Tjalling C. Koopmans Research Institute

How to reach the authors

Please direct all correspondence to the first author.

Stephanie Rosenkranz
Utrecht University
Utrecht School of Economics
Vredenburg 138
3511 BG Utrecht
The Netherlands
CEPR
Email: s.rosenkranz@econ.uu.nl

Utz Weitzel
Utrecht University
Utrecht School of Economics
Vredenburg 138
3511 BG Utrecht
The Netherlands
E-mail: u.weitzel@econ.uu.nl

IS IT TRUST WE MODEL? AN ATTEMPT TO CALCULATE THE NON-CALCULATIVE

Stephanie Rosenkranz[a]
Utz Weitzel[b]

[a]Utrecht School of Economics
Utrecht University
CEPR

[b]Utrecht School of Economics
Utrecht University

Abstract
In this paper we characterize a situation in which non-calculative trust has to play a role in the decision to cooperate. We then analyze the given situation in game theoretical terms and distinguish those aspects of players' decisions that are cooperative from those that may be interpreted as being trustful. We argue that the cooperative aspect relates to incentives while the trustful (and thus non-calculative) aspect of the decision is related to the framing of the situation.

1   Introduction

Despite a generally accepted definition of "trust", the growing economic literature on this phenomenon clearly indicates that economists consider it to play an important role in economic relations.[1] "Trust" seems to influence incentives for ex post opportunism and thus alter moral hazard problems, it seems to influence behavior in situations characterized by adverse selection, and it determines (in-)completeness of contractual arrangements as well as the feasibility of self-enforcing agreements.

The *formal* economic analysis of trust relations usually proceeds along the following lines.[2] A situation is characterized in which players may choose between a cooperative and a non-cooperative action. In the traditional "trust game" the decisions are assumed to be taken sequentially, which turns the situation into a unilateral trust relation.[3] In the prisoners' dilemma decisions are taken simultaneously which is then labeled a bilateral trust relation. Most importantly, in most models the cooperative action is referred to as 'placing trust' or 'being honest' and 'honoring trust'. Straightforward analysis of those situations reveals that the cooperative action is not chosen in equilibrium.

In order to relate the analysis to real-world examples in which people obviously do cooperate quite often despite opportunistic incentives,[4] several theoretical solutions are offered of how to alter the analyzed games such that

---

[1] Since the first contributions by Kreps (1990) , Williamson (1996) and Dasgupta (1988) a huge literature has evolved in economics on that topic.

[2] See e.g. Kreps (1990) as well as Dasgupta (1988) and Williamson (1996) who were among the first to touch the topic systematically in economics, while Parkhe (1998a), Parkhe (1998b) or Noteboom, Berger, and Noorderhaven (1997) are examples for a more fundamental business approach to trust. More recently, Harvey (2002) provided economic research with an excellent survey on trust and trustworthines, and Bohnet and Croson (2004) edited an interesting special issue on trust and trustworthiness in the same journal.

[3] See Kreps (1990) for the trust game.

[4] For experimental evidence for cooperative behaviour in PD situations see Dawes (1980) or Axelrod (1984).

the cooperative action is in fact chosen in equilibrium[5]. The simplest solution is that people write explicit or implicit contracts implying that behaving opportunistically leads to sanctions which make such behavior unattractive.[6] Thus incentives for opportunism are removed. Alternatively, it is assumed that players' preferences are such that it is in fact in their self-interest to cooperate (due to social preferences, hostages, or just non-opportunistic preferences).[7] A third group of models refers to the temporal embeddedness of the relation. The relation has to be seen in a repeated context and if the time preferences of the players is such that present opportunistic behavior can be sanctioned in the future, cooperative behavior is an equilibrium outcome. Finally, also the social embeddedness of the trust relation may induce cooperative behavior in equilibrium. Provided that agents can observe behavior of other players in previous trust situations, updating of beliefs over the player's preferences may induce cooperative actions in equilibrium.[8]

All those contributions greatly advance our understanding of how to model "trust". However, confronted with the less formal discussion of what the distinctive nature of "trust" is, several questions still arise. In sociology as well as in economics, there is a common notion that "trust" is *not* purely calculative.[9] According to Noteboom (2002) we can define trust as a disposition towards trusting behavior, i.e. behavior with limited safeguards, accepting vulnerability, based on the expectation that the risk is limited.

---

[5]See Harvey (2002) for a detailed discussion of the various formal solutions to the trust problem characterized by the prisoners' dilemma.

[6]Feasibility and efficiency of explicit contracts to prevent opportunistic behaviour is extensivly discussed in the principal-agent literature. See e.g. the contributions by Hart and Holmström (1987), Bull (1987) and MacLeod and Malcomson (1989).

[7]For examples for social preferences see the survey of Fehr and Schmidt (2003). Hostages have be analyzed in this context by Raub and Keren (1993) and Raub (2003). Non-opportunistic preferences are the driving force in Dasgupta (1988).

[8]See Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982) or Buskens (2003).

[9]See Williamson (1996) who first stated this insight rather drastically. See also Coleman (1990), or Gambetta (1988) and Heimer (2001). See also Noteboom (2002), Hardin (2001) or Sabel (1993) for excellent examples of the more sociological literature.

Trustful behavior is boundedly rational in the sense that it is largely calculative, but relies on routines and tolerance levels, using subjective probabilities, that are determined or influenced by non-calculative elements. The key element of "trust" is the fact that it is to a certain degree *non-calculative*.

If we reconsider the above mentioned theoretical solutions explaining cooperative behavior in situations of trust, we realize that the first three modeling approaches (writing contracts, making assumptions on preferences and using sanctions in repeated relations) rule out the non-calculative aspect of trust. But is it then really "trust" that is modeled? Based on the understanding discussed above these models primarily analyze *cooperative behavior*. Only if we include a non-calculative element, cooperative behavior and trustful behavior are not perfectly the same. The forth class of models, offering social embeddedness as a solution, transfers "trust" to the players' beliefs about the other player's preferences. In fact, those beliefs can be highly subjective, which introduces an element of non-calculativeness. However, at the same time these models employ "reputation" as a synonym for "trust" although these two concepts are not equivalent. Where does reputation, the subjective probability not to be confronted with an opportunist, come from? And, moreover, how do we handle isolated encounters between individuals?

