# Submission Number: PET11-11-00310

## Getting and Staying out of a Low Trust Trap: Roles of Tipping and Institutions

Thomas B Singh
*University of Guyana*

## *Abstract*

Given that there are significant endogenous effects in trust that cause people to follow trust norms when deciding whether to trust others, trust can be modelled as a supermodular game or a game with strategic complementarities. As such, even though the strategy choices are binary and discrete, the game is amenable to monotone comparative static analysis in which institutional reform can be represented as an increase in a parameter in the game of trust. The analysis pessimistically implies that formal institutional reform has only a limited role to play in increasing the level of trust because low trust is in fact an equilibrium or a 'low trust trap.' What is first required is a tipping mechanism, which involves only a relatively small number of influential agents adopting the strategy of trusting others even in a low trust environment. Formal institutions, appropriately reformed, will then be required to ensure that shocks to trust, even if they did occur, would not accumulate and tip society back into low trust.

**Submitted:** March 16, 2011.

# Getting and Staying Out of a Low Trust Trap: The Roles of Tipping and Institutional Reform[1]

*By*

Thomas B. Singh

Institute of Development Studies

University of Guyana

January 2011

**Abstract**

Given that there are significant endogenous effects in trust that cause people to follow trust norms when deciding whether to trust others, trust can be modelled as a supermodular game or a game with strategic complementarities. As such, even though the strategy choices are binary and discrete, the game is amenable to monotone comparative static analysis in which institutional reform can be represented as an increase in a parameter in the game of trust. The analysis pessimistically implies that formal institutional reform has only a limited role to play in increasing the level of trust because low trust is in fact an equilibrium or a 'low trust trap.' What is first required is a tipping mechanism, which involves only a relatively small number of influential agents adopting the strategy of trusting others even in a low trust environment. Formal institutions, appropriately reformed, will then be required to ensure that shocks to trust, even if they did occur, would not accumulate and tip society back into low trust.

**Key Words**: Monotone Comparative Statics, Policy, Institutions, Tipping Mechanisms

**JEL Codes**: Z10, C72, C73, B52, D83, D84

[1]For the 2011 Annual Conference of the *Scottish Economic Society*, Perth, Scotland. Conference Website: http://www.ses-conference.co.uk/.

# Getting and Staying Out of a Low Trust Trap: The Roles of Tipping and Institutional Reform


## Abstract

This paper exploits the connection between supermodularity and the tipping of equilibria to discuss raising the level of trust in low trust countries. Given that there are significant endogenous effects in trust that cause people to follow trust norms when deciding whether to trust others, trust can be modelled as a supermodular game or a game with strategic complementarities. As such, even though the strategy choices are binary and discrete, the game is amenable to monotone comparative static analysis in which institutional reform can be represented as an increase in a parameter in the game of trust. The analysis implies that formal institutional reform has only a limited role to play in increasing the level of trust because low trust is in fact an equilibrium or a 'low trust trap.' What is first required is a tipping mechanism, which involves only a relatively small number of influential agents adopting the strategy of trusting others even in a low trust environment. Formal institutions, appropriately reformed, will then be required to ensure that shocks to trust, even if they did occur, would not accumulate and tip society back into low trust.


## 1. Introduction

If low, medium and high average trust were somewhat arbitrarily defined as below 0.25, between 0.25-0.5, and above 0.5 respectively, then 53% of the countries surveyed in Wave IV (1999 – 2004) of the *World Values Surveys* (*WVS*) would exhibit low trust, 36% medium trust, and 11% high trust. Raising the level of trust is an issue because more than 50% of the world's economies are characterised by low trust and what is more, low trust tends to persist

in low trust economies. The tendency for the persistence of low trust, despite the wave of large-scale institutional reforms in transiton and developing countries over the last couple of decades, is related to the notion that low trust countries might be caught in a low-trust traps, while the preponderance of low trust in the economies surveyed in the *World Values Surveys*, is related to the idea that low trust equilibria are stochastically stable. The basic argument of this paper is that a tipping mechanism is required to raise the level of trust in countries caught in a low trust traps. While the reform of formal institutions cannot take a country out of low trust, good formal institutions are important in sustaining the level of trust medium and high trust countries because medium trust is unstable and there is a long run tendency for the collapse of trust in even high trust countries.

These results are in fact closely related to the 'stylised facts' of within-country conformity in trust decisions and global diversity in average trust across countries, strikingly evident in the *World Values Survey* data on interpersonal trust among strangers (Singh 2010). Young (2003) explains the simultaneous existence of local conformity and global diversity effects as the outcome of norm-based behaviour that arises from the cumulative effect of uncoordinated decisions made by many interacting individuals. The within-country conformity would arise if people followed trust norms when making trust decisions and, as trust norms vary from country to country, there would naturally be stark variation in average trust across countries. Raising the level of trust in low trust economies would therefore require norm-shifting or tipping mechanisms, as against institutional reform, which merely sustains trust in medium trust economies and prevents the long run collapse of trust high trust in countries. This latter result on the relative ineffectiveness of formal institution reform in securing desirable social equilibria and changing individual behavior is now fairly well established in the literature on the economic significance of social norms and culture in the quest for cooperation and for

influencing economic outcomes in general.

When applied to trust, the Young (1993) model of the evolution of conventions would involve randomly matched pairs of boundedly rational individuals playing a 2 × 2 coordination game that is subject to persistent shocks to its best response adaptive dynamics. High trust and low trust are absorbing states or norms that correspond to high and low average trust levels, and conformity to the trust norm prevailing within a country is explained as convergence to one of these absorbing states, which are strict (pure strategy) Nash equilibria of the trust game. As such, the high and low trust norms do not have to be enforced by any external authority or even by social sanctions against those who depart from the norm. Trust norms are self-enforcing because they involve shared beliefs and expectations about how others in society will behave. If the norm is for everyone to be trusting, then "everyone trusts, everyone expects others to trust, and everyone has good reason to trust because trusting is in each person's *best interest* when everyone one else plans to trust," to modify Young (1996)'s classic description of conventions as a regularity of behaviour in the sense of Lewis (1969); and similarly for cynicism.

The stability of high and low trust norms is not the only prediction of the model. Medium trust is also a Nash equilibrium of the trust game in which the population shares playing trusting and cynical strategies are equal to the probabilities given in the mixed strategy equilibrium of the trust game, but the medium trust equilibrium is highly unstable, and therefore is not a norm. Medium trust in countries such as the USA and Germany is only sustained by good formal institutions. Moreover, while both high and low trust norms are stable, the low trust equilibrium is risk dominant and has the larger basin of attraction, and is therefore *stochastically stable*. Hence in the long run all societies will tend to collapse to low

trust. Without good formal institutions even high trust economies would collapse to low trust in the long run.

The empirical counterpart of norm-based behaviour is the existence of strong endogenous social interactions effects (Manski, 1993; Brock and Durlauf, 2001). Using the Brock and Durlauf (2001) discrete-choice-with-social-interactions model, Singh (2010) establishes the existence of endogenous effects in trust decisions, implying that people do indeed tend to follow the trust culture or norm prevailing in their particular society. This explained the within-country conformity in trust decisions. Moreover, given that the endogenous effects are strong enough to produce multiple equilibria, the global diversity in trust that has hitherto been an unresolved puzzle in the literature on trust, is now explainable as coordination on trust norms that differ from one country to the next.

While this article is recognises that the low trust equilibrium is stochastically stable, and is therefore concerned with raising the level of trust in countries that are already caught in low trust traps, an important strand of the literature considers the dynamics of the game and the interaction structures that might in fact preclude the stochastic stability of low trust. This includes Skyrms and Permantle (2000), which allows the very structure of interactions to evolve by reinforcement learning, meaning that people in the trust game learn who to interact with; the possibility of population clustering discussed in McCabe *et al* (2003), though the discussion here is limited to Prisoner's Dilemma type trust games and evolutionary dynamics; and Berninghaus and Haller (2010), which maintains the local interaction of Young (1993) and Ellison (1993), but also allows the structure of the interactions to change over time due to stochastic shocks, to capture the idea the social ties are not rigid.

