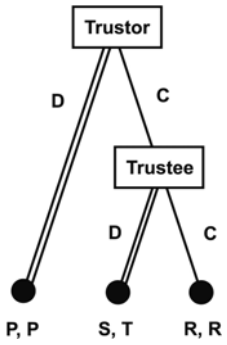
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CHAIRS OF SOCIOLOGY

Signaling Trustworthiness. Theory, Models and Experiments

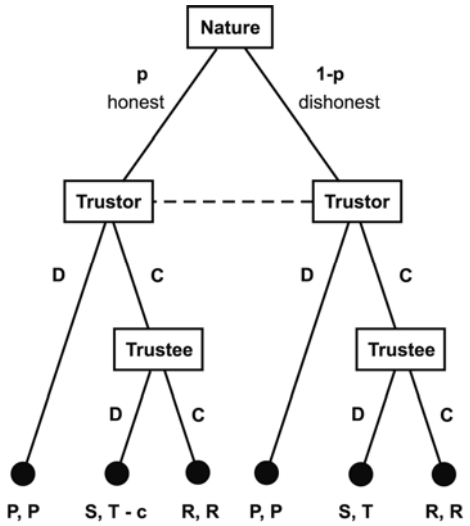
Andreas Diekmann and Wojtek Przepiorka
December 6th 2006

Trust Game



$$T > R > P > S$$

Trust game with incomplete information (Model II)



$$u_D = P$$

$$u_C = p \cdot R + (1-p) \cdot S$$

$$u_C > u_D \Rightarrow p > \frac{P-S}{R-S}$$

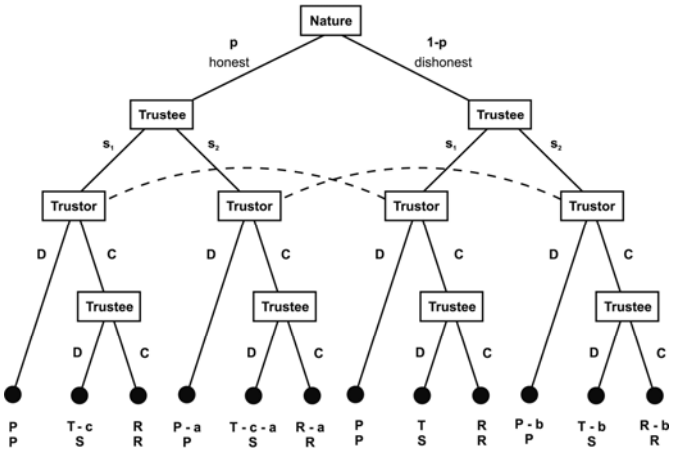
$$R > T - c$$

Costly Signals: Labels

Trust game with signals of trustworthiness (Model III)

$c > T - R$

$s_1 = no\ signaling,$
$signaling\ cost\ = 0$



$s_2 = signaling$
$signaling\ cost\ = a,\ if\ honest$
$signaling\ cost\ = b,\ if\ dishonest$

$P < R - a$

$P > T - b$

Conditions for separating equilibrium

Separating Equilibrium

if :

c > T – R

a < R – P

b > T – P

→ a < b

Equilibrium strategy:

Honest trustee signals (s2), dishonest trustee does not signal (s1).

Trustor chooses Cooperation (C) if s2, otherwise Defection (D).

Honest trustee cooperates.

Expected value

P(Θ|s2) = 1 (honest, if signal)
P(Θ*|s1) = 1 (dishonest, if no signal)

Expected utility of trustor = P(Θ|s2) P(Θ) R + P(Θ*|s1) P(Θ*) P
= pR + (1-p)P
Expected utility of honest trustee = R – a
Expected utility of dishonest trustee = P

For small proportion of dishonest actors (p is large) signaling redistributes payoffs from trustees to trustor

1. **p > (P-S)/(R-S).** Redistribution of wealth.
Expected utility of trustor **without signaling= p R + (1-p)S. With signaling: pR + (1-p)P.**
Expected utility of **honest trustee = R, with signaling R-a.**
Expected utility of **dishonest trustee = T, with signaling P.**

With signaling, both honest and dishonest trustees loose while trustor gains (1-p) (P-S).