This brings us to the central question of this analysis. Can we incorporate a *non-calculative* element of trust into a formal model? And if so, how? In this paper we characterize a situation in which non-calculative trust has to play a role in the decision to cooperate. We then analyze the given situation in game theoretical terms and distinguish those aspects of players' decisions that are *cooperative* from those that may be interpreted as being *trustful*. We argue that the cooperative aspect relates to incentives while the trustful (and thus non-calculative) aspect of the decision is related to the framing of the situation.

Four types of non-calculative elements have been identified e.g. by Lane (2000) or Noteboom (2002): unreflective trust (neglect of risk due to naivety or cognitive dissonance), trust based on affect, trust based on routinized

4

behavior, and trust based on a default or norm. The results of our model emphasize the latter and thus point into the direction of Messick and Kramer (2001) but with the virtue of a formal analysis. This basic idea is also supported in a study by Willinger, Keser, Lohmann, and Usunier (2003), who found that it is possible to distinguish different levels of trust (as a form of social capital) in different societies.

The paper is structured as follows: In section 2 we formally model and analyze a specific situation that implicitly involves non-calculative "trust". In subsection 2.1 we present a one-shot game in which firms decide on co-operating in a joint project and in which profits are generated in a contest. Although cooperation is pay-off dominant, cooperative behaviour can also lead to unilateral spillovers. A firm's know-how (including such spillovers) determines its probability of success in this contest. In subsection 2.2 we show that, both, cooperation and non-cooperation are always equilibria. We furthermore characterize conditions under which only a specific type of firm cooperates. In section 3 we discuss the (calculative) cooperative and the (non-calculative) trustful dimensions of the results. Finally, in section 4, we outline how non-calculative trust could not only implicitly, but also explicitly be integrated into formal theoretical analysis.

2   A Situation of "Trust"

When managers negotiate a possible cooperation between their firms they often face a dilemma: In order to evaluate the full potential of the cooperation both companies might be forced to partially reveal sensitive know-how or any other kind of information without being able to effectively safeguard this disclosure, e.g. with a perfectly secure (complete) cooperation contract. Managers may well assume that the other company is an important potential partner, but without the mutual disclosure of some knowledge both parties do not know precisely how capable and competent the other company is. Even if both parties know that a successful cooperation will generate the highest profits compared to all other options, due to the public good char-

acteristics of know-how, such a disclosure is potentially dangerous. Thus, in the dilemma under consideration a firm's know-how has to be shared in order to enjoy the potential benefits of a cooperation, but this disclosure bears the risk of being exploited.[10]

According to the definition of Coleman (1990) the above described relation between the two firms qualifies as a 'trust relation': (a) Placing trust by the truster allows the trustee to honor or to abuse trust, (b) the truster regrets placing trust if trust is abused, but benefits from honored trust, and (c) the truster voluntarily places resources in the hands of the trustee without formal safeguards. Note that in our scenario, both firms are in the role of the truster as well as in the role of the trustee so that we are considering a bilateral trust relation.

For our argument it is useful to distinguish two relevant concepts: the trustfulness of the truster (the extent to which the truster places trust) and the trustworthiness of the trustee (the extent to which the trustee honors or abuses trust). The trust game introduced by Kreps (1990) suggests that the trustfulness of the truster largely depends on the expectation of the trustworthiness of the trustee. The same applies to many other trust models, including ours. This leaves the determination of the sources of trustworthiness of the trustee as one of the most important points of analysis. In general the trustee might behave trustworthy out of self-interest, which can be guided by his own material advantage in a specific relation (shadow of the future, reputation, hostages) or due to possible sanctions from authority and contractual obligations that apply apart from a specific relation. Trusting behavior on the part of the truster is then mostly calculative. On the other hand, the trustee might behave trustworthy due to motives which lie beyond his direct self-interest. In a specific relation this could be due to bonds of friendship, kinship, routines or empathy and altruism. On the macro level (i.e. not relation specific) he could behave trustworthy due to

---

[10]Note that, given complementary know-how, even firms with very little know-how face this dilemma.

values, social norms of proper conduct and moral obligation.[11] The purpose of the following analysis of a specific situation is to identify non-calculative sources of trustworthiness and to integrate them into formal modelling.

In the following two subsections we first depict the formal setting and the analytical results, before we turn to the main interpretations and insights in terms of which aspects of the analyzed decisions are cooperative and which are based on trust.

2.1 The Model

Assume there are two firms $i = 1, 2$ with types v_i, which can be interpreted as know-how levels. Both firms believe that the types v_i are independent and only have two possible values v_l (low know-how level) and v_h (high know-how level), with $v_l < v_h$. Furthermore, they believe that p is the probability that v_i equals v_h and that $(1 - p)$ is the probability that v_i equals v_l. Each firm's type v_i is private information only to that firm. In the following stochastic framework a firm's know-how represents its probability of success in a contest. The firms play the following one-shot game:

They both have two possible actions, 'cooperate' or 'do not cooperate', denoted by $A_i \in \{C, D\}$ on which they decide simultaneously. If a firm chooses action C (cooperate) it decides to signal its willingness to cooperate by disclosing (some of) its know-how. If the firm chooses D it does not disclose any know-how. If both firms choose C then they form an *alliance*. Please note that this wording allows a firm to 'cooperate' without an alliance actually being formed (provided the other party chooses D). For simplicity it is also assumed that firms cannot strategically choose the amount of know-how that spills over. Should they decide to cooperate, they reveal their type completely and truthfully and a fraction α of their know-how is disclosed.[12] Once know-how is disclosed is not possible to take it back, but it is also

---

[11]See Noteboom (2002) and Williamson (1996) for this scheme for the sources of trustworthiness.