For countries already caught in low trust traps, getting out of low trust is actually facilitated

by the empirical finding of strong endogenous effects that produce multiple equilibria. Eichenique (2001) shows that any 2 × 2 game with multiple equilibria is a supermodular game, or a game with strategic complementarities, so that even though the game of trust discussed here involves binary strategy choices, to trust or not to trust others, it is possible to develop sufficient conditions for tipping out of low trust equilibria and to use monotone comparative statics to assess the effectiveness of institutional reform in raising trust. This is precisely what was done Heal and Kunreuther (2006) in the context of homeland security and computer activity, and the analysis here draws heavily on that discussion.

Section 2 outlines the stage game of trust with Stag Hunt payoffs that is played recurrently by randomly matched pairs of boundedly rational agents. Section 3 sketches an argument that the adaptive play of the trust game that is based on Young (1993) produces results that are equivalent to an analysis of the trust game as one with strategic complementarities, which in turn reflect the idea that there are endogenous effects in the trust decision. Section 4 discusses the role of formal institutional reform in transforming low trust countries into high trust ones, noting that this role is limited to increasing the level of trust without catapulting such countries into high trust. What is actually required to do so is a 'tipping mechanism,' and this is discussed in Section 5. Section 6 outlines a two-part policy for transforming low trust countries into high trust ones. The first part addresses tipping out of the low trust equilibrium, while the second part addresses policy to maintain high trust, given the (very) long run tendency for low trust to be stochastically stable. Section 7 concludes.

## 2. The Recurrent Game of Trust among Strangers

While trust is usually represented by cooperation in a Prisoner's Dilemma (**PD**) game, the willingness to trust others being the basis (in preferences or beliefs) of cooperation, trust is

neither necessary nor sufficient, and it might even damage the prospects, for cooperation.[2] Moreover, the specific assumptions about behavior that undergird the representations of trust in **PD** games vary widely, from the rational cooperation of Kreps *et al* (1982) to the reciprocity-as-a-primitive hypothesis of Berg *et al* (1995), without any consensus about or general theory of the behavior that supports and enhances trust. To address these concerns it is explicitly assumed that people make decisions or choices about whether to trust others, and that their decision making is characterized by learning as the basic game of trust is played recurrently by randomly matched pairs of agents. The behavioural assumption therefore is that agents are boundedly rational, though they ultimately behave rationally because they learn to do so[3] as they play the game of trust.

As emphasized by Skyrms (2001) it is the Stag Hunt game, in which two players must choose either to cooperate with each other to hunt a highly valued stag or to defect and hunt a lower valued hare independently, that is really about whether a hunter decides to trust or not to trust the other hunter. In the Stag Hunt game shown in *Fig. 1*, two players must choose either to cooperate with each other to hunt a highly valued stag, or to defect and hunt a lower valued hare independently. Unlike the **PD**, cooperation is privately more appealing than unilateral defection so there is no incentive to defect. Stag hunting or cooperation by both players is itself a Nash Equilibrium because neither would want to unilaterally switch to hunting hare when the other is hunting stag.

[2] For a comprehensive review of the literature that addresses this issue, see Raymond (2006).
[3] For an interesting discussion of this point, see Binmore (2006).

| | | ***Player 2*** | |
|---|---|---|---|
| | | **Trust Others** | **Be Cynical** |
| ***Player 1*** | **Trust Others** | $\pi^1_{tt}, \pi^2_{tt}$ | $\pi^1_{tc}, \pi^2_{tc}$ |
| | **Be Cynical** | $\pi^1_{ct}, \pi^2_{ct}$ | $\pi^1_{cc}, \pi^2_{cc}$ |

*Fig.1. A Game of Trust*

The payoff to *Player* ($k$ = 1, 2) is $\pi^k_{ij}$, where $i, j$ refer to the simultaneous decisions by *Players* 1 and 2 respectively either to trust others ($t$) or to be cynical ($c$). Thus, $\pi^1_{tt}$ is the payoff to *Player 1* from trusting others when *Player 2* also trusts others and $\pi^2_{ct}$ is the payoff to *Player 2* from trusting others when *Player 1* is cynical. Payoffs in the game of trust satisfy the inequalities $\pi^1_{tt} > \pi^1_{ct}$, $\pi^1_{cc} > \pi^1_{tc}$, $\pi^2_{tt} > \pi^2_{tc}$ and $\pi^2_{cc} > \pi^2_{ct}$, so the Stag Hunt game of trust is a coordination game. The Stag Hunt game of trust is *symmetric*, in the sense that the particular identities of the players are immaterial for the payoffs and agents do not condition their strategy choices on the particular identities of the agents against whom they are playing (Samuelson, 2002). Thus, $\pi^1_{tt} = \pi^2_{tt} = \pi_{tt}$ and $\pi^1_{cc} = \pi^2_{cc} = \pi_{cc}$; and $\pi^1_{ct} = \pi^2_{tc}$ and $\pi^2_{ct} = \pi^1_{tc}$.

Mutual trust is the 'payoff dominant' or Pareto Optimal N. E. that yields the highest aggregate payoff. Players will hunt stag if and only if they trusted each other to always hunt stag, and this would involve trusting each other to trust each other to always hunt stag, *ad infinitum*. Among other things, this would also involve trusting each other never to hunt hare either as a mistake or as a safeguard against unilateral defection by the other player, or for any other reason. Thus, any stag-hunting society will necessarily be characterised by a high level of trust, where people believe in the general trustworthiness of others; and the belief-

based and the behaviour-based definitions of trust will be consistent with each other.

Mutual cynicism is the risk dominant N. E. that satisfies the general condition $(\pi^1_{cc} - \pi^1_{tc})(\pi^2_{cc} - \pi^2_{ct}) \geq (\pi^1_{tt} - \pi^1_{ct})(\pi^2_{tt} - \pi^2_{tc})$ or, given the symmetry of the game, $\pi^1_{ct} + \pi_{cc} \geq \pi^1_{tc} + \pi_{tt}$. If this inequality holds strictly, mutual cynicism will be strictly risk dominant. The risk dominant equilibrium is therefore equivalent to the strategy pair, in this case mutual cynicism, which secures the highest expected payoff when the two strategies (trust and cynicism) are played with equal likelihood: As the probability of playing trust and cynicism are both 1/2, the expected payoff to cynicism will be ½ $(\pi^1_{ct} + \pi^1_{cc})$, which is larger than the expected payoff to trust, which will be ½ $(\pi^1_{tt} + \pi^1_{tc})$. As it is difficult for a player acting alone to capture a stag, the payoff is lowest for an independent stag campaign. Without any guarantee that *Player 1* will hunt stag, *Player 2* will therefore be concerned that she might end up with the lowest payoff if *Player 1* for some reason (even a miscalculation) chose to hunt hare. Each player would therefore be tempted to hunt hare for a minimum possible payoff if she could not *trust* the other player to always hunt stag, even if a hare were to cross his path. Thus, with a general lack of trust among players, hare hunting by both players will also be a Nash Equilibrium. As with Stag hunting, where neither player would want to switch unilaterally to hare hunting once both players are hunting stag, neither player would want to switch unilaterally to stag hunting once they are both hunting hare.

In the game of trust with Stag Hunt payoffs, mutual trust and mutual cynicism are the two pure strategy Nash Equilibria in the sense that for each and every player, trusting others (cynicism) yields the highest payoff, and is therefore the best reply, when the other player also trusts others (is cynical). There is also a mixed strategy Nash Equilibrium given for ***Player 1*** as, "Trust others with probability $x = \frac{\pi^1_{cc} - \pi^1_{tc}}{(\pi^1_{tt} + \pi^1_{cc}) - (\pi^1_{ct} + \pi^1_{tc})}$ and Be Cynical with

probability $(1 - x)$." The corresponding mixed strategy would have ***Player 2*** trust others with probability $y = \frac{\pi^2_{cc} - \pi^2_{ct}}{(\pi^2_{tt} + \pi^2_{cc}) - (\pi^2_{ct} + \pi^2_{tc})}$.