Efficiencey gains by signaling if large proportion of dishonest actors

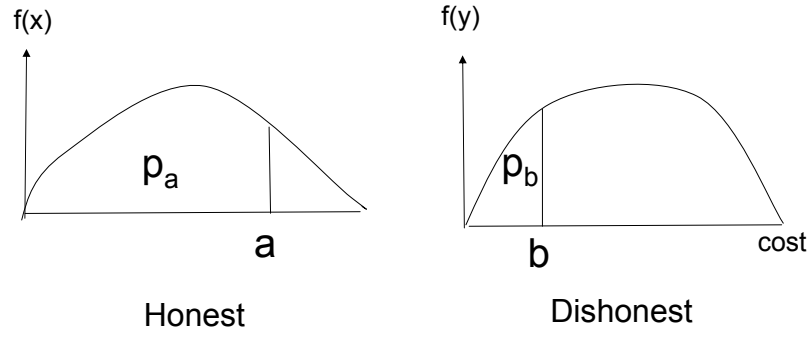
2. However, **if p < (P-S)/(R-S)** trustor and honest trustee will profit from signaling.

Without signaling, all actors receive **P**.

With signaling, **both trustor (pR+(1-p)P) and honest trustee (R-a)** win, dishonest trustees receive same payoff **P** as before. Signaling mechanism transforms the situation such that a Pareto improvement is attainable.

H: Probably, signaling mechanisms will develop if p is small, i.e. there is a large proportion of dishonest actors.

Distribution of signaling cost for honest and dishonest actors (Model IV)



Extension of he model: Distribution of signaling costs (model IV)

Strong assumption: **All honest actors have low signaling costs a, all dishonest actors have high costs b.**

Assume a distribution of signaling costs x (y) for honest (dishonest) actors. a (b) is the threshold and p_a (p_b) the area "left" of the threshold. The probability of signaling for honest (dishonest) actors is p_a (p_b) with $p_a > p_b$.

Compute expected utility of the trustor by Baysean updating:

$P(\Theta|signal) = (p_a\, p)/(p_a\, p + p_b(1-p))$

$P(\Theta^*|signal) = (p_b\,(1-p)/(p_a\ p + p_b(1-p))$

$E(Trustor) = p_a pR + p_b(1-p)S + (1- p_a)pP + (1-p_b)(1-p)P > P$

Trustor chooses C if there is a signal, otherwise D on condition that expected utility E(Trustor) > P

Condition for semi-separating equilibrium

$$(p_a/p_b)(p/(1-p) > (P-S)/(R-P)$$

or

$$(R-P)p_a p > (P-S)p_b(1-p)$$

Signaling if

1. Large efficiency gain (R-P).
2. Large probability for signaling costs below threshold for honest actors relative to dishonest actors (p_a/p_b).
3. The larger the ratio of honest and dishonest actors p/(1-p).

Trust game with incomplete information and signaling provides a theoretical framework

(1)Generate the structure of the strategic situation in a lab and study the behavior (signaling, response to signals, acceptance of an institution of signaling)

(2)Identify similar situations in the field and study the behavior.

(3)Intervene in the field and study the behavior (fieldexperiment)

Some References

Model I, the original trust game (TG), was proposed by Dasgupta (1988) and Kreps (1990). Also, Dasgupta (1988) outlined the TG with two types of actors (honest and dishonest) and incomplete information (Model II). Furthermore, he computed a threshold according to Coleman‘s formula using a numerical example. The TG with two types and incomplete information is further discussed in papers by Camerer and Weigelt (1988), Voss (…), Buskens (1999), Bacharach and Gambetta (2001), and Raub (2004). For a derivation of the Coleman threshold from a TG with incomplete information see Dasgupta (1988), Voss (…), and Raub (2004). The signaling model III is a special case of the hostage-model presented in Raub (2004). Bacharach and Gambetta discuss the existence of a semi-separating equilibrium in more general terms, focussing on mimic behaviour. They do not outline the extensive game explicitly. However, the assumptions are similar to model IV. (In contrast to model IV they assume that honest people will always signal with probability one.)