[12]In a possible extension of the model one could relax this rather restrictive assumption so that players disclose as little as necessary to "signal" their type.

not lost for the one who disclosed it.[13] Let $\alpha \in [0,1]$ denote the degree of spillovers which is common knowledge to both firms. Whenever one firm j unilaterally reveals its know-how, the other firm i's know-how is increased by αv_j. If both firms cooperate, their know-how is always perfectly transferred at no cost. The know-how of the alliance corresponds to the sum of both firms' individual know-how.

Firms' payoffs are determined in a stochastic contest model and represent functions of their know-how. Note that a strategy of a firm i is a mapping from the set of types $\{v_l, v_h\}$ into actions $A_i \in \{C, D\}$. Let $v_i \in \{v_l, v_h\}$ with $i = 1, 2$ and $i \neq j$ be statistically independent probabilities of success for the two firms conditional on their know-how levels. For illustration purposes, assume that 'success' is defined as an invention on the basis of own or joint know-how generating a payoff of V. A firm i's payoff, given it is of type v_i and has chosen action A_i, and given firm j is of type v_j and has chosen action A_j (with $i, j = 1, 2$), is denoted by $\Pi_i(A_i, A_j, v_i, v_j)$.

Firms can find themselves in four different situations depending on the choice of both their actions at the first stage: Either none of the two firms cooperate, only firm A cooperates, only firm B cooperates, or both cooperate and an alliance is formed. For each of these situations payoffs are defined as follows:

If neither of the two firms decides to cooperate (D,D), a firm i's payoff is given by:

$$\Pi_i(D, D, v_i, v_j) = v_i (1 - v_j) V.$$

It is important to understand that although neither firm cooperates, each of them can still be successful in the contest model (i.e., both come up with a discrete invention). However, due to Bertrand competition on the product market, marginal profits erode to zero and the value V is lost for both firms if they are successful independently.

In case of unilateral know-how disclosure the revealing (cooperating, C)

---

[13] We abstract from any spillovers due to backward engineering or espionage. Therefore there is no leakage of know-how if players choose not to reveal.

firm earns

$$\Pi_i(C, D, v_i, v_j) = (1 - \alpha)v_i(1 - v_j)V,$$

which is 0 in case of complete spillovers, $\alpha = 1$.

If firm i is the receiving (non-cooperating, D) firm, i's payoff is given by:

$$\Pi_i(D, C, v_i, v_j) = v_i(1 - v_j)V.$$

Notice that the receiving firm does not directly benefit from the revealing firm's know-how, because the receiving firm only wins V if the other firm is *not* successful. Therefore, any know-how which is unilateraly revealed is "unproductive" for both firms. This is due to the fact that if a firm's know-how leads to success it does so independently of who owns this know-how.[14]

In case of simultaneuos two-sided cooperation (C,C), each of the two firms will only employ their know-how once and may be successful with either one. In such an alliance firms share the value V according to their initial contributions of know-how.[15]

$$\Pi_i(C, C, v_i, v_j) = (1 - (1 - v_i)(1 - v_j))\frac{v_i}{v_i + v_j}V$$

Now equilibria of the game can be characterized.

2.2 Analytical Results

In this section only the analytical results of the model are presented. A more detailed and intuitive discussion of the results will follow in the next section. The technically not interested reader may skip this section and directly proceed to the next one.

---

[14]If the fraction αv_j is successful both firms make the invention but Bertrand competition on the product market leads to zero profits. If the fraction $(1-\alpha)v_j$ is successful the receiving firm earns no profits because the revealing firm wins the contest.

[15]Here we assume that once know-how is revealed, the firms' types are common knowledge such that the players can adjust the sharing rule proportionately to their initial contributions. Of course also a 50:50 sharing rule can be taken into account as some parts of the literature suggest. In the remainder of the paper it will become obvious that our findings are qualitatively independent of this assumption.

In order to characterize Nash-equilibria of this game we first present the equilibrium conditions which have to be satisfied for a firm i of a certain type $v_i \in \{v_l, v_h\}$. The Nash-equilibrium strategies $A_i^*(v_i)$ of both firms are defined by the mutual best-response property. The analysis is confined to symmetric equilibria in pure strategies: Whenever a firm i of a given type does (not) cooperate, firm j of the same type does also (not) cooperate. This means that, for instance, firm i of type v_i prefers action $A_i = C$ over action $A_i = D$ (for a given action of the other firm) if this action leads to a higher expected profit, taking into account the probability to be confronted with a firm j of the same type, also choosing action $A_j = C$. In equilibrium an analogous condition has also to hold true for the other firm. We establish the equilibrium conditions for any possible outcome for each type of a firm in the appendix so that we can easily summarize the following results:

Proposition 1 *There always exist two equilibria: One in which firms of both types do not cooperate, $A_i^* = D$, and one in which firms of both types cooperate, $A_i^* = C$.*

Proof. Note that $\Pi_i(C, D, v_i, v_j) < \Pi_i(D, D, v_i, v_j)$ for all v_i, v_j. Furthermore also $\Pi_i(C, C, v_i, v_j) > \Pi_i(D, C, v_i, v_j)$ holds, since for all $i, j = 1, 2$:

$$\left((v_i + v_j - v_i v_j) \frac{v_i}{v_i + v_j} V \right) - (v_i (1 - v_j) V) = V v_i \frac{v_j^2}{v_i + v_j} > 0.$$

∎

Proposition 2 *For an intermediate level of prior probabilities $p \in [\hat{p}, \check{p}]$ only firms with low know-how cooperate, and $A_i^*(v_l) = C, A_j^*(v_h) = D$ is an equilibrium; with*

$$\hat{p} \equiv \frac{v_l}{2\alpha(1 - v_h) + v_l} \ \text{ and } \ \check{p} \equiv \frac{v_l^2}{\alpha(1 - v_h)(v_l + v_h) + v_l^2}.$$