## 3. Endogenous Effects and the Supermodularity of the Trust Game

Assuming random pairwise matching and adaptive play, Young (1993) showed that best-reply-to-recent-sample-evidence dynamics (BR-RSE adaptive play) assures by its convergence properties that people would conform to high or low trust norms and that medium trust is unstable, but the supermodularity of the trust game would itself have suggested these very results. To demonstrate this, the game of trust is first presented in a general way and adaptive play is briefly reviewed, then the equivalence of the theory of supermodular games and the evolutionary theory of trust is established.

*The Game of Trust, Played Adaptively*

In BR-RSE adaptive play of, individuals who are boundedly rational and do not actually know the trust stance that will be adopted by others, choose to trust or not to trust others as a best response to the trust stance they *expect* those others to adopt. Their payoffs are lower when their expectations are incorrect, and this causes them to adjust their strategies. A complication is that the trust game is characterised by strategic uncertainty due to the existence of multiple equilibria. For individuals to play best response strategies, they must all therefore form mutually consistent strategies at the *same* equilibrium point. An assumption of rational expectations would resolve the coordination problem quite simply of course, but it would do so by imbuing individuals with a hyper-rationality that would border on omniscience.

Instead, BR-RSE adaptive play assumes only that individuals can adjust their strategies in

favour of those that yield higher payoffs. As the game is played recurrently with different rivals, individuals are exposed to the strategies adopted by others more generally in society and this influences their choices so that ultimately everyone learns to play the game's equilibrium. Trust norms emerge from this adaptive process, allowing individuals to coordinate their trust decisions by following the existing trust norm. The possible trust norms correspond to states in which (almost) everyone is trusting or (almost) everyone is cynical. Medium trust, which corresponds to a mixed strategy equilibrium in which the probabilities of choosing one or the other action is equivalent to the population shares choosing those actions, is unstable and therefore is not a norm.

Coordination on some particular trust norm is possible as long as there are no mistakes, or as long as the mistakes are not too 'dense' to accumulate into a large shock. When there are persistent shocks, society will lurch from one trust norm to the other but in the long run society will tend to coordinate on the risk-dominant low trust norm. This result was proved under the assumption that shocks became vanishingly small, but Young (1998) clarified that it would also hold for non-vanishing shocks. This is because the risk-dominant equilibrium has the larger basin of attraction so once society is in that state it will tend to remain there. A relatively larger number of shocks or mutants will be required to tip society out of the low trust state than would be required to tip a high trust society into a low trust state.

BR-RSE adaptive play of the game of trust therefore implies that societies will in the short and medium term tend to coordinate on either high or low trust, but in the long run, most societies will be characterised by a culture of low trust.

*The Results Equivalence of Supermodularity and Adaptive Play of the Trust Game*

Supermodular games are games with strategic complementarities or games for which the marginal return to adopting a strategy increases as others adopt that strategy. In other words, supermodular games have increasing best response functions. It is well known that the multiplicity of equilibria in coordination games is due to the presence of strategic complementarities, which produce a positive feedback from social to individual behaviour. The results equivalence of adaptive play and the supermodularity properties of the trust game will be discussed in terms of the stage game $\mathcal{T} = (\Omega_i, \pi^i; i \in N)$ of the recurrent game of trust. $\Omega_i = \{0, 1\}$ is the strategy set available to each player $i = \{1, 2, \ldots, N\}$ and $\pi^i: \Omega = \Omega_i \times \Omega_j \rightarrow \mathbb{R}$ is the payoff function, $j$ referring to any player other than player $i$. Each player chooses a pure strategy $\omega_i \in \Omega_i$.

This stage game with Stag hunt payoffs is itself a supermodular game because $\Omega_i$ is a subset of real numbers with the natural order, according to which $1 > 0$, and the payoff function exhibits *increasing differences* in the sense that the incremental return from trusting others is greater as others are also more trusting. *Fig. 2* below illustrates the nature of strategy space a bit more formally. Writing $i$'s choice of trusting and not trusting respectively as $1_i$ and $0_i$, and correspondingly for $j$, increasing differences means that

$$\pi^i(1_i, 1_j) - \pi^i(0_i, 1_j) \geq \pi^i(1_i, 0_j) - \pi^i(0_i, 0_j) \qquad (1)$$

This simply says that a player's payoff to increasing her strategy and becoming more trusting is greater if rivals also have increased their strategies and have become more trusting. The game of trust in *Fig. 1* above has increasing differences: Given that $\pi^i_{cc} > \pi^i_{tc}$, the incremental payoff to becoming trusting when others are cynical is actually negative while, given that $\pi^i_{tt}$

$> \pi^{i}_{ct}$, the incremental payoff to becoming trusting is positive when others are trusting.

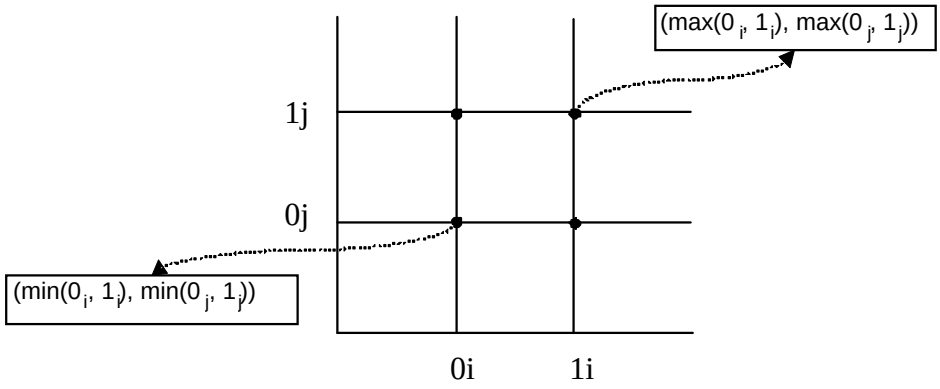


*Fig. 2. The Strategy Space for the Trust Game as a Sublattice of* $\mathbb{R}^2$

The theory of supermodular games has the following basic results and implications for the trust game, which will be stated without proofs, as in Vives (2007). These are that the supermodularity of the trust (stage) game assures that the payoff function has a fixed point and that the corresponding set of equilibria has bounds in the form of a 'largest' element (mutual trust) and a 'smallest' one (mutual cynicism). These equilibria are symmetric, but more importantly, mutual trust is Pareto Optimal and mutual cynicism is Pareto inferior.

These basic results extend to the recurrently played game of trust discussed above. Milgrom and Roberts (1990) show for a wide class of adaptive strategy adjustment processes that would include BR-RSE adaptive play,[4] the bounds of the solution set in supermodular games

[4] See Hofbauer and Sandholm (2002) for a related discussion; also Hopkins (2002). Milgrom and Roberts (1990) specifically note that the class of adaptive processes they consider includes fictitious play, where the latter involves people computing the observed frequency distribution of plays in the *entire* history to date as the

are always the greatest and smallest Nash equilibria (and that these Nash equilibria are also the bounds of the set of serially undominated strategy profiles). Milgrom and Roberts (1990) thus go beyond existence of the extremal (Nash) equilibria to elaborate on how these equilibira might be achieved. The dynamic processes they identify will converge on mutual trust or mutual cynicism – just as in the application of the Young (1993) model to the unperturbed game of trust. Similar results are proven by Dixon and Somma (1999) in a model of simultaneous play of a game such as $\mathcal{T}$ or a "mean-defined supermodular game" in which the payoff to each player depends explicitly on the average of the population play (as against perceptions of the population frequency of strategies) and on her own strategy. Under payoff monotonic dynamics, the asymptotically stable[5] equilibria of mean-defined supermodular games are their strict symmetric Nash equilibria, which correspond to the norms of mutual trust and mutual cynicism.

