Modelling relationship between trust and trustworthiness using a finite repeated trust game

Andrea Gelei (andrea.gelei@uni-corvinus.hu)
Corvinus University of Economics, Institute of Business Economics

Imre Dobos
Budapest University of Technology and Economics, Institute of Economic Sciences

Levente Dudás
Corvinus University of Economics, Institute of Business Economics

Abstract
Our hypothesis suggests that in business relationships with high level of perceived trustworthiness, the willingness to be involved in risky situations is higher than in relationships in which actors do not believe their partner is highly trustworthy. We specified the mathematical-statistical model indicating the hypothesized relationship between trust and trustworthiness. We developed a modified version of the repeated Trust Game, and carried out the experiment. One unique features of our game design are that the *ECU* amount cumulated in previous iterations could be reinvested. The second unique feature is the payout function that –supposing rational actors– facilitated a cooperative behavior.

The project is supported by the Hungarian Scientific Research Fund (OTKA), project No. K 115542.

Keywords: Trust, Trustworthiness, Repeated Trust Game

Introduction

Trust has a long-standing research history, which encompasses several methodological approaches from case studies to behavioral game designs. The research discussed in this paper applies this latter, because we think the contextual factors have crucial importance in research programs of trust, a special relational attribute. We are mainly interested in business relationship and the role trust plays in it. Our theoretical background is linked to the Transaction Cost Theory (TCT, Williamson, 1979), but also builds on the B2B marketing literature. Trust is interpreted as a governance mechanism (Jarvenpaa et al., 2000), a safeguard the partners use in their relationships against opportunism behavior. However, we distinguish between trust and trustworthiness. Trust is interpreted as the trustor's willingness to act and be actually engaged in a risky behavior with the counterpart (Mayer et al.; 1995). Trustworthiness on the other hand is a kind of stock, indicating the level of accumulated trustworthiness of the trustee at a given point in time (Barney – Hansen, 1994). In a risky situation, the level of accumulated trustworthiness (stock) is evaluated and confronted with the actual situation. In case the trustee's level of trustworthiness perceived by the trustor is sufficient, the trustor might decide to take the risk associated to the actual situation and be engaged into the risky behavior. This engagement, the actual behavior of the trustor indicates the presence of trust between partners in a given relationship. In such a case the trustworthiness becomes the instrument of governance. Trust in this conceptualization is a flow type of phenomenon reflected by the actual movements or actions of the trustor. Our overall objective to test this concept, specifically the following hypothesis: In business relationships characterized by high level of perceived trustworthiness, the willingness to be involved in risky situations is higher than in relationships in which actors do not believe their partner is highly trustworthy.

We have specified the model describing the above presented conceptualization of trust and trustworthiness and developed a game design for our experiment. We dynamized the classic Trust Game (Berg et al., 1995). We have a sample of 49 pairs. This paper presents the mathematical model and the game design and discuss the experiment and the database development process. We cannot discuss our results here; however, preliminary outputs of the experiment will be presented at the conference.

Modelling the hypothesized relationship between trust and trustworthiness – Game specification

In order to test our hypothesis, we found it necessary to 'dynamize' the game – play it as a repeated game. The repeated feature of the game allows us to observe the level of perceived trustworthiness among players throughout the multiple iterations of the game, and its effect on actual trust (action measure by the invested amounts). Dynamic trust games can be played with the players not knowing the end condition (usually: iteration count), in other cases, this end condition is disclosed at the beginning, so the players can use this information for their strategy. In the game we played, the end condition –iteration 10 times– was revealed at the beginning of the game, because we wanted the players to calculate accordingly, as it was an important aspect considering the game theory strategies. The cumulative feature of our experiment differs significantly from previously seen repeated games. It is unique that in our game the *ECU* (Experimental Currency Unit) amount cumulated in previous iterations can be reinvested. This way the *ECU* gains can grow exponentially throughout the game.

Decision variables and payout structure

In this segment we present the variables and parameters used in the modified version of the game and define the incentive structure for the players.