Proof. See the appendix. ∎

For the ease of exposition the model presented in this paper is reduced to the simplest case. However, with respect to the underlying assumption on competition the results are very robust.[16]

---

[16] A one-shot contest with Cournot competition yields the same qualitative results. Com-

3 Discussion of Results

3.1 The Cooperative Decision

The first result of the model applies to all players under consideration, regardless whether they have a high or low level of know-how: (a) the building of an alliance (bilateral cooperation, (C,C)) is always an equilibrium, however, (b) the bilateral rejection to enter an alliance (D,D) also turns out to be an equilibrium. These two equilibria exist irrespective of the probability with which a player encounters a partner with a high level of know-how. This result seems contradictory, but is actually quite straightforward: If there exist spillovers, and given that the other player does not reveal know-how, the gains from not revealing outweigh the gains from cooperating for all players. While the receiving player i does not directly benefit from the other player's know-how, the unilaterally revealing player is worse off. Hence, if a player assumes that the other party will maliciously exploit any revealed know-how, not to cooperate is its best option and bilateral rejection to enter an alliance (D,D) constitutes an equilibrium.

On the other hand, given that the other player reveals know-how (cooperates), it is profitable to also reveal know-how. A player who received know-how only benefits from the other player's know-how if it shares the value V in a cooperation. For this to happen, it also has to cooperate and to reveal its know-how. Since the payoff in an alliance is always greater than independently generated payoffs, this result does not hinge on the player's own type.[17] Hence, if players assume that the other party will be cooperative, entering an alliance (C,C) constitutes an equilibrium. Note that alliances also represent the efficient outcome of the game. Hence, if players

petition could even be modeled as a first price auction or as a patent race in which the v_i represent the Poisson intensities of invention. The results also do not differ drastically from those of the second-price auction. The probability to meet a certain type is less important whereas the degree of spillovers becomes a critical factor. The calculations for those forms of competition are available upon request from the authors.

[17]Even a partner with extremely little extra know-how still raises the payoff. Note that this is subject to proportional sharing rules in an alliance and complementary know-how.

could coordinate, they would agree to coordinate on the cooperative equilibrium. The implications of this result on the role of trust will be discussed below.

The second result of the model is type-dependent: It states that players with little know-how will *always* cooperate and (at the same time) players with a high level of know-how will *not* cooperate, if the probability p to meet a player with a high level of know-how is "not too small and not too big" (with $\check{p} < p < \hat{p}$). To understand this result it is important to note that players with low know-how find it beneficial to cooperate among themselves even if competitors with high know-how do not cooperate, as long as the danger of meeting such a cheating player with high know-how is not too high ($p < \hat{p}$). For an illustration of this effect please refer to the top bar of Figure 1. On the other hand, since a player with high know-how has much more to loose, it does not find it beneficial to cooperate (although players with low know-how cooperate), unless the probability to meet a cheating player with high know-how is small ($p > \check{p}$). This is illustrated with the middle bar of Figure 1. Combined, the two effects lead to a range of medium probabilities (lowest bar in Figure 1) in which only low level players cooperate while high level players do not cooperate, so that in effect, only alliances of players with low level of know-how exist.[18]

---

[18] As the remark in the appendix shows, there does not exist an equilibrium in which only firms with high know-how cooperate. For players with high know-how to reveal their know-how, the probability to meet a cooperating player of the same type has to be sufficiently high. On the other hand, players with low know-how will then also always have an incentive to cooperate so that the equilibrium of type-independent alliances is reached. There is no interval of prior probabilities for which high know-how types cooperate and players with low know-how do not have an incentive to also cooperate.
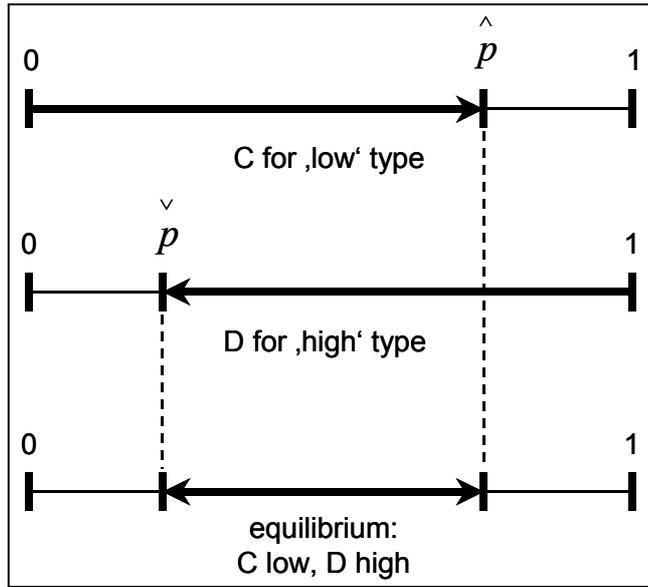
Figure 1

Figure 2 presents a summary of the multiple pure strategy equilibria explained above, which players potentially face when they consider know-how based cooperations with no or insufficient legal safeguards.
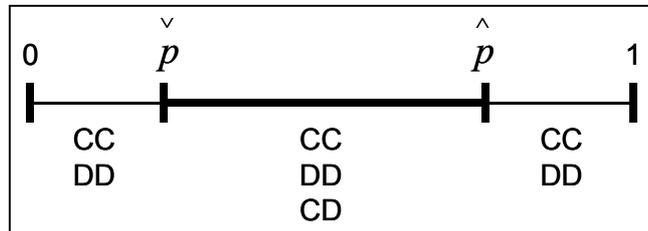
Figure 2

Finally there also exists a mixed strategy equilibrium, in which both types of players allocate more weight to the non-cooperative action.

