Will Banks Promote Trade? Equilibrium Selection for the Trust Game with Banks*

Werner Güth† Peter Ockenfels‡

July, 1998
revised version: December, 1998

Abstract

The Trust Game describes a situation where mutually beneficial trade is endangered by opportunistic exploitation. In the Trust Game with Banks this dilemma can be avoided by banks guaranteeing that sellers will be paid. This outcome is, however, not the only possible solution. Bank interference as an equilibrium outcome can coexist with another equilibrium according to which banks are not used at all. By applying the theory of equilibrium selection it is analysed which of the two competing outcomes should be expected, i.e. whether or not banks can indeed promote trade.

*The authors acknowledge the helpful comments by Martin Schweizer. Support from the Deutsche Forschungsgemeinschaft (SFB 373, Quantifikation und Simulation ökonomischer Prozesse) is gratefully acknowledged.

†Humboldt-University of Berlin, Department of Economics, Institute for Economic Theory III, Spandauer Str. 1, D - 10178 Berlin, Germany

‡J. W. Goethe-University, Department of Economics (02), Institute for Economic Theory, PF 11 19 32 (81), D - 60054 Frankfurt/Main, Germany

1. Introduction

Already before business banks were assuming a major role in the last century, it was predicted that they will help to overcome social dilemma situations due to opportunistic exploitation and thus promote cooperation which will further weaken incentives for exploitation (Hildebrand, 1864). An attempt to rigorously analyse this claim has been made by Güth (1998), based on a proposal by Schils (1998) and on previous studies of the trust game (Güth and Kliemt, 1994).

Here we will rely on what has been called the further reduced game of trust with banks (see Figure II.4 of Güth, 1998) to which we now refer simply as the **Trust Game with Banks**. This game (see Figure I) results from a more complex game by eliminating dominated strategies. It will be interpreted here without referring to the more complex situation. The upper payoff is that of the seller, the lower the one of the buyer. Compared to Güth (1998) the way of describing payoffs is further simplified.

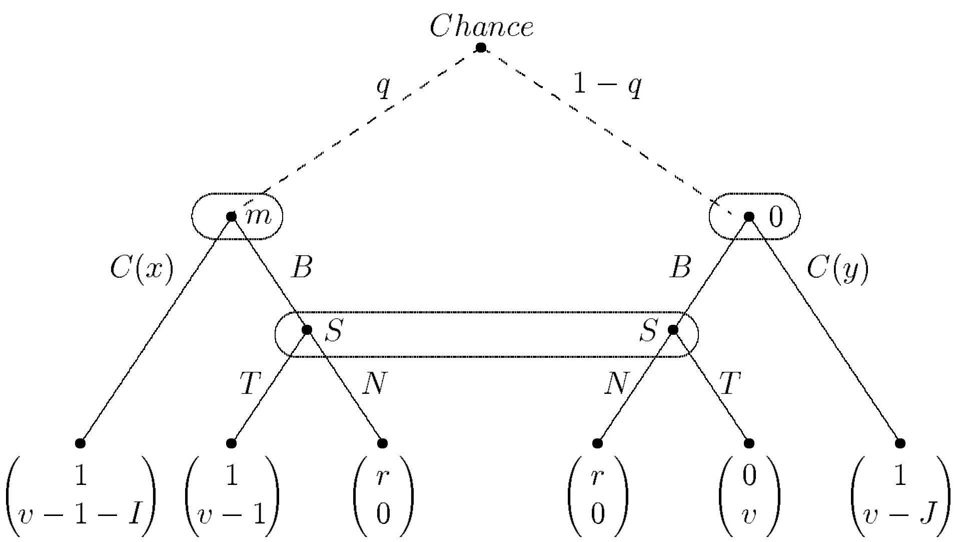

Figure I: The Trust Game with Banks

The game starts with a fictitious (therefore the dashed lines) chance move selecting the buyer's type m, who can be trusted, or 0 who would exploit the seller if possible. Knowing his own type m or 0 the buyer can either apply (the move

$C(x)$ by m and $C(y)$ by 0) for a price guarantee or not (the move B by m, respectively 0). The banks themselves are not explicitly modelled as strategic players, but only captured by the implications of their decision behavior. More specifically, it is assumed that the application by m is granted with probability x whereas the one by 0 is accepted with probability y where

(I.1) $\quad 1 > x > y > 0.$

The interpretation of assumption (I.1) is that banks know more, respectively obtain better information so that they can at least stochastically separate types of customers where a usual seller is unable to do so.