The theory of supermodular games goes even further. Though the BR-RSE adaptive process only envisages people playing pure strategies, mixed strategies are important as because they reflect a fundamental continuity in trust decisions that seems intuitively appealing. In that case, the (unstable) mixed strategy referred to population shares of individuals playing pure strategies. Supermodular games allow people to play properly mixed strategies. Echenique and Edlin (2004) show that mixed strategy Nash equilibria are unstable for myopic learning dynamics in games with monotonically increasing best response functions, which would

---

basis for making best response replies. Young (1998b) discusses the close relationship between fictitious play and BR-RSE adaptive play.

[5] A state is asymptotically stable if the dynamics ensure that any solution path that starts close enough to that state will necessarily converge to that state. A related but weaker stability concept, Lyapunov stability, requires that the dynamics keep solution paths that enter a neighbourhood of a stable state close to that state, without requiring convergence.

certainly include BR-RSE adaptive play of the trust game with Stag hunt payoffs.

The implication of the Echenique and Edlin (2004) result is of course that, while the pure strategy Nash equilibira in high trust and low trust are stable, medium trust equilibria, interpreted either as mixed strategy equilibria or as population shares corresponding to the probabilities in the mixed strategy equilibria of individuals playing high and low trust pure strategies, are unstable. As with BR-RSE adaptive play, when medium trust is observed, it is not a self-enforcing equilibrium but some exogenous mechanism or some set of monitoring and enforcement institutions must sustain it, preventing medium trust countries from collapsing to low trust.

## 4. The Limited Role of Formal Institutional Reform in Raising Trust

If formal institutions function to prevent the collapse of trust in unstable medium trust countries, could these institutions then also help to increase trust in low trust economies, either by making them medium trust economies or by catapulting low trust economies to self-enforcing high trust states? If so, then adopting these institutions where they do not now exist, or improving them where they are weak, should be enough to raise the level of trust in low trust countries.

The discussion in this section places emphasis on formal economies and formal institutions, but most countries have informal economies, some of which are very significant, that are governed mainly by informal institutions such as norms of trusting others, honour, honesty – or their opposites. Importantly, agents who operate in the formal economy are governed both by formal and informal institutions. The informal institutions that govern behaviour in the formal economy might be different from or similar to the ones that operate in the informal

economy. Many interesting issues[6] pertaining to trust and raising the level of trust in an economy are overlooked by the focus on the formal economy, but the clarity of the results hopefully justifies the simplification. Additionally, the discussion in this section is comparative-static in the sense that it abstracts from the process of institutional reform that has taken place for example in transition economies. It is therefore about institutional reform and change 'after the dust would have settled.' In doing so, many of the details of the process of reforming institutions are being suppressed and many assumptions are being made. In particular, the assumption here, that institutional reform and strengthening in countries with weak institutions would constrain others to be more trustworthy, is premised on an assumption that people would have confidence that the revamped formal institutions will indeed work efficiently and fairly, or that people would have 'political' trust in institutions.[7]

The first step in the analysis is to note that good formal institutions constrain others (in the formal economy) to be trustworthy or enhance the trustworthiness of others, albeit extrinsically. Weak formal institutions, a defining characteristic of most developing countries, allow others to take advantage of individuals who are sufficiently trusting to enter explicit and implicit contracts. It is not that others will necessarily take advantage of a trusting person when formal institutions are weak; and indeed, trusting others enough to enter into contracts may even be to everyone's mutual advantage. It is just that with weak institutions, if others wanted to take advantage of a person's trusting behaviour, they *could* do so, and could do so with impunity. A person who would otherwise be willing to enter into

[6] An important question for future research for example, is the relationship between institutional reform, trust and the size and growth of the informal economy.

[7] For a useful discussion of the trust and the process of reforming formal institutions, see Rose-Ackerman (2001a, b).

transactions that require her to trust others might therefore be more cautious when the formal institutional framework is weak, judging that the system is incapable of ensuring that others would be trustworthy.

Even if *intrinsic* trustworthiness were crowded out by stronger formal institutions, a possibility noted by Bohnet and Baytelman (2007), an individual in such a reformed environment would have greater confidence that others would act 'as if' they were trustworthy whether they were so inclined or not. Uslaner (2002) refers to this as a 'top down' approach or view, which holds that good (efficient and fair) formal institutions would cause people to have greater 'faith' in each other. True, Uslaner (2002) invokes a different transmission channel from good formal institutions to greater trustworthiness than the one discussed here. For him, good formal institutions would limit the corrupt behaviour of elites whose actions would otherwise erode trustworthiness in the wider society by a trickle-down effect that he calls the 'rotting fish' effect, in reference to the Chinese proverb that the 'fish rots from the head down.' But whether formal institutions function by preventing fish from beginning to rot or enhancing effective trustworthiness, the overall effect is the same. It is even likely that institutional change would generate network externalities such that as more individuals have greater confidence in the trustworthiness of others, it becomes more rewarding to trust others in the sense of assuming that they are trustworthy.

As individuals become more willing to trust others on account of an increase in their trustworthiness, the average level of trust in the economy and particularly, the average level of trust in the formal economy,[8] would increase. Measuring the quality of institutions with a

[8] The statement that institutional reform might ultimately lead to higher average trust in the formal economy doesn't preclude high average trust in the informal economy, maintained by various social norms. Nor does it

numerical index,[9] what this means is that trust increases monotonically with the quality of institutions, which is a *parameter* in the game of trust. But more than this, one could use the experience of transition countries as a natural thought experiment, to ask whether the marginal return to trusting others increased in the post-reform period, after institutions were revamped. It is difficult to draw any conclusion about this because the reform process is still ongoing in most countries, and at any rate, people still have not fully adjusted to the new environment even in those countries where the process is most advanced. Indeed, as the adjustment takes place, there is some evidence of a reduction in trust among close friends and family, which in the pre-reform period had been an important mechanism for coping with the earlier dysfunctional institutional system. Rose-Ackerman (2001a) notes however that the average level of interpersonal trust (i.e., trust among relative strangers) is not low in post-socialist countries. She reports that in the New Democracies Barometer of Rose and Haerpfer (1998), average trust is 51% in the post-socialist countries covered in the survey. This survey uses a different methodology from the World Values Surveys (2006) data on which this thesis research is based, but the latter also suggests that though there has been no dramatic increase in interpersonal trust after the early reforms in post-socialist countries,

---

necessarily mean that the informal economy will shrink and participation in the formal economy will increase. This is because there are other factors, mainly tax rates and the efficiency and fairness of laws and regulations in the formal economy, which determine the relative size and growth of the two sectors. Finally, it is even possible for average trust in the informal economy to be inversely related to trust (and institutions) in the formal economy, so that an improvement in formal institutions crowds out trust in informal sector, even to extent of lowering the national level of trust.

[9] Such indices include the *International Country Risk Guide* on the risk of expropriation and corruption; the *Business Environmental Risk Intelligence* on contract enforceability and infrastructure quality; the *World Bank's* Governance Matters summary measure of governance indicators; *Freedom House's* Political Rights summary index and *Polity IV's* various summary indicators of democracy scores and constraints on the executive.

there has been no collapse either. Taken together, these two survey results would seem to indicate that the level of trust, and particularly, the optimal or equilibrium trust strategy for individuals in an economy, is monotone non-decreasing in the quality of the formal institutions of that economy.

As a supermodular game, the recurrent game of trust with random matching and BR-RSE adaptive play therefore has monotone best response functions or increasing differences in the trust decisions of others. It is also monotone non-decreasing, and is quite possibly monotone increasing, in the quality of formal institutions. In this game, which is the stage game $\mathcal{T}$ played recurrently, individuals' payoffs $\pi^i(\omega_i, \omega_{-i}; \theta)$ depend on their own actions $\omega_i \in \{0, 1\}$, the actions of all others $\omega_{-i} \in \{0, 1\}^{N-1}$ and the quality of formal institutions $\theta$; and the payoff function has increasing differences in $(\omega_i, \omega_{-i})$ and non-decreasing differences in $(\omega_i, \theta)$. While it is true that the strategy set is discrete, the strategies are ordered as discussed above, with trusting behaviour being greater than cynicism. Clearly, institutional reform and change will be represented by an increase in the parameter[10] $\theta$, but it is equally clear that the result of that reform cannot be predicted by using the Implicit Function theorem to do a comparative static analysis.