3.2 The Trustful Decision

Let us now discuss the game in light of a trust relation. First of all, we model a bilateral situation, in which trust and cooperation are separated. Cooperation is efficient, given the other player cooperates as well, but cooperating unilaterally makes the players vulnerable. The decision to cooperate

is risky since not cooperating is also a rational strategy for the other player. Thus, this decision involves trusting behavior (i.e. behavior with limited safeguards, accepting vulnerability, based on the expectation that the risk is limited). Without any risk or uncertainty, there cannot be any room for trusting behavior.[19] Trustworthiness is here given as a preference and thus primarily based on self-interest. The degree of self-interest and thus the preferences of the players are themselves also uncertain. This allows us to also explicitly discuss the role of reputation as it is modeled in other models on "trust". The parameter p can be interpreted as being the "reputation" of a player to possess a high level of know-how. In our analysis this does not automatically imply more cooperative behavior or even more trustworthy behavior of this player. The self-interest of a player to cooperate in this framework does not only depend on his own type but also on the type of the other player. Thus in this model reputation and trustworthiness are not identical concepts.[20]

With respect to the role of trust in relationships the implications of the results in Figure 2 are twofold: On the one hand the analysis clearly reveals that trust is not an imperative for cooperative behavior, even if cooperative behavior includes the danger of being exploited. Thus, with regard to the hen-or-egg-discussion "whether trust leads to cooperative behavior or vice versa" Parkhe (1998a), our model provides a general framework in which cooperative behavior is not only possible without trust, but in some situations $(\check{p} < p < \hat{p})$ even more likely than non-cooperative behavior. From a market perspective the implication is straightforward: If bilateral cooperation is the efficient outcome of the game, then alliances can be "born" without any involvement of trust and will be observable within a given population. Hence, cooperation is not identical to "trust".

---

[19]See Noteboom (2002) on the information paradox. The research by Eckel and Wilson (2004) reveals a very interesting aspect. Their experiments show that trustful behaviour of individuals seem not to be determined by the individuals' risk attitude.

[20]Although our model is closely related to the basic model of Dasgupta (1988), this is a very important difference, making it possible to disentangle all the concepts that Dasgupta uses synonymously: trust, honesty, reputation, and cooperation.

On the other hand, from an individual's viewpoint this argumentation is flawed, since decisions in very specific situations have to be taken. Being uncertain about the behavior (and type) of the opponent induces a player to put more weight on the non-favourable (thus non-cooperative) action in a mixed strategy equilibrium. When leaving the market perspective and taking the view of an individual player, the model shows that trust can be an important additional factor that facilitates cooperations in order to overcome more inefficient equilibria. Thus, with regard to the above quoted hen-or-egg discussion, the results of our model do not simply argue against the importance of trust, but rather enable us to identify aspects of a decision problem that are in fact trustful and thus important to facilitate cooperation: Due to the multiple equilibrium outcomes as summarized in Figure 2 individuals are likely to be confronted with a serious coordination problem. Trust can here be implemented as a *selection device* of the players to coordinate each other's strategies to the preferred (efficient) outcome.

As suggested earlier we can look for the sources of trustworthiness (being the determinants of trustful behavior) on the macro level in which the relation is embedded. We can use e.g. the level of trust in a society as a selection device: The higher such social capital, the more firms should be able to coordinate their strategies towards the cooperative outcome. Fukuyama (1995) defines trust accordingly as "the expectation that arises within a community of regular, honest, and cooperative behavior, based on commonly shared norms, on the part of members of that community". He shows that trust within the society affects the performance of its institutions, including firms, as it determines people's ability to cooperate and contribute to socially efficient outcomes.[21]

---

[21]Furthermore, it has been argued by La Porta et al. (1997) that trust is more essential when interactions are infrequent or involve strangers, as in our pre-contractual framework. LaPorta et al. (1997) identify different levels of trust for Germany and France on the basis of answers to attitudinal questions, characterizing Germany as a high-trust society and France as a low-level society. Our analysis implies that, despite the stated dilemma, managers in Germany more often coordinate on the efficient cooperative equilibrium. Hence, it should be possible to verify empirically whether alliance formation is more

4 Prospect and Conclusion

Our simple situation of a trust relation between firms that are planning to engage in an alliance poses very basic questions about the role of trust in cooperative behavior. In most formal arguments that were made to clarify the role of trust with the help of game theoretical terms, trustful behavior cannot be explained in an isolated encounter. In the trust game by Kreps (1990) trust will not be placed. In contrast to Kreps (1990) our model does not assume completely irrational trust in order to explain cooperative behavior in a one shot setting. It is based on a modified perspective on trust, in which trust can be employed as a selection device for cooperative behavior without being a *conditio sine qua non.*

In contrast to the isolated encounter, in the repeated version of the game by Kreps (1990) temporal embeddedness (enabling the truster to control the behavior of the trustee) and social embeddedness (enabling the truster to modify his expectations about the incentives of the trustee) give reason to place trust. But even in those games the main weakness is that placing trust is only one of several possible outcomes. Those games all share the problem of the multiplicity of equilibria. The cooperative or trustful outcome is just one of them. This is a severe problem for the predictive power of the games. Our approach therefore turn the weakness into a virtue and emphasizes the role of trust as a selection device, to help the parties to focus on one - trustful - outcome.

In line with many other formal analyses of trust, we neither depict nor model trust explicitly in this paper (as it is not explicitly included in the game). Although we are aware of the fact that this may be interpreted as a weakness rather than a strength, we believe that the introduction of trust as a simple parameter would bring us back to ignoring the fact that trustful behavior *should* be characterized by non-calculative elements in order to distinguish it from the mere decision to cooperate.