Decision variables of the model

- I_t^A *ECU*[1] available for player *A*, at the end of iteration *t*
- I_t^B *ECU* available for player *B*, at the end of iteration *t*

- x_t *ECU* given by player *A*, to player *B* during iteration *t*
- y_t *ECU* given by player *B*, to player *A* during iteration *t*

T denotes the final iteration count, in our case *T*=10

With the help of these variables, we can construct the following equations, considering that player *A* starts the game with 10 *ECU* and player *B* does not have any amount at the beginning:

$$I_t^A = I_{t-1}^A - x_t + y_t, \quad I_0^A = 10\,ECU$$
$$I_t^B = I_{t-1}^B + 3 \cdot x_t - y_t, \quad I_0^B = 0\,ECU$$

(t=1,2,…,T).

We set non-negative investments, *ECU* given in each round, as a condition, so: $x_t \geq 0$, $y_t \geq 0$

Assuming, that the players goal is to maximize the *ECU* collected together (adding the earnings of player *A* and *B*), then we can formulate the goal function, with the above terminology, the following way:

$$I_T^A + I_T^B \rightarrow \max .$$

The solutions of the above dynamic optimization problem are the Pareto optima. The maximum obtainable final *ECU* is $I_0 \cdot M^T$, where *M* is the investment multiplicator, the coefficient, with which we multiply the amount player *A* gives to player *B*. With the parameters we defined this is $10 \cdot 3^{10}$, when both players fully commit all *ECU* available to them, up until the last step where the transferred *ECU* multiplies by rule. This is true for all steps except for the last decision player *B* makes (10th iteration). Because of this, if the players cooperate to obtain the collectively maximum state of the game, player *B* will have the choice to make the division of the winnings among the players. Player *B* decides, how much *ECU* can player *A* get $y_T^o = 10 \cdot 3^T - \tilde{x}$

Payout structure

The function defining the payouts for the subject in the experiment was not the usual, linearly associated with the final *ECU* amount, but instead formulated in the following way

$$F_A\left(I_T^A\right) = G \cdot \delta\left(I_T^A - I_T^B\right) + \frac{I_T^A + I_T^B}{10 \cdot M^T} \cdot K ,$$

and

$$F_B\left(I_T^B\right) = G \cdot \delta\left(I_T^B - I_T^A\right) + \frac{I_T^A + I_T^B}{10 \cdot M^T} \cdot K ,$$

where

G premium for the 'winner' player with relatively more *ECU* at the end,
K premium coefficient based on the collectively cumulated *ECU* amount,

M investment multiplicator
T iteration count.

In our experiment we used the values as G=500 HUF (about 2 USD), K=1000 HUF, M=3 and T=10.

A $\delta(.)$ function - for the winner premium allocation - is defined the following way:

$$\delta(z) = \begin{cases} 0 & z < 0, \\ \frac{1}{2} & z = 0, \\ 1 & z > 0. \end{cases}$$

Strategies and stability

In our experiment, we deliberately chose the parameters to disrupt the usual game theory equilibrium. In the classic experiment, the individual strategy for both players in a stable, Nash equilibrium state is not to cooperate at all. In the following, we analyze the decisions of individual players in reverse order (starting with the last decision of player B) to see, what are their maximizing strategies.

Assuming rational decision makers with *ECU* maximizing goal, we can easily conclude that the payout function dictates $y_T = 0$ in any (I_t^B, I_t^A) cases. Therefore player A knows in the last iteration, that the goal is to max(F_A), with the assumption of not getting back anything from player B. The last decision for player A can be extracted to the following problem

$$\max_{x_T} = \left\{ G \cdot \delta\left((I_{T-1}^A - x_T) - (I_{T-1}^B + x_T \cdot 3)\right) + \frac{(I_{T-1}^A - x_T) + (I_{T-1}^B + x_T \cdot 3)}{10 \cdot M^T} \cdot K \right\}.$$

The solution of this optimization problem is dependent on the state of the game (I_t^B, I_t^A) and the optimal x_T value is different in the following cases.

1) If $\left(I_{T-1}^A \leq I_{T-1}^B\right)$ player A does not have room for optimization and have no interest in cooperation: $x_T^{\bullet} = 0$.