Field Experiment

2 x 2 experiment

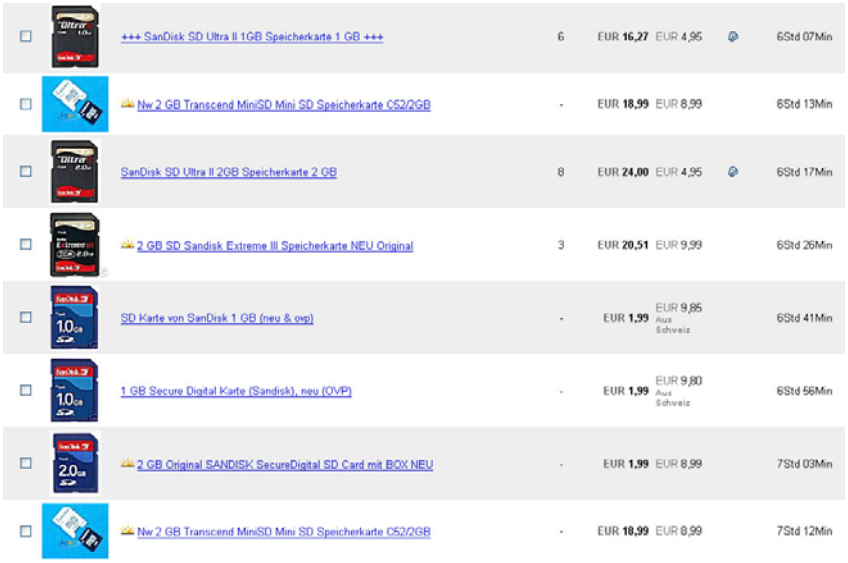
eBay auction with 84 sellers of 1 GB memory cards

(1 pair per day, 42 days)

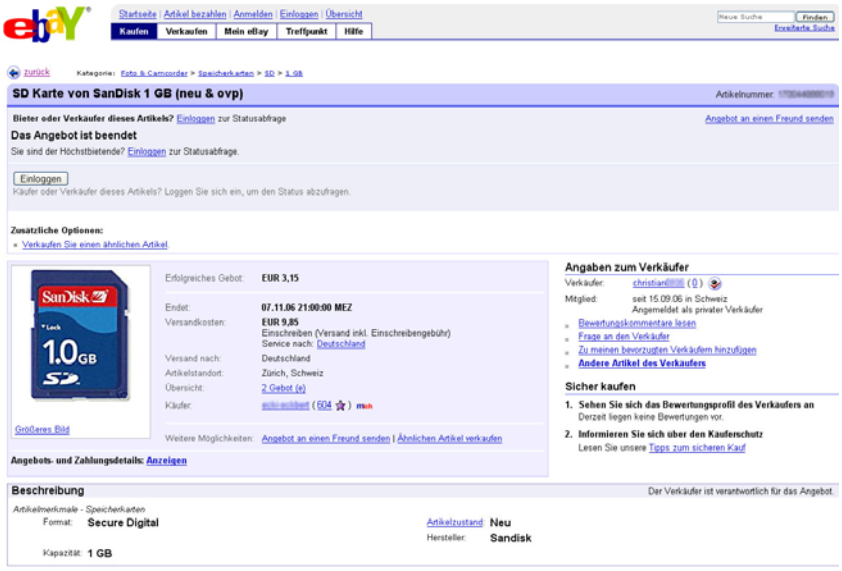
Sellers start with zero reputation (assumption: signals are more important if seller has no reputation)

