Fairness Evolves with Partial Information of Partner Choice

Paul Rauwolf and Joanna J. Bryson

Theoretical background:

In the trust game, an *investor* is given one unit of money and is offered the opportunity to invest with a *trustee*. If invested, the money is multiplied by some factor, *b*, and the trustee can return some portion to the investor. In a one-shot game, the subgame-perfect equilibrium is to reject investment, as there is no reason for the trustee to return any. Despite this, humans consistently choose to invest.

Recently, analysis has shifted to include offering investors information on a trustee's historical return rate prior to deciding whether to invest. The rational play is to invest only if the trustee's return rate is greater than 1/b. However, in empirical studies, individuals often request a return rate of 50%, deemed a 'fair' split.

Research question:

Manapat, et. al., (2012) demonstrated that if all investors collude and require a high return rate, the trustees evolve to return the requested rate. However, this is not evolutionarily stable. An agent demanding the rational return rate of just over 1/b invades a population of colluding agents. If fairness is not evolutionarily stable, how did it evolve?

Method:

We extend the evolutionary game theory model of Manapat, et. al., (2012), who demonstrated that trust evolves with partner choice. We analyze whether partner choice and partial information are sufficient for evolving fairness. An investor is offered the opportunity to invest with one of K partners. There is q probability that the investor knows the return rate of a trustee. The investor can either a.) invest with the trustee who is offering the highest known rate of return, b.) invest with a trustee whose return rate is unknown, or c.) keep the money.

Each agent possesses a demand threshold, *d*. If none of the known trustees are offering a rate of return above *d*, then the investor may take a risk and trust someone whose information is unknown. We analyze whether, for varying values of q and K, a demand evolves which is 'fair'.

Results:

We show that when $q<1$ and $K>1$, an agent benefits from rejecting unfair offers and trusting unknown partners. Such behaviour forces a higher return from the trustees.

Conclusion:

Humans often trust without information of another's return rate, but demand fairness when given information. We demonstrate that if humans evolved in social contexts with incomplete information, both traits are advantageous and evolutionary stable.