Fortunately, the theory of montone comparative statics of Milgrom and Shannon (1994) applies to supermodular games with increasing best response functions and non-decreasing differences in the salient parameter, without requiring continuous choice variables, or for that matter, any of the usual assumptions of convexity and differentiability. By this theory, ***both*** *the largest and smallest equilibrium points in the recurrent game of trust, respectively low and high trust, will increase with an increase in the value of the institutional parameter; and*

[10] Note that representing institutions as a parameter implies that institutions are inflexible. One way to develop this discussion is to allow for flexible institutions but representing institutions as a variable.

*starting from either equilibrium best reply (and BR-RSE) dynamics will lead to a larger or higher equilibrium level of trust after institutional reform and change*. High trust and low trust economies will remain in their high and low trust states however. In particular, for some increase in $\theta$, low trust economies will not be catapulted into high trust states and the increase in the low trust equilibrium might be quite marginal. This is an even more sobering result when it is considered that the direct costs (e.g., redrafting legislation, Parliamentary meetings and debates) and the indirect costs (e.g., social unrest) of institutional reform might be very high.

What this simply means is that a country that is caught in a vicious cycle of low trust will be unable to attain high trust as an equilibrium by simply undertaking institutional reforms, however comprehensive and committed those reforms might be. Emphasising that "The major new thrust of development economics in the 1990s was recognition of the crucial role of institutions in permitting an economy to function effectively," Kuczynski and Williamson (2003) identify a number of institutional reforms as key to 'restarting growth and reform' in Latin America, and by implication in any other poor, developing countries. They contended that the 'first-generation' reforms of the so-called Washington Consensus were geared at liberalising and stabilising these economies, but that growth had come to a stop after some initial success because these countries were prone to crises and needed to embark on 'second-generation' reforms to their institutions. Of course, as institutional reform was needed to liberalise and stabilise the economies in question in the first instance, what they were simply recommending was *more* institutional reform. The monotone comparative static results noted above imply however that the reform of formal institutions will have only limited success in raising the level of trust, which is itself regarded as fundamental to economic growth (Knack and Keefer, 1997; Zak and Knack, 2001). It is not of course that institutional reform will not

raise trust in low trust economies; it will. But institutional reform will only raise the average level of trust associated with the low trust equilibrium, and will do so only marginally. It will not be enough to dislodge countries from low trust traps and catapult them into high trust equilibria.

If the 'rotting fish' effect of Uslaner (2007), that good formal institutions would limit the corrupt behaviour of elites whose actions would otherwise erode trustworthiness in the wider society by a trickle-down effect, implies that institutional reform will enhance trust in the economy, then this effect will be rather small. What prevents it from being larger is the *raccomandazione* effect, also discussed in Uslaner (2007), which refers to the Italian practice or 'culture' of soliciting favours of senior officials and other people with influence in a society. The *raccomandazione* effect is about all those norms, corruption included, which have the effect of eroding trustworthiness in a society and creating a low trust norm. The monotone comparative static result that formal institutional reform will increase the low trust equilibrium in low trust economies effectively involves both the positive 'rotting fish' effect and the negative *raccomandazione* effect. Formal institutions, which operate in the formal economy, are just the 'hardware' that could create a trusting environment; but if the 'software' of trusting informal institutions were missing (or only available in the informal economy for example), the economy would remain stuck in a low trust equilibrium even after a positive institutional shock.[11]

[11] Paraphrasing a comment by an Indian journalist about the 'uncertain health' of democracy in India, that "We have the hardware of democracy, but not the software, and that can't be borrowed or mimicked," quoted by Constable (1999) and referred to by Uslaner (2007).

The limited role of formal institutions was actually implied by the social interactions models that were estimated in Singh (2010), which emphasised that exogenous effects cannot generate multiple equilibria because they do not involve a feedback from social to individual trust. Institution reform at best makes everyone a little more trusting, but low trust is rather like a poverty-trap, and what is required for countries to get out of this trap is something in the nature of a 'tipping mechanism.'

## 5. Tipping: From Low Trust to High Trust

With most of the world characterized by low trust, it would be useful to consider whether low trust countries could become high trust even though low trust is an equilibrium or a trap. BR-RSE adaptive play suggested that this is indeed possible if there are enough 'errors' in which people irrationally choose to trust others in a low trust environment. This section elaborates on this idea using the supermodularity properties of the game of trust.

Heal and Kunreuther (2006) show that the increasing differences property of supermodular games is key to 'tipping' in the sense of being a sufficient condition for a society to jump from one equilibrium to another. In particular, they show that if and only if a symmetric game such as the game of trust has 'sufficient increasing differences' would its equilibria exhibit the tipping phenomenon. Imagine that the 2 × 2 stage game specified above is played simultaneously such that a player's payoff depends on her decision to trust or not to trust others and on the trust decisions made by everyone else. As before, let $1_i$ and $0_i$ denote $i$'s decision to trust or not to trust others respectively and let $\Omega_{-i}$ be the vector of trust decisions made by everyone else. Assume that there is some vector $\Omega'_{-i} \geq \Omega_{-i}$ such that there is at least one more trusting agent in vector $\Omega'_{-i}$ than in $\Omega_{-i}$. Increasing differences in the game of trust implies that

$$\pi^i(1_i, \Omega'_{-i}) - \pi^i(0_i, \Omega'_{-i}) \geq (>)\ \pi^i(1_i, \Omega_{-i}) - \pi^i(0_i, \Omega_{-i}), \text{ where } \Omega'_{-i} \geq (>)\ \Omega_{-i} \qquad (2)$$

Increasing differences are implied by the existence of strategic complementarities in the game of trust. Supermodularity itself implies comprehensive coordination, or coordination by everyone, on either the high or the low trust equilibrium but the policy maker's interest is in whether a low trust society can be made to 'tip' into a high trust equilibrium. Of course, in a high (low) trust economy there will always be some small share of the population that would choose low (high) trust strategies but this share will not be able to disturb the high (low) trust equilibrium, which would be self-enforcing and would tend to perpetuate itself.

Let $T \subset N$ be an arbitrary subset of players. As with every other player in $N$, players in $T$ choose between being cynical and trusting others, i.e., $\Omega_{i \in T} = \{0, 1\}$. If the set $T$ is such that when all $i \in T$ become trusting everyone else in the game of trust also switches to trusting others as a best response strategy, then the set $T$ is said to be a tipping set $T^S$. In other words, $T^S$ is a subset of players, for whom $\Omega_{i \in T^S} = \{1\}$, which makes coordination on high trust the only equilibrium of the trust game. The adoption of trusting strategies by $T^S$ ensures that society 'tips' into the Pareto optimal trusting equilibrium.

What determines whether $T$ is a tipping set or not is the number of persons who adopt the trusting strategy. Clearly it is trivial, and would be particularly discouraging for policy makers, if $T^S$ and $N$ were of the same cardinality, or if $T^S = N$. If this were the case, there would be no tipping set, but coordination on high trust would occur because each individual was responding to the trust fundamentals and not to the actions of others in society; or as noted by Skyrms (2004), because each individual adopted the belief that others would be trusting. The policy maker would instead be interested in some $T^S$ that is a subset of $N$, and

especially in the smallest such tipping set; and the policy maker would hope that the smallest $T^S$ is indeed small relative to $N$. To the extent that a tipping set existed, the social planner would need only influence a relatively small number of individuals to adopt high trust strategies for society to tip into a high trust equilibrium.