2) If $\left(I_{T-1}^A > I_{T-1}^B\right)$ player A should consider whether it is worth to increase the cumulatively collected *ECU* amount for the collective premium (K), or ensure the winners premium (G). This depends on the ratio of the parameters (G:K) and also on the α ratio of the cumulatively collected *ECU* amount compared to the maximally collectable amount:

   $$\alpha = \frac{I_{T-1}^A + I_{T-1}^B}{10 \cdot M^{T-1}}$$

   a. If $\alpha < 1$, meaning, in any of the previous iterations the transferred amount by any of the players was not giving the full amount available, then the maximizing strategy for player *A* is to give only so much, that *A* will still be the winner assuming y_T = 0 (ensuring the winners premium):

$$x_T^{\bullet} = (I_{T-1}^A - I_{T-1}^B)/3 - 1$$

b. If $\alpha = 1$, meaning, in all previous iterations the transferred amount by both players was the full amount available, then even with the assumption of y_T = 0 there is no x_T, that would maximize payouts by ensuring the winners premium. The optimal solution is in this case, is to transfer the full amount available $x_T^{\bullet} = (I_{T-1}^A)$, improving the collectively earned *ECU* amount, because:

$$G + (I_{T-1}^A - x_T^{\bullet} + 3 \cdot x_T^{\bullet}) \cdot K < K$$

We can conclude based on the above, that at *T-1*, assuming that the players contributed the full amount in every iteration until that point, the maximizing strategy for player *B* is to transfer the full amount, as he knows, player *A* is incentivized to give the full amount again, that *B* can keep at the end. We can retract with backward logic that all prior decisions have the same logic, to contribute the full amount if the full amount was contributed in all steps before. The first decision is not excluded from this line of thinking x_1, meaning that the classic equilibrium of the game is disrupted, and because of the design of the payout function, full cooperation is the maximizing individual strategy. This results that the players individual Nash equilibrium strategy leads to Pareto optimal state considering the overall end-state of the game

Extension of the model with the concept of trustworthiness

In the base model of the trust game the concept of trustworthiness was not included. As stated in the theoretical introduction, this concept can be interpreted as a stock variable, meaning that it can be determined at any stage. In our experiment we prompted our participating players at each iteration with questions concerning the perceived trustworthiness level to test our hypothesis on empirical data.

We extend the set of decision variables with the following:

- TW_t^A the trustworthiness of player *A*, as perceived in iteration *t* by player *B*
- TW_t^B the trustworthiness of player *B*, as perceived in iteration *t* by player *A*

Beside the above, stock type variables, we introduced an expectation function for the players, to track what amount do they expect to get in the next iteration from their partner. The two expectation functions respectively:

- $x_t^e(y_t)$ the expectation of player *B*, concerning the amount player *A* transfers at iteration *t*
- $y_t^e(x_t)$ the expectation of player *A*, concerning the amount player *B* transfers at iteration *t*

This expectation value as well as the trustworthiness value is only known to the experimenter, it is not presented to the other player. With the help of the above expectation functions we can model the changes in trustworthiness through the iterations of the game. We capture the changes in the perceived trustworthiness based on the difference between the expected and the actually transferred amount. Based on this, we define trustworthiness with the following function:

- $f^A\left(y_t^e(x_t), y_t\right)$ denotes the differential function of the perceived trustworthiness (showing how player *B* perceives) player *A* during iteration *t*. This depends on the expected amount of player *B* and the actually received amount from player *A*.
- $f^B\left(x_t^e(y_t), x_t\right)$ denotes the differential function of the perceived trustworthiness (showing how player *A* perceives) player *B* during iteration *t*. This depends on the expected amount of player *B* and the actually received amount from player *B*.

With the above terminology we formulated a discrete differential equation system of the changes in the perceived trustworthiness in the following way:

$$TW_{t+1}^A = TW_t^A + f^A\left(y_t^e(x_t), y_t\right), TW_0^A = 0,$$
$$TW_{t+1}^B = TW_t^B + f^B\left(x_t^e(y_t), x_t\right), TW_0^B = 0.$$

We assumed the initial value of the perceived trust to be minimal, hence the players were assigned in pairs randomly and secretly. Any interaction between them was only possible through the sole input of the amount invested in the computer game.