	German name	Turkish name
Seller ID	**21**	**21**
No seller ID	**21**	**21**

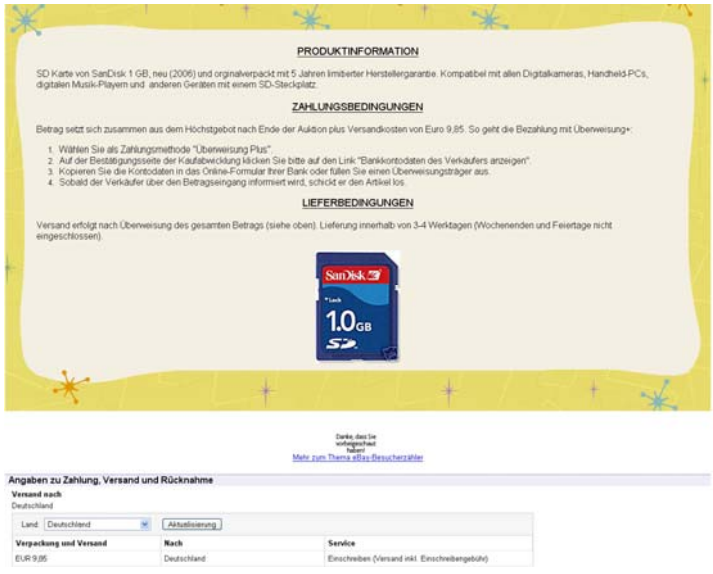
Listing

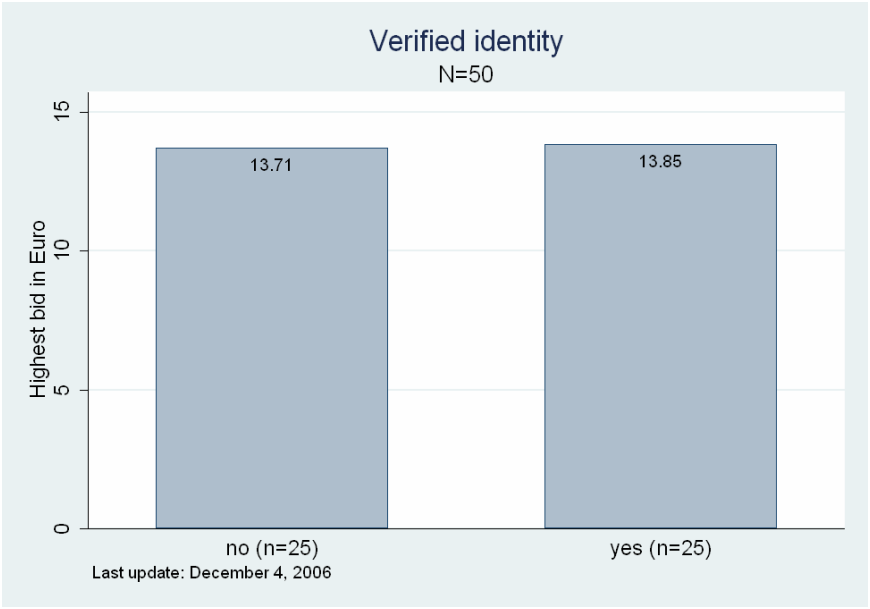


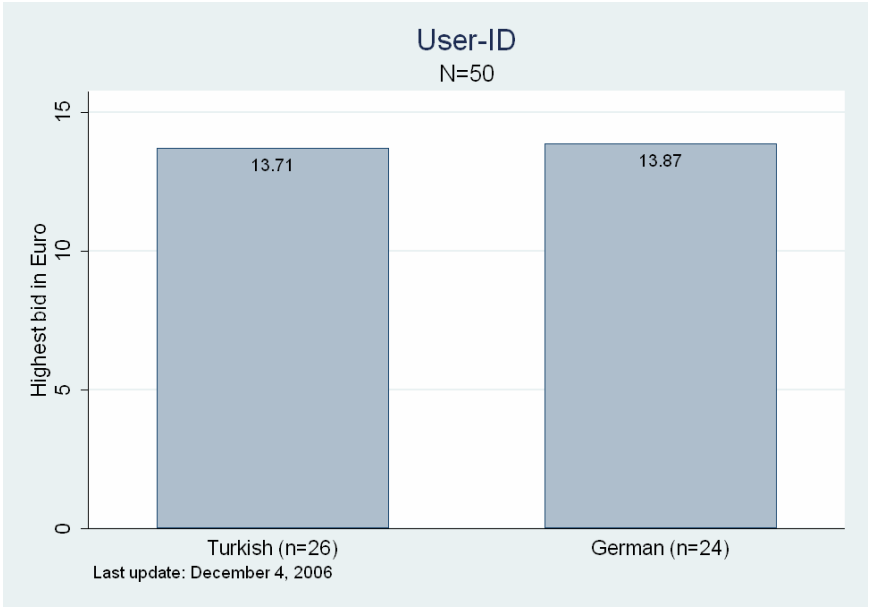
Item, Part 1



Item, Part 2




Verified identity
N=50
Highest bid in Euro
0
5
10
15
13.71
13.85
no (n=25)
yes (n=25)
Last update: December 4, 2006


User-ID
N=50
Highest bid in Euro
0
5
10
15
13.71
13.87
Turkish (n=26)
German (n=24)
Last update: December 4, 2006

Field Experiment (N=50)

2 x 2 experiment

eBay auction with 84 sellers of 1 GB memory cards

(1 pair per day, 42 days)

Sellers start with zero reputation (assumption: signals are more important if seller has no reputation)