In our model we obviously need additional assumptions (e.g. focal points,

---

successful in Germany.

norms) to explain why an individual chooses 'cooperation' in this situation as the relevant equilibrium strategy. But this model allows us to separate the influence of trust on an individual's decision to cooperate beyond pure self-interest, and also to distinguish trust from reputation. Assuming that it is in my partner's self-interest to cooperate, I must additionally belief that we perceive the situation in the same way and that we share the same norms of conduct. This insight relates our analysis to all previously discussed formal attempts to solve the "trust problem". In fact in the repeated trust game or prisoners' dilemma, and even in the 'reputation game', players have the same problem of choosing among multiple equilibria. Choosing to cooperate over a longer period is just *one* of many possible equilibrium strategies in those games. Furthermore, even solutions of games of one-shot trust situations as the prisoners' dilemma and the trust game can be related to our analysis: If we allow for a small degree of incomplete information in these game over the other player's preferences (his direct monetary payoff, or the degree of altruism or inequality aversion due to social preferences, or the subjective value of a hostage) we are immediately in the strategic framework of the situation above. Only under very restrictive assumptions we would expect such games to have a unique outcome. Hence, describing a trust relation as a situation with multiple equilibria allows us to introduce the "non-calculative element" of trust into the formal analysis by relating it to the selection principle players apply in those situations.

At this point it might be argued that in those cases in which one equilibrium is clearly payoff dominant, it should be easy for the players to coordinate their strategies on that equilibrium. The argument ignores the fact that this equilibrium is often also risk dominated.[22] If we analyze the situation in terms of the influence of trust on the decision, this implies that more trust is needed to go for the payoff dominant equilibrium. Experimental evidence suggests that individuals seems to coordinate more often on the

---

[22] In the model outlined above the cooperative equilibrium is for example the payoff dominant outcome, while the non-cooperative outcome is risk dominant for players of both types for low values of α or for players with high types for high values of α.

17

risk dominant equilibrium than on the payoff dominant outcome.[23] More-over, there is a common understanding in game theory that focal points guide human behavior and recently several attempts have been made to formalize the underlying idea of salience.[24] Schelling (1960) suggests that "[m]ost situations provide some clue for coordinating behavior, some focal point for each person's expectation of what the other expects him to expect to be expected to do." He introduces the notion of focal points to denote the systematic use of non-strategic features in strategic contexts, especially in coordination games. The clues he refers to are vaguely characterized as prominent or conspicuous, and unique, and are what Lewis (1969) originally termed salience - unique in some conspicuous respect. The notion of salience turned out to be difficult to formalize, as according to Schelling it should capture both imagination and logic. Successful approaches in this respect are the variable frame theory of Bacharach (1993), framed strategic forms of Casajus (2000), as well as the approach of Janssen (2001). Casajus (2000) nicely points out that "[i]n real life people do not perceive their decision problems only in terms of payoff differences between strategies. Besides the payoff structure, other features of the game are systematically exploited by the players to determine their actions. It matters how a player perceives the game and how he expects the others to describe it." The "imagination" is captured by the frame the players apply to the strategic situation and the "logic" is the mechanism leading to the salient outcome. The basic ideas of variable frame theory are that each player perceives the strategies as possessing a number of attributes and that the set of attributes restricts the players' perceived options by some principles of reasoning. Furthermore a commonly known probability distribution over those sets of attributes is assumed, and that players cannot impute attributes to other players that they are not aware of themselves. For simple games with ranked outcomes (so-called hi-lo Schelling games), this approach prescribes the efficient out-

---

[23] See Haruvy and Stahl (1998) for a comprehenisive test for comparing different selection principles.

[24] See e.g. Bacharach (1993), Janssen (2001) and Sugden (1995).

come as a unique solution.[25] Thus, applied to our context it would again be a purely calculative decision to cooperate without any need to refer to trust as a determinant. Confronting this result to intuition as well as to experimental results suggests that there is still a missing link. The test of Haruvy and Stahl (1998) revealed that the selection principle used most often is the so-called "level-1 bounded rationality". This means that players choose strategies with the highest expected payoff given a diffuse prior about opponent behavior. In our game we interpret this "diffuse belief" as the influence of non-calculative elements on the trustful decision. That is why we suggest that the variable-frame level-n theory of Bacharach and Stahl (2000) provides the adequate framework to model trust formally. The authors combine the fully rational variable frame theory of Bacharach (1993) and the boundedly rational level-n theory of Stahl and Wilson (1995). This theory predicts the salient outcome (to cooperate) only if sufficiently rational players' frames express the focal point. The aspect of non-calculativeness is captured by the notion that players have to determine a *subjective* probability about the behaviour of at least one pivotal boundedly rational (level 0) player in order to determin the focal point.

This would mean in our context that the players perceive the situation as one in which trust will be placed (by at least one boundedly rational player). We consider the most promising next step in modelling trust to be the extension of the variable frame level-n theory to simple games of incomplete information.

---

[25]Alternatively the idea of focal points has been formalized by Casaju (2000). His analysis of framed strategic forms is also very promising but bears the same problem as the variable frame theory if we apply it to our considerations on "trust".

5 Appendix

Equilibrium conditions

A firm of type $v_i \in \{v_l, v_h\}$ does *not* cooperate, $A_i^* = D$, if the following inequality holds:

$$
\begin{aligned}
p & \ (\Pi_i(C, D, v_i, v_h) - \Pi_i(D, D, v_i, v_h)) & + \\
(1-p) & \ (\Pi_i(C, D, v_i, v_l) - \Pi_i(D, D, v_i, v_l)) & < 0
\end{aligned}
\tag{1}
$$

A player of type $v_i \in \{v_l, v_h\}$ does cooperate, $A_i^* = C$, if the following inequality holds:

$$
\begin{aligned}
p & \ (\Pi_i(C, C, v_i, v_h) - \Pi_i(D, C, v_i, v_h)) & + \\
(1-p) & \ (\Pi_i(C, C, v_i, v_l) - \Pi_i(D, C, v_i, v_l)) & > 0.
\end{aligned}
\tag{2}
$$

