous in nature - ranging from toxin producing microorganisms to law enforcement institutions. But how can initially rare, costly punishment behavior gain a foothold in a population? In nature, individuals carefully select their interaction partners and refuse to participate in risky enterprises. Such voluntary participation prevents deadlocks in states of mutual defection and thus promotes cooperation - but fails to stabilize it. However, the combined efforts of punishment and volunteering can change the odds in favor of cooperation. Under the stochastic dynamics of finite populations the freedom to withdraw leads to prosocial coercion. To date, theory and experiments emphasize the role of such peer-punishment. At least in human societies peer-punishment has been largely superseded by sanctioning institutions and vigilantism deemed illegal. This can be modeled by introducing the opportunity for pool-punishment, which represents a precursor of executive power and echoes Elinor Ostrom's principles for 'Governing the Commons'. Pool-punishment always incurs costs to those committed to it even if no one requires reprimanding. Interestingly, our model predicts that individuals nevertheless trade the higher efficiency of peer-punishment for the increased stability of pool-punishment to maintain cooperation.