Generalized Trust in the Mirror - a model on the Dynamics of Trust

Description

High levels of trust have been linked to a variety of benefits including the well-functioning of markets and political institutions or the ability of societies to solve public goods problems endogenously. While there is extensive literature on the macro-level determinants of trust, the micro-level processes underlying the emergence and stability of trust are not yet sufficiently understood. We address this lacuna by means of a computer simulation. With this model, conditions under which trust is likely to emerge and be sustained are identified. We focus our analysis mainly on the individual characteristics of agents: their social or geographical mobility, their attitude towards others or their general uncertainty about the environment. Contrary to predictions from previous literature, it turns out that a low geographical or social mobility is detrimental to both, the emergence and robustness of trust. The model also helps to reveal a hidden link between trusting others and being trustworthy.

Layout:

There is a fix number of agents, controlled by *number_of_agents*. Each agent serves two roles: First, they try to determine whether others are, in generally, trustworthy. More precisely, they try to find whether there is enough trustworthiness in society society to make trusting others a rational thing to do. Second, each agent has their own type of behavior when being trusted by others: they are trustworthy or not. Let's start with the latter. Agents do not change their behavior within a model run, they are either always trustworthy or never. The reason for this, briefly, is that some real life agents' trustworthiness develops on a much larger time scale than the one we are interested in. The share of untrustworthy agents is determined by the slider *percent-untrustworthy*.
Independent of their own trustworthiness, each agent tries to estimate whether overall trustworthiness is high enough do warrant trust in others. This estimate is expressed as a number in the interval [0,1]. At the beginning of a simulation run, agents start with an initial estimate, drawn from a normal distribution with variance 0.2 around the input parameter *initial_trust_expectation*. Agents then update their estimates through interaction with others. Agents that estimate the overall amount of trustworthiness high enough (at least 0.5) are colored green, else they are colored red.

Model Dynamics

Each interaction round consists of three phases. In the first phase, agents randomly pair up with a neighbor chosen from their Moore Neighborhood. If no such neighbor is available or every neighbor is already teamed up with somebody else, agents remain unpaired. In the second phase, agents interact with their chosen partner. Such interaction starts with one partner deciding whether to put trust in the other person or not. She does so iff her estimate of trustworthiness is above 0.5. In this case, the second interaction partner will act trustworthy or not, depending on her trustee ype. Both agents then update their beliefs with the learning rule described below. Finally, in the third phase of a round, every agent moves a fix distance, controlled by the input parameter *mobility.* Each spot can only be occupied by one agent; an agent will move until she finds a free space.
The simulation continues for a maximum of 1000 rounds. Simulations are aborted early if all agents are red or all but at most ten agents are green. In this case, running the simulation longer would not produce further change.

Information updates

Within the personal interaction described above, both agents may gain additional information about overall trustworthiness. If the first agent decides to place trust, she learns whether her counterpart is trustworthy (I=1) or not (I=0). But also the latter agent gains new information. In observing whether the first agent is willing to place trust, she learns whether this agent deems overall trustworthiness high enough to place trust (I=1) or not (I=0). Both feedbacks are incorporated into the agents belief by a weighted average. For direct learning, i.e. the information collected as trustor, this weighted average is:

$$estimate\_new = (1-w) * estimate\_old + w*I$$

In this formula, w denotes the weight an agent is willing to attribute to new information. This weight is controlled by the input variable *weight_new_information.*
For indirect learning, the new information is discounted by a factor controlling how much more or less the agent values indirect information compared to direct information. This is controlled by the slider *social-learning-factor.* The formula for the weighted average is then:

$$estimate\_new = (1-w)*s * estimate\_old + w*s*I$$

where s denotes the value of social-learning-factor

Shocks

The shock button reduces the estimate of each agent by a random number between 0 and 0.5

Things to notice

- In the long run, all agents will obtain the same color, i.e. they will all agree in their estimate whether others are, overall, trustworthy or not
- At a mobility of 1, local clusters of green and red agents can form. In this case, red clusters expand slowly, i.e. agents begin to belief to lose trust in others. As a consequence, less simulations end up with all agents green than with higher mobilities. The existence of clusters is tracked by the Monitor ‘Segregation’, calculating the Bray-Curtis index of dissimilarity for the distribution of green and red agents.