	German name	Turkish name
Seller ID	**13.64** (13.25) 1.61 n=12	**14.05** (11.94) 3.26 n=13
No seller ID	**14.09** (13.94) 2.29 n=12	**13.37** (12.40) 1.76 n=13

Fixed-effects models

Variables	Model 1	Model 2
Identity (0/1)	**1.340**	1.125
Verified identity (1=yes)	(2.46)	(1.97)
Name (0/1)	-0.214	-0.272
User name (1=German)	(-0.39)	(-0.50)
First (0/1)	-	0.511
First in listing (1=yes)		(1.14)
constant	**13.217**	**13.096**
	(31.75)	(30.70)
R^2 (within)	0.225	0.273
R^2 (between)	0.077	0.104
R^2 (overall)	0.001	0.003
rho	0.702	0.706
N	50	50

Fixed-effects regression, dependent variable Highest bid (in Euro). t-values in parenthesis.

Effect of signal (reputation) on price

(Diekmann and Wyder 2002)

	Modell 1 mit absoluter Anzahl Bewertungen	Modell 2 mit logarithmierter Anzahl Bewertungen	Modell 3 mit absoluter Anzahl Bewertungen und Heckman-Korrektur
Reputation (Anzahl Bewertungen)	0,671** (3,19)	10,755** (2,62)	0,720*** (3,46)
Mindestpreis	0,055 (1,19)	0,075 (1,64)	–0,045 (–0,89)
Versandkosten	–2,549* (–2,48)	–3,111** (–2,82)	–2,048* (–2,03)
Dauer der Auktion	–0,200 (–0,16)	–0,569 (–0,45)	–0,080 (–0,067)
Mindesterhöhung	3,313*** (4,29)	3,635*** (4,64)	2,923*** (4,05)
Anzahl der Gebote	1,278 (1,89)	1,597* (2,39)	0,685 (1,02)
Konstante	505,79*** (16,88)	496,50*** (16,51)	529,45*** (17,98)
Adj. R^2	0,261	0,237	0,408
Lambda	–	–	72,042*** (4,95)
N	99	99	99

OLS-Regression mit der abhängigen Variable Verkaufspreis (netto, ohne Versandkosten). t-Werte in Klammern. Signifikant für: $\alpha = 0{,}05$ (*), $\alpha = 0{,}01$ (**), $\alpha = 0{,}001$ (***) bei zweiseitigem Test. Modell mit logarithmierter Reputation: Reputation = ln(Anzahl Bewertungen +1). Heckman-Korrektur: Probit-Schätzung des Verkaufserfolgs mit Prädiktoren „Reputation", „Mindestpreis" und „Versandkosten". Die geschätzte Korrelation zwischen dem Fehlerterm der Regression und der Probitgleichung ist größer als eins und wurde durch eins ersetzt.

Anderhub, Vital, Dirk Engelmann, and Werner Güth, 2002: An experimental study of the repeated trust game with incomplete information. Journal of Economic Behavior and Organization 48:197-216.

Bacharach, Michael, and Diego Gambetta, 2001: Trust in Signs. S. 148-184 in: Karen S. Cook (Hg.): Trust in Society, New York: Russell Sage Foundation.

Bolle, Friedel, and Jessica Kaehler, 2005: The Conditional Efficiency of Signals for Trustworthiness. Workingpaper: Europa Universität Viadrina Frankfurt (Oder).

Buskens, Vincent, 1999: Social Networks and Trust, Amsterdam: Thela Thesis.

Camerer, Colin and Keith Weigelt, 1988: Experimental Test of a Sequential Equilibrium Reputation Model. Econometrica 56: 1–36.

Coleman, James S., 1990: Foundations of Social Theory. Cambridge (MA): The Belknap Press of Harvard University Press.

Dasgupta, Partha, 1988: Trust as a Commodity. S. 49-72 in: Diego Gambetta (Hg.): Trust: Making and Breaking Co-operative Relations, Oxford: Basil Blackwell.

Kreps, David M., 1990: Corporate Culture and Economic Theory. S. 90-143 in: James E. Alt und Kenneth A. Shepsle (Hg.): Perspectives on Positive Political Economy, Cambridge, Mass.: Cambridge University Press.

Raub, Werner, 2004: Hostage Posting as a Mechanism of Trust: Binding, Compensation, and Signaling. Rationality and Society 16: 319-365.

Snijders, Chris, 1996: Trust and Commitments. Amsterdam: Thesis Publishers.