Only player i of type v_l cooperate, i.e. $A_i^*(v_l) = C$ and $A_i^*(v_h) = D$, if:

$$
\begin{aligned}
p & \ (\Pi_i(C, D, v_h, v_h) - \Pi_i(D, D, v_h, v_h)) & + \\
(1-p) & \ (\Pi_i(C, C, v_h, v_l) - \Pi_i(D, C, v_h, v_l)) & < 0,
\end{aligned}
\tag{3}
$$

holds and simultaneously also

$$
\begin{aligned}
p & \ (\Pi_i(C, D, v_l, v_h) - \Pi_i(D, D, v_l, v_h)) & + \\
(1-p) & \ (\Pi_i(C, C, v_l, v_l) - \Pi_i(D, C, v_l, v_l)) & > 0
\end{aligned}
\tag{4}
$$

holds. On the other hand, only players i of type v_h cooperate, i.e. $A_i^*(v_h) = C$ and $A_i^*(v_l) = D$, if:

$$
\begin{aligned}
p & \ (\Pi_i(C, C, v_h, v_h) - \Pi_i(D, C, v_h, v_h)) & + \\
(1-p) & \ (\Pi_i(C, D, v_h, v_l) - \Pi_i(D, D, v_h, v_l)) & > 0,
\end{aligned}
\tag{5}
$$

holds and at the same time also

$$
\begin{aligned}
p & \ (\Pi_i(C, C, v_l, v_h) - \Pi_i(D, C, v_l, v_h)) & + \\
(1-p) & \ (\Pi_i(C, D, v_l, v_l) - \Pi_i(D, D, v_l, v_l)) & < 0
\end{aligned}
\tag{6}
$$

holds. It is easily seen that whenever $\Pi_i(C, C, v_i, v_j) > \Pi_i(D, C, v_i, v_j)$ for all $i, j = 1, 2$ and $i \neq j$ then (2) is satisfied and $A_i^* = C$ is always an

equilibrium. Analogously, if $\Pi_i(C, D, v_i, v_j) < \Pi_i(D, D, v_i, v_j)$ then (1) is satisfied and $A_i^* = D$.

On the other hand, if $\Pi_i(C, C, v_i, v_j) > \Pi_i(D, C, v_i, v_j)$ holds and at the same time also $\Pi_i(C, D, v_i, v_j) < \Pi_i(D, D, v_i, v_j)$ then both these equilibria exist, which is known as a 'coordination game'. Due to incomplete information in this framework there is also the potential for equilibria with $A_i^*(v_l) = C, A_j^*(v_h) = D$ or $A_i^*(v_l) = D, A_j^*(v_h) = C$ depending on the parameter values of p. These are obviously the more interesting cases.

Proof of Proposition 2. Inequality (5) is satisfied for:

$$
\begin{aligned}
0 \quad > \quad & p\left((1-\alpha)v_h(1-v_h)V - v_h\left(1-v_h\right)V\right) + \\
& (1-p)\left((v_h + v_l - v_h v_l)\frac{v_h}{v_h+v_l}V - v_h\left(1-v_l\right)V\right) \\
\Leftrightarrow \quad & p > \hat{p} \equiv \frac{v_l}{2\alpha(1-v_h)+v_l},
\end{aligned}
$$

while inequality (6) is satisfied for:

$$
\begin{aligned}
0 \quad < \quad & p\left((1-\alpha)v_l(1-v_h)V - v_l\left(1-v_h\right)V\right) + \\
& (1-p)\left((v_l + v_l - v_l v_l)\frac{v_l}{v_l+v_l}V - v_l\left(1-v_l\right)V\right) \\
\Leftrightarrow \quad & p < \check{p} \equiv \frac{v_l^2}{\alpha(1-v_h)(v_l+v_h)+v_l^2}.
\end{aligned}
$$

The interval $[\hat{p}, \check{p}]$ exists since:

$$
\check{p} - \hat{p} = \frac{v_l}{2\alpha(1-v_h)+v_l} - \frac{v_l^2}{\alpha(1-v_h)(v_l+v_h)+v_l^2} =
$$

$$
\frac{v_l \alpha\left(v_h-1\right)\left(v_l-v_h\right)}{\left(2\alpha(1-v_h)+v_l\right)\left(\alpha(1-v_h)(v_l+v_h)+v_l^2\right)} > 0 \text{ for } v_h \geq v_l, \text{ and } \alpha > 0.
$$

∎

Remark 1 *There is never an equilibrium in which only players with high know-how reveal.*

21

Proof. Inequality (3) is satisfied for:

$$
0 \; > \; p\left((v_h + v_h - v_h v_h)\frac{v_h}{v_h + v_h}V - v_h\left(1 - v_h\right)V\right)
$$
$$
+ \left(1 - p\right)\left((1 - \alpha)v_h(1 - v_l)V - v_h\left(1 - v_l\right)V\right)
$$
$$
\Leftrightarrow \quad \frac{v_h\alpha(1 - v_l)}{\frac{1}{2}v_h^2 + v_h\alpha(1 - v_l)} < p
$$

while (4) is satisfied for:

$$
0 \; < \; p\left((v_l + v_h - v_l v_h)\frac{v_l}{v_l + v_h}V - v_l\left(1 - v_h\right)V\right)
$$
$$
+ \left(1 - p\right)\left(0 - v_l\left(1 - v_l\right)V\right)
$$
$$
\Leftrightarrow \quad \frac{(v_l + v_h)\,\alpha\,(v_l - 1)}{-v_h^2 + (v_l + v_h)\,\alpha\,(v_l - 1)} > p.
$$

There is no interval for p for which both inequalities hold at the same time:

$$
\frac{(v_l + v_h)\,\alpha\,(v_l - 1)}{-v_h^2 + (v_l + v_h)\,\alpha\,(v_l - 1)} - \frac{v_h\alpha(1 - v_l)}{\frac{1}{2}v_h^2 + v_h\alpha(1 - v_l)} =
$$

$$
\frac{v_h\alpha\,(v_l - 1)\,(v_h - v_l)}{\left(-\alpha(v_l + v_h - v_l v_h) - v_h^2 + \alpha v_l^2\right)\left(-v_h - \alpha(2 + 2v_l)\right)} < 0
$$

∎

References

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books, New York.