The following proposition by Heal and Kunreuther (2006) is therefore important:

PROPOSITION 1. *Under the assumption of increasing differences there exists a nontrivial* ***tipping set*** *that would cause the equilibrium to flip from low trust to high trust.*

PROOF (due to Heal and Kunreuther (2006)). Increasing differences implies that as there are more trusting persons in society, the incremental return to $i$ of switching to trust increases. Hence, writing the number of players choosing a particular strategy as a superscript and the identity of the player choosing a particular strategy as a subscript,

$$\pi_i(0^{N-1},1_i)-\pi_i(0^{N-1},0_i)<\pi_i(0^{N-2},1_1,1_i)-\pi_i(0^{N-2},1_1,0_i)<\pi_i(0^{N-3},1_1,1_2,1_i)-\pi_i(0^{N-3},1_1,1_2,0_i)< \\ ...<\pi_i(1_1,1_2,...,1_{N-1},1_i)-\pi_i(1_1,1_2,...,1_{N-1},0_i). \quad (3)$$

Clearly, the first inequality in (3) is negative because if everyone else is cynical (low trust), $i$'s best response would be to be cynical as well, not to become trusting. The next inequalities indicate that as the number of trusting persons in society increases, $i$'s marginal return to switching to trust is not just positive (which is the basic supermodularity property) but *increases*. Increasing differences in the trust game therefore implies that the incremental gain to switching from trust moves from being negative initially, and becoming less negative as more persons in society are trusting, to eventually becoming positive. The final inequality is

certainly positive because if everyone else is trusting (high trust), the incremental gain from becoming trusting can only be positive.

Writing the first positive difference as $\pi_i$ ($0^{N-t-1}$, $1_1$, …, $1_t$, $1_i$) - $\pi_i$ ($0^{N-t-1}$, $1_1$, …, $1_t$, $0_i$), the incremental gain moves from being negative when the first $t$ agents are trusting, to being positive when agent $i$ becomes trusting. The first $t$ agents therefore form a tipping set $T^S$, which necessarily exists because the sequence in (3) must switch from being negative to being positive.■

Heal and Kunreuther (2006) show further that the *smallest* tipping set consists of those agents whose decisions to switch to high trust generate the greatest "external influences" on the trust decisions of others, or exhibit the 'greatest' increasing differences. Consider the decision of an agent $i$ in (3). By switching to trust, that agent would effectively be imposing a positive externality on agent $i + 1$ that would be realised if the latter were to switch to trust. Assume that the magnitude of that externality were independent of the particular identity of agent $i + 1$ and of the decisions made by others earlier in the sequence given in (3). These assumptions mean that the size of the externality is a function only of the identity of agent $i$.

Now imagine a new sequence that starts with the same initial difference as in (3), the largest negative difference where only one person has chosen a strategy of trusting others. What is different in this new sequence is the *order* on all subsequent differences. Recall that though some of these subsequent differences are negative as well, the incremental return to trusting is always positive and by the assumptions made above, depends on the identity of the particular agent who decides to become trusting. Specifically, assume that with the same initial largest negative difference, the other differences in (3) were re-arranged in *decreasing*

*order* of the magnitude of the external influence generated by the various agents. Then the smallest tipping set would be given by the first subset of agents that would cause the new, re-ordered sequence to move from being initially negative to being positive. The smallest tipping set will therefore comprise those agents whose trust decisions have the greatest influence on the decisions of others.

The proof of PROPOSITION 1 does not make clear why anyone in $T$ would switch to high trust in a low trust environment when such a change in strategy produces a negative payoff. To be sure, the increasing differences property ensures that the negative incremental return to switching strategies is decreasing up to the point that a tipping set makes the incremental return positive, but it would nonetheless be irrational for anyone to switch to trust before this point. Low trust is an equilibrium state and the rational decision for (almost) everyone in a low trust environment is to be cynical. Likewise, the rational decision for anyone in a medium trust environment is to randomise between trusting other and being cynical. So if agents $i = 1, \ldots, t$ in $T$ will have no compelling reason to switch to a strategy of trusting others, how or by what mechanism can there ever be a tipping set which, as was seen, must exist? This is a question for public policy, and is considered in the next section.

An important theoretical point is that the Young (1993) model of stochastic adaptive dynamics emphasised mistakes in discussing norm shifts. In the context of trust, these mistakes would involve trusting decisions by boundedly rational agents in low trust environments, to effect the tipping into high trust. Heal and Kunreuther (2006) go beyond this by recognising that the identity of the agent making the mistakes is important in the sense that some agents generate greater increasing differences than others do. The implication for the discussion of trust is that there are some agents whose decisions to switch to high trust in

a low trust environment would generate greater "external influences" on the trust decisions of others, or exhibit greater increasing differences. The smallest tipping set discussed above comprises precisely these agents.

Thus, unlike the Young (1993) model, which implied that the risk dominant equilibrium was stochastically stable because of its larger basin of attraction, and therefore required a large number of trusting agents to tip a society out of the low trust equilibrium, the Heal and Kunreuther (2006) model implies that only a relatively small number of 'influential' agents is needed to form a tipping set. This is not to say however that effecting the transformation of low trust to high trust equilibria will at all be easy, or that the two models are contradictory. In both models agents will be equally averse to adopting trusting strategies in low trust environments. What the Heal and Kunreuther (2006) does is to give individuals distinct identities according to the degree to which they can influence others. Clearly, the Heal and Kunreuther (2006) result is more applicable to small countries than to large ones and it may mean that public policy to raise the level of trust may be 'easier' in the former than the latter.

## 6. Public Policy

For the policy maker to influence the level of trust in a society there must be some variables at hand that are causally related to the level of trust and that are amenable to manipulation by policy. One strategy for identifying such variables is to consider the statistically significant explanatory exogenous/contextual or environmental variables in regressions that seek to explain variations in trust or the likelihood of trusting. The main concern that arises from such an approach though is that improvements in the contextual environment can only produce a marginal increase in trust, but cannot remove a country from a low trust equilibrium. Another other strategy is to infer policy from theoretical models.

Regarding the first strategy, Zak and Knack (2001) identify several exogenous/contextual determinants of trust that are potentially policy-relevant. These include formal 'governance-type' institutions related to contract enforcement and social and ethnic heterogeneity; income (and income equality); social norms that restrain cheating; and wealth. In their discussion of the policy-relevant variables that can raise trust, Knack and Zak (2003) find however, that ethnic heterogeneity is not related to the quality of governance and formal institutions. They also exclude social norms and wealth because the former are difficult to measure and the latter is difficult to distinguish from income effects. Knack and Zak (2003) therefore conclude that education, redistributive transfers and civil liberties are the main policy variables that can be manipulated to increase trust. Of course there are several other empirical studies that also identify contextual determinants of trust, many of which are potentially policy-relevant.

The theoretical models of trust that give insight into policies for raising the level of trust effectively make a distinction should between 'long-term' and 'short-to-medium-term' policies. Short-to-medium term policy follows immediately from the supermodularity properties of the game of trust. Long-term policy follows from the limiting properties of the BR-RSE adaptive play of the trust game, and in particular involves the selection of the Pareto Optimal Nash equilibrium as one that would survive persistent shocks.

The following policy rule, based on inferences from theoretical models, indicates the general approach that must be adopted if low trust countries are to achieve high trust in a sustainable fashion.

TWO-PART POLICY RULE: 1. *Enact short term policy **on a tipping set** to dislodge a low trust equilibrium and catapult a low trust society into a high trust state.* 2. *After tipping has been accomplished, use long term policy to ensure that high trust will in fact be a stochastically stable equilibrium, so that even if there were persistent random shocks that caused people to choose non-best-response low trust strategies in the 'new' high trust state, the latter equilibrium will be sustained.*

Short-term policy requires getting a tipping set of players to deliberately and persistently 'err' by choosing to trust others in a low trust environment. By way of analogy, Heal and Kunreuther (2006) were actually concerned with how to get most if not all airlines to invest in protective measures for computer and other kinds of electronic networks that would enhance airline security in the aftermath of the September 11, 2001 attacks. They concluded that only a relatively small number (the tipping set comprising the 'most influential') of the airlines needed to adopt the protective measures for there to be a massive increase in security. This is because of tipping, due to the presence of strategic complementarities in security investment decisions among airlines. In this model, tipping could have been achieved by setting security standards for airlines that were determined to comprise the tipping set.