The experiment and database development

The game designed detailed above promised to be sufficient to test our hypotheses. The question remained, how to properly tailor the experiment to our game, needs and resources.

Attracting subjects for the experiment was the first problem at hand. Although, subjects from a wide range of background (preferably with appropriate business experience) would have been the ideal solution, but given our limited budget, communication channels and network, we targeted mainly the students of our university. Pervious research results indicate that students studying business and/or economics have statistically similar behaviour compared to practicing managers (Bolton et al., 2012).

Advertisements for a playful but serious game for actual monetary rewards were circulated through the official internal newsletter, which is available for all citizens of the university. Besides that, targeted promotions were made a few days in advance of each of the scheduled experiments (altogether 7 occasions), mainly during the beginning of classes for business students. These have been followed up by an email through the official learning system (Moodle). Students who were interested in participating could use an online sign-up sheet, to take empty slots at an upcoming event (signup.com). This was a flexible way to organize participation because there were fix places per experiment due to room size. Additional benefit of the signup online tool was the built-in option for participants to join a waiting list, cancel and even to swap places without extra administration from the organizer. An email reminder was sent 24 hour before each experiment to increase subject show-up rate.

The experiment had to be conducted in a controlled environment, where it is ensured, that no other type of communication took place. As the medium of the interactions was developed accordingly, an application in MS Excel environment communicating through

VBA macros and an interim data file. The program was deployed on the internal network of the university, to ensure stable connectivity. The advantage of this solution was that subjects were familiar with this kind of user interface, it allowed to log all investment decisions (ECU transfers) as well as logging trustworthiness levels and expectations.

The data recorded by the application were: identification of the pair (using a code), game turn, received amount, accumulated amount, amount to be given to the other player in this turn, perceived trustworthiness of the partner on Likert scale of 5 options and the expected amount to be received next turn in percent of the partner's accumulated amount. Besides displaying all of the above, the players could see also the current ECU accumulation for both players, updated live, each time the 'amount to be given' was modified. We anticipated, that this feature will result more conscious choices. When the player was committed to the ECU amount to pass over, they had to hit a 'save' button to convey the information to their partner. Additional advantage of the tool was, that it advised the subjects what tasks they need to complete throughout the game, using message boxes. A display at the end of the game prompted the player to ask the experimenter to approve the score then jumped to the exit survey. Upon completion of the survey, the experimenter directed the subjects to the pay-out desk.

As the venue of the experiment we had to acquire a large enough room with sufficient devices to have so many players simultaneously, that they cannot easily guess who their in-game partner is. A requirement above this was stable network connectivity. The computer lab that was available for our research project was a long room with 4 computers in a row that seats 36 people. The number of participants per experiment varied between 6 and 24 (altogether 98 participants on 7 occasions), which is rather low to ensure secret pairs, but the large enough room, along with monitors obstructing the view, proved to be able to conceal assignment of the pairs.

As stressed above, we deemed highly important that he playing partner of the subjects remain hidden. The random assignment of the pairs was defined by a label on the instruction sheets. These sheets were handed to the players by the experimenter using a simple heuristic to maximize distance between the paired subjects. After the distribution of the instruction sheets we demanded silence in the room, further clarification questions were only allowed by show-of-hand and private discussion with the experimenter. Other purposes of the instruction sheet were to communicate the aim of the game without explaining the research question and communicate the pay-out structure without explaining game theory strategies. For administrative purposes a verification section about the result of the game (ECU) was included as well. This was filled by the experimenter who remained in the computer lab throughout the experiment and was used to handle pay-outs by the second experimenter, who proceeded to the pay-out desk midway through game completion. In order to adhere with university policy, the instruction sheet prompted the players to sign a statement that they did not play this game before, they accept to use the log on their behaviour for research purposes, and also accept the cash reward based on their performance that is paid immediately after, and have no claim later on.

At the end of the game we asked the participants to complete an online 'exit' survey. The main purpose of this was to capture their perceived risk preferences and what kind of strategies they recognized, considered and played eventually. To measure risk taking and collaborate - compete aptitude we used a 7 grade Likert scale rating as well as an open question about their own strategies and how they perceived risk level through the phases of the game. We anticipated, that this self-reported data could further enrich the data analysis on their behaviour. The exit survey was also an opportunity to log basic information about the players: pairing number (for identification purposes), gender (only two options) and the major of their studies. With larger number of participants, it could be interesting to see whether the game results are mediated by the latter two aspects.