Bacharach, M. (1993). Variable universe games. In K. Binmore, A. Kirman, and P. Tani (Eds.), *Frontiers of Game Theory*, Cambridge, MA, pp. 255 – 275. MIT Press.

Bacharach, M. and D. O. Stahl (2000). Variable-frame level-n theory. *Games and Economic Behavior 32*, 220 – 246.

Bohnet, I. and R. Croson (2004). Trust and trustworthiness. *Journal of Economic Behavior and Organization 55*(4), 443 – 445.

Bull, C. (1987). The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics 102*, 147 – 159.

Buskens, V. (2003). Trust in triads: Effects of exit, control, and learning. *Games and Economic Behavior 42*, 235–252.

Casajus, A. (2000). Focal points in framed strategic forms. *Games and Economic Behavior 32*, 263 – 291.

Coleman, J. S. (1990). Relations of trust. In *Foundations of Social Theory*, Cambridge, London, pp. 91–116. The Belknap Press of Harvard University Press.

Dasgupta, P. (1988). Trust as a commodity. In D. Gambetta (Ed.), *Trust - Making and Breaking Cooperative Relations*, New York, pp. 49–72. Basil Blackwell.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology 31*, 169–193.

Eckel, C. C. and R. K. Wilson (2004). Is trust a risky decision? *Journal of Economic Behavior and Organization 55*, 447–465.

Fehr, E. and K. M. Schmidt (2003). Theories of fairness and reciprocity - evidence and economic applications. In S. T. M. Dewatripont and L. P. Hansen (Eds.), *Advances in Economics and Econometrics, Eighth*

World Congress of the Econometric Society, Volume 1, Cambridge, pp. 208–257. Cambridge University Press.

Fukuyama, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity*. Free Press.

Gambetta, D. (1988). Can we trust trust? In D. Gambetta (Ed.), *Trust - Making and Breaking Cooperative Relations*, New York, pp. 213–238. Basil Blackwell.

Hardin, R. (2001). Conceptions and explanations of trust. In K. S. Cook (Ed.), *Trust in Society*, New York, pp. 3–39. Russell Sage Foundation.

Hart, O. and B. Holmström (1987). Chapter 3: The theory of contracts. In T. Bewley (Ed.), *Advances in Economic Theory, Fifth World Congress*, Cambridge. University Press.

Haruvy, E. and D. O. Stahl (1998). An empirical model of equilibrium selection in symmetric normal form games.

Harvey, S. J. J. (2002). The trust paradox: A survey of economic enquiries into the nature of trust and trustworthiness. *Journal of Economic Behavior and Organization 47*, 291–307.

Heimer, C. A. (2001). Solving the problem of trust. In K. S. Cook (Ed.), *Trust in Society*, New York, pp. 40–88. Russell Sage Foundation.

Janssen, M. C. W. (2001). Rationalizing focal points. *Theory and Decision 50*, 119–148.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory 27*, 245–252.

Kreps, D. and R. Wilson (1982). Reputation and imperfect information. *Journal of Economic Theory 27*, 253–279.

Kreps, D. M. (1990). Corporate culture and economic theory. In J. Alt and K. Shepsle (Eds.), *Perspectives on Positive Political Economy*, Cambridge, pp. 90–143. Cambridge University Press.

Lane, C. (2000). *Trust Within and Between Organizations*. Oxford University Press.

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

MacLeod, B. and J. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica 57*, 447–80.

Messick, D. and R. M. Kramer (2001). Trust as a form of shallow morality. In K. S. Cook (Ed.), *Trust in Society*, New York, pp. 89–118. Russell Sage Foundation.

Noteboom, B. (2002). Forms of trust. In *Trust; Forms, Foundations, Functions, Failures and Figures*, Cheltenham, Northampton, pp. 36–54. Edward Elgar Publishing.

Noteboom, B., H. Berger, and N. G. Noorderhaven (1997). Effects of trust and governance on relational risk. *The Academy of Management Journal 40*(2), 308–338.

Parkhe, A. (1998a). Building trust in international alliances. *The Journal of World Business 33*(4), 417–437.

Parkhe, A. (1998b). Understanding trust in international alliances. *The Journal of World Business 33*(3), 219–239.

Raub, W. (2003). Hostage posting as a mechanism of trust: Binding, compensation and signaling. *Institute for the Study of Cooperative Relations, Discussion Paper 208*, 1–32.

Raub, W. and G. Keren (1993). Hostages as a commitment device. *The Journal of Economic Behavior and Organization 21*, 43–67.

Sabel, C. F. (1993). Studied trust: Building new forms of cooperation in a volatile economy. In R. Swedberg (Ed.), *Explorations in Economic Sociology*, New York, pp. 104–144. Russell Sage Foundation.

Schelling, T. C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Stahl, D. O. and P. Wilson (1995). On players' models of other players: Theory and experimental evidence. *Journal of Economic Behavior and Organization 25*, 309 – 327.

Sugden, R. (1995). A theory of focal points. *Economic Journal 105*, 533 – 550.

Williamson, O. F. (1996). Calculativeness, trust, and economic organization. In O. F. Williamson (Ed.), *The Mechanisms of Governance*, New York, Oxford, pp. 250–275.

Willinger, M., C. Keser, C. Lohmann, and J.-C. Usunier (2003). A comparison of trust and reciprocity between france and germany: Experimental investigation based on the investment game. *The Journal of Economic Psychology 24*, 447–466.