In the case of trust, companies, organizations, agencies and sectors that have an extensive reach in society, are all potential tipping agents. These might include banks and other financial institutions, large companies that are major employers and/or have a large market share in the local economy, large hospitals, the post office, etc. Thus, instead of attempting to raise the level of trust in the entire society, the policy maker could adopt policies that will encourage the tipping agents to be more trusting of others. The state may however have to underwrite the trusting activities of tipping agents.

*Long Term Policy*

Without any attempt at comprehensiveness, this section indicates some of the equilibrium selection mechanisms or 'policy parameters' that would make the long run resting point of the system the payoff-dominant (high trust) equilibrium. Not all of the mechanisms are equally amenable to implementation by policy, but many are feasible options. Skyrms (2004) presents an insightful discussion of most of the policy parameters discussed in the literature, while Weibull (1998) suggests a supporting research agenda.

The first policy parameter of interest is the payoff structure itself, which might perhaps be influenced by some combination of taxes and subsidies to transactions that especially require trust. Kandori, Mailath and Rob (1993) have shown that the Pareto Optimal equilibrium will be stochastically stable if the payoffs in games such as the game of trust in *Fig. 1* were such that $\pi^k_{ij} = 0$ whenever $i \neq j$. With this payoff structure, the return to mis-coordination would be zero. Such a modified game would be a pure coordination or common interest game, which is also a supermodular game. The example that immediately comes to mind is credit to persons who might not meet the formal requirements for a loan but whose informal ownership of assets is significant (de Soto, 2000). Banks routinely do not lend to such persons because their risk of default is calculated to be very high. Essentially, banks adopt a Prisoner's Dilemma framework when assessing this type of loan, but to the extent that many of the people who apply might actually be good risks, the game is essentially a pure coordination one. Once it is achieved then, a policy that makes these loans a pure coordination problem will sustain high trust.

Various combinations of the interaction structure and game dynamics would also make high trust stochastically stable, but the combinations are manifold and complex. By 'interaction structure' is meant the type of matching, the spatial extent (local/global) of interactions, and mobility across and within neighbourhoods. Game dynamics involve assumptions about behaviour and learning. Learning rules are themselves complicated, with game theory (Camerer, 2003) identifying one set of rules and computer science (Kimbrough and Lu, 2005; Fang *et al*, 2002) another.

Robson and Vega-Redondo (1996) identified the combination of random matching and imitation learning, which involves players copying the successful behaviour of *others*, as achieving the long-run selection of the Pareto Optimal equilibrium. The emphasis on learning to play the game by imitation, as against playing (myopic) best-response strategies is the source of bounded rationality in this model, but is precisely what makes the difference for long run equilibrium selection. Binmore and Samuelson (1994) find a similar result when they assume 'noisy' learning that involves imitation of a randomly chosen individual as against generally imitating the successful behaviour of others. Skyrms (2004) and Kimbrough (2005) show that efficient long run selection was possible if players were arranged on a lattice such that the local interaction or neighbourhood structure is two-dimensional and they employed an "imitate-the-best-neighbour" strategy. This latter result compares with the Ellison (1993) model of players arranged in a circle and employing myopic best-response strategies, which resulted in the selection of the risk-dominant equilibrium.

High trust will also be sustained by making the interaction structures of these models to be dynamic, thereby allowing for the 'co-evolution' of strategies and structures. This will

happen when players can simultaneously update strategies and variously choose the neighbourhoods to which they will belong (Ely, 2002), or the players with whom to interact within (Skyrms, 2004)[12] locations or across (Oechssler, 1997; Dieckmann, 1999) locations. While the 'across locations' models of partner choice assume myopic best response dynamics though, the 'within location' model assumes reinforcement learning, according to which strategies that yielded high (low) payoffs for the *decision-maker* in the past are adopted (avoided).

Identifying possible combinations of interaction structure and game dynamics is not useful unless there are policy measures that can be used to achieve the desired combination. The type of learning that people adopt might be influenced for example by education, socialisation, the structure of formal institutions, and more generally, social structure.[13] There is in fact a significant body of research, not very well known to economists, on the use of social learning theory to influence behaviour or skills in diverse contexts, e.g., problem solving (Dandurand *et al*, 2004), and managing natural resources (Pahl-Wostl *et al*, 2004), and crime (Akers, 1998). More familiar to economists though is the extensive literature on how the interaction structure might be influenced by the nature and extent of fiscal decentralisation, zoning regulations, transportation and communications systems, political systems and so forth. Designing mechanisms to influence the interaction structure, game dynamics and social learning to achieve the long run selection of high trust will clearly be a *context specific* multidisciplinary undertaking.

[12] As noted by Skyrms (2008) however, more sophisticated learning processes such as myopic best-responding will simply lead to faster achievement of high trust as a rest point.

[13] Economists have always made assumptions about learning, though often without realizing it. Chamley (2004) represents the significant progress made in economics in dealing with social learning. For a more general discussion of the economic significance of learning and cognition, see Denzau and North (1994).

The general point to note is that even if institutional *reform* is not an encouraging policy option because of its limited ability to raise the level of trust in society, formal institutions are clearly important for ensuring that once attained, high trust would be sustained as a stochastically stable equilibrium in the long run.

## 7. Conclusion

As a supermodular game or a game with strategic complementarities, the game of trust becomes amenable to comparative static analysis. A key question is whether or not the reform of formal institutions, taken to be an increase in a parameter in the supermodular game of trust, will suffice to increase the level of trust in low trust economies. The answer is that formal institutional reform has only a limited role to play in increasing the level of trust in such environments because low trust is in fact an equilibrium, or a 'low trust trap.'

What is first required is a tipping mechanism, which involves only a relatively small number of 'influential' agents to adopt the strategy of trusting others even in a low trust environment. Large-scale behavioural changes can occur for reasons other than the existence of a set of tipping agents, for example by a charismatic leader, or changes in laws, regulations or standards that cannot be circumvented, or even the momentousness of a particular incident. But when, as with interpersonal trust, change cannot be achieved in any of these other ways, tipping will be particularly important.

Beyond this, good formal institutions are required to keep countries in the high trust state once this has been attained. This is because of the long run tendency for all countries to

become low trust. Formal institutions will ensure that shocks to high trust, even if they did occur, will not accumulate into a shock of tipping proportions. Hence, not all countries will degenerate to low trust. Indeed, because of good institutions high trust countries will always exist despite the dynamics.

By that very token, there will also be medium trust economies. Such economies will be dynamically drawn to coordinate on low trust both because of their inherent short run instability and because of their stochastic instability in the very long run, but again, good institutions will prevent their collapse to a low trust state.

**References**

Akers, Ronald L. (1998). *Social Learning and Social Structure: A General Theory of Crime and Deviance,* Boston: Northeastern University Press.

Berg, J., J. Dickhaut and K. McCabe (1995). "Trust, Reciprocity and Social History," *Games and Economic Behaviour,* Vol. 10, pp. 122-142.

Berninghaus, S. and H. Haller (2010). "Local Interaction on Random Graphs," *Games* 2010, No. 1, 262-285.

Binmore, K. and L. Samuelson (1994). Muddling Through: Noisy Equilibrium Selection," University of Wisconsin-Madison. Available at http://129.3.20.41/eps/game/papers/9403/9403005.pdf

Bohnet, I. and Y. Batelman (2007). "Institutions and Trust: Implications for Preferences, Beliefs and Behaviors," *Rationality and Society* Vol. 19, No. 1, pp. 99-135.