In order to assess how well the players understood the game design and reward structure, a control question was also included at the end of the survey. This simple scenario description, detailing accumulated ECU amounts for two hypothetical players, asked the subjects to determine the rewards for these players. The subject's ability to answer correctly was used to determine their understanding of the game, and their ability to play according to its logic. Arguably, this question should have popped up at the beginning the first financial decision. Measuring this at the end of the game does not show the level of understanding at the time of the first turn's decisions.

After completing the game and the survey as well, subjects were directed to the pay-out desk. Given the sensitive nature of the task we aimed to handle it separately, discreetly and as quickly as possible. The location was a desk in a hall down a stairway from the computer lab. Each player could collect their reward, one at a time. The instruction sheet had a section designated to confirm the end result that was approved by the experimenter in the computer room, and based on this, the experimenter at the pay-out desk calculated the reward. The reward formula was defined in a continuous way, but the pay-outs were rounded up to 100 HUF bits (less than 0.5 USD), for easier and swifter handling.

Conclusion

In the paper the authors showed an ambitious alteration of a classic experiment. The game design was modified in three major ways. Firstly, the reporting of the perceived trustworthiness level by the subjects about their partner, in order to contrast these values with the actual trust expressed by the invested amounts. Secondly the accumulative feature of the investments, meaning an exponential growth in the potential ECU earnings. The reason for this was the intention to create the perception of a highly risky situation, as previous research showed that the level of trust between the partners only plays a role when the perceived risk level is high. Although it can easily be shown mathematically, that non-cooperation in either of the rounds, have the same affect on the collectively gathered ECU amount, the exponential growth in numbers did create a higher risk perception according to the exit survey. High risk perception was important to press in game design and experiment setup (e.g. professional organising, enforcing silence to increase tension during gameplay), because the setting as a laboratory experiment for small cash rewards does reduce the stake compared to real life business decisions.

Thanks to the winner's premium component of the reward function, the third altered aspect, the perceived risk growth was further enhanced by the fear of opportunistic behaviour on the partner's account. However, the modification of the reward function with the component based on the collectively accumulated amount intended to make all pairs of rational decision makers cooperate throughout the game. The combination of these elements successfully created high risk perception and intense gameplay, as reported by the subject both in the exit survey and at the cash desk. This intensity contributed the end results as well, as only 6 out of the 49 pair of players followed the optimal strategy, dictated by game theory, and fully cooperated through the game. It is worth to note however, that because of the cumulated nature of the game, even small mistakes are not tolerated, meaning that the optimal strategy cease to be the cooperative one, even after one occasion of not fully committing the available amount. This feature naturally contributed to lower cooperation rates. Participants need to understand this aspect as early as turn one in our version of the game. This stresses the importance of ensuring very clear instructions and mandating participants to complete a test on game design awareness at the beginning of the gameplay.

The project is supported by the Hungarian Scientific Research Fund (OTKA), project No. K 115542.

References

Barney, J. and Hansen, M. (1994), Trustworthiness as a source of competitive advantage, *Strategic Management Journal* 15, Winter Special Issue, 175 – 190.

Berg, J., Dickhaut, J. and McCabe K. (1995), Trust, reciprocity, and social history, *Games and Economic Behavior*, 10, 122-142.

Bolton, G.E. – Ockenfels, A. – Thonemann, U.W. (2012): Managers and Students as Newsvendors, Management Science, 58 (12), 2225-2233.

Jarvenpaa, S.L., Tractinsky and N.,Vitale, M. (2000), Consumer trust in an internet store, *Information Technology and Management* 1, 45-71.

Mayer, R.C., Davis, J.H. and Schoorman, F.D. (1995), An Integrative Model of Organizational Trust, *Academy of Management Review* 20(3), 709 – 734.

Williamson, O. E. (1979), Transaction-cost economics: the governance of contractual relations, *Journal of Law and Economics*, 1979. October, Vol. 22. No 2, 233-261