Brock, W. and S. Durlauf (2001). "Discrete Choice with Social Interactions," *Review of Economic Studies*, Vol. 68, pp. 235-260.

Camerer, C. (2003). *Behavioural Game Theory: Experiments in Strategic Interaction*, Princeton University Press.

Chamley, C. (2004). *Rational Herds: Economic Models of Social Learning*, Cambridge University Press.

Constable, P. (1999). "India's Democracy in Uncertain Health; Upheaval Underlines System's Elasticity," *Washington Post*, April 21, 1999.

Dandurand, F., M. Bowen and T. Schultz (2004). "Learning by Imitation, Reinforcement and Verbal Rules in Problem Solving Tasks," The Salk Institute for Biological Studies, UCSD.

Denzau, A. and D. North (1994). "Shared Mental Models: Ideologies and Institutions," *Kyklos*, Vol. 47, pp. 3-31.

de Soto, H. (2000). *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*, Basic Books.

Dieckmann, T. "The Evolution of Conventions with Mobile Players," *Journal of Economic Behaviour and Organisations*, Vol. 38, pp. 93-111.

Dixon, H.D. and Ernesto Somma (1999). "Coordination and Equilibrium Selection in Mean Defined Supermodular Games under Payoff Monotonic Selection Dynamics," *Discussion Papers in Economics*, Department of Economics and Related Studies, University of York.

Echenique, F. (2001). "A Characterisation of Strategic Complementarities," *Institute of Business and Economic Research*, Paper E0-299, Department of Economics, University of California, Berkeley. Available at http://repositories.cdlib.org/iber/econ/E01-299.

Echenique, F. and Aaron Adlin (2004). "Mixed Equilibria are Unstable in Game of Strategic

Complements," *Journal of Economic Theory*, Vol. 118, pp. 61-79.

Ellison, G. (1993). "Learning, Local Interaction, and Coordination," *Econometrica*, Vol. 61, No. 5, pp. 1047-1071.

Ely, J. C. (2002). "Local Conventions," *Advances in Theoretical Economics*, Vol. 2, Issue 1, pp. 1-30.

European and World Values Surveys (2006). Four-Wave Integrated Data File, 1981-2004, v.20060423, 2006, *The European Values Study Foundation and World Values Survey Association*. Aggregate File Producers: ASEP/JDS, Madrid, Spain/Tilburg University, Tilburg, the Netherlands. Aggregate File Distributors: ASEP/JDS and ZA, Cologne, Germany. Available at http://www.worldvaluessurvey.org/

Fang, C., S. Kimbrough, A. Valluri, Z. Zheng and S. Pace. (2002). "On Adaptive Emergence of Trust Behaviour in the Game of Stag Hunt," *Group Decision and Negotiation*, Vol. 11, pp. 449-467.

Heal, G. and H. Kunreuther (2006). "Supermodularity and Tipping," Working Paper 12281, *National Bureau of Economic Research*. Available at http://www.nber.org/papers/w12281.

Hofbauer, J. and W. Sandholm (2002). "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, Vol. 70, No. 6, pp. 2265-2294.

Hopkins, E. (2002). "Two Competing Models of How People Learn to Play in Games,"

*Econometrica*, Vol. 70, No. 6, pp. 2141-2166.

Kandori, M., G. Mailath and R. Rob (1993). "Learning, Mutation and Long Run Equilibria in Games," *Econometrica*, Vol. 61, No. 1 (Jan. 1993), pp. 29-56.

Kimbrough, S. O. (2005). "Foraging for Trust: Exploring Rationality and the Stag Hunt Game," *Lecture Notes in Computer Science*, Vol. 3477, pp. 1-16. Paper Presented to the Third International Conference, iTrust 2005, Paris, France, May 23-26, 2005. Available at http://opim-sun.wharton.upenn.edu/~sok/.

Kimbough, S. and M. Lu (2005). "Simple Reinforcement Learning Agents: Pareto Beats Nash in an Algorithmic Game Theory Study," *Information Systems and E-Business Management*, Vol. 3, pp. 1-19.

Kuczynski, P. and J. Williamson (Eds.) (2003). *After the Washington Consensus: Growth and Reform in Latin America*, the Peterson Institute of International Economics, Washington.

Knack, S. and P. Keefer. (1997). "Does Social Capital Have and Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, Vol. 112, No. 4 (Nov.), pp. 1251-1288.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982). "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, Vol. 27, pp. 245-252.

Lewis, D. (1969). *Convention: A Philosophical Study*, Cambridge, Mass: Harvard University Press.

Manski, C. (1993). "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, Vol. 60, No. 3, pp. 531-542.

McCabe, K., M. Rigdon and V. Smith (2007). "Sustaining Cooperation in Trust Games," *The Economic Journal*, Vol. 117, No. 922, pp. 991-1007.

Milgrom, P. and J. Roberts (1990). "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, Vol. 58, No. 6 (Nov. 1990), pp. 1255-1277.

Milgrom, P. and C. Shannon (1994). "Monotone Comparative Statics," *Econometrica*, Vol. 62, No. 1, pp. 157-180.

Oechssler, J. (1997). "Decentralisation and the Coordination Problem," *Journal of Economic Behaviour and Organisation*, Vol. 32, pp. 119-139.

Pahl-Wostl, C., and M. Hare**.** (2004). Processes of social learning in integrated resources management. *Journal of Applied and Community Psychology* **14**:193-206.

Raymond, L. (2006). "Cooperation without Trust: Overcoming Collective Action Barriers to Endangered Species Protection," *The Policy Studies Journal*, Vol. 34, No. 1, 37-57.

Robson, A. and Vega-Redondo (1996). "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," *Journal of Economic Theory*, Vol. 70, pp. 65-92.

Rose and Haerpfer (1998). *New Democracies Barometer V: A 12-Nation Survey*, Studies in Public Policy 306, Centre for the Study of Public Policy, University of Strathclyde, Glasgow.

Rose-Ackerman, S. (2001a). "Trust and Honesty in Post-Socialist Countries," *Kyklos*, Vol. 54, pp. 415-444.

Rose-Ackerman, S. (2001b). "Trust, Honesty and Corruption: Reflections on the State-Building Process," *Yale Law School, John M. Olin Centre for Studies in Law, Economics and Public Policy*. Available at http://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?article=1013&context=lepp_papers.

Singh, T. (2010). "To Trust or Not to Trust Others: Norms, Interactions and Institutions," PhD Thesis, Department of Economics, University of Kent.

Skyrms, B. (2001). "The Stag Hunt," Presidential Address, Pacific Division of the American Philosophical Association.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*, Cambridge University Press.

Skyrms, B. (2008). "Trust, Risk and the Social Contract," *Synthese*, 160: 21-25.

Skyrms, B. and R. Permantle (2000). "A Dynamic Model of Social Network Formation,"

*Proceedings in the National Academy of Sciences of the USA*, 97, pp. 9340-9346.

Uslaner, E. (2002). "Trust and Corruption," Paper presented at the 2002 Annual Meeting of the American Political Science Association, Boston, USA.

Vives, Xavier (2007). "Supermodularity and Supermodular Games," *Occasional Paper* No. 07/18, May 2007, IESE Business School, University of Navarra.

Weibull, J. (1998). "What Have we Learned from Evolutionary Game Theory so far?" *Scandinavian Working Papers in Economics*, No. 487. Available at http://swopec.hhs.se/iuiwop/abs/iuiwop0487.htm.

Young, H. P. (1993). "The Evolution of Conventions," *Econometrica*, Vol. 61, No. 1, pp. 57-84.

Young, H. P. (1998). "Individual Learning and Social Rationality," *European Economic Review*, Vol. 42, pp. 651-663.

Zak, P. and Knack, S. (2001). "Trust and Growth," *The Economic Journal*, Vol. 111, No. 470, pp. 295-321.

Zak, P. and Knack, S. (2003). "Building Trust: Public Policy, Interpersonal Trust and Economic Development," *Supreme Court Economic Review*, Vol. 10, pp. 91-107.