If the application is accepted, the game is over: The seller receives for sure the standardized sales price of 1. In case of m the payment is made by the buyer whose evaluation for the sales object is v. The cost I is what m has to pay to the bank for guaranteeing that the price will be paid. In case of the buyer type 0 the seller receives the price of 1 from the bank. The payoff of type 0 is his value v for the sales object minus the cost J of credit use and a possibly positive security deposit. Our analysis is based on the assumption

(I.2) $\quad v > 1 + I,\ I > 0,\ 1 > J > 0.$

The first inequality means that, of course, $v > 1$ holds (otherwise there exists no incentive for trade) and that the positive costs I of bank interference are not prohibitively high either. The other assumption simply states that it is indeed advantageous (for an 0-type customer) not to pay the price, but to fool the bank.

If the payment is not guaranteed by a bank or if the customer does not apply for such a guarantee at all (it is assumed by Figure I that both events yield the same consequences), the seller S must decide whether to trust (the move T) or not (the move N). In case of T and the m-type, the price will be paid whereas in case of the 0-type customer the move T leads to exploitation (the customer obtains v without paying the price of 1). In case of N no trade takes place, i.e.

the customer earns nothing whereas the seller S keeps the object of sale which he evaluates by r with

$$(I.3) \qquad 1 > r > 0.$$

Clearly, $1 > r$ is needed for providing an incentive to trade whereas $r > 0$ expresses that the object of trade is not useless for the seller S.

The two pure strategy equilibria (see Güth, 1998)

$$(I.4) \qquad U = (C(x), C(y), N) \text{ and } V = (B, B, T)$$

coexist for

$$(I.5) \qquad r < q < \frac{(1-y)r}{1-yr-(1-r)x}.$$

Because of (I.1) and (I.3) condition (I.5) defines a non-empty interval. To exclude that U implies an expected loss of the bank (which would question the institution of such a bank), one has to restrict the range of q further by

$$(I.6) \qquad q \geq \frac{y(1-J)}{xI+y(1-J)}.$$

The condition that the two restrictions (I.5) and (I.6) together define a non-empty interval for the probability $q \in (0,1)$ is

$$(I.7) \qquad x > 1 - \frac{1-y}{y} \cdot \frac{r}{1-r}(xI - yJ).$$

In view of conditions (I.1) and (I.3) the right hand side of (I.7) is smaller than 1 for $I \geq J$. In the following we simply assume

$$(I.8) \qquad r \geq \frac{y(1-J)}{xI+y(1-J)}$$

to make sure that the equilibrium U does not imply an expected bank loss and that the institution of banks is sustainable. Thus we solve the model satisfying conditions (I.1), (I.2), (I.3), (I.5), and (I.8). Whenever we speak of the Trust Game with Banks in the following, all these requirements are assumed.

Even with these restrictions a complete (unique) solution of all games would require a very complicated case distinction which would go beyond what can be represented in one article. We will partly restrict ourselves to just demonstrating the method of selecting an equilibrium. More specifically, all parameters except the crucial screening probabilities x and y of bank interference will then be numerically specified in order to reduce the complexity of the case distinction.

2. The reduced game for comparing U and V

All three players S, m, and 0 in Figure I choose different strategies according to U and V and thus are active players in the reduced game for comparing U and V. Since all players have only two pure strategies, this means that the reduced game for comparing U and V is the original Trust Game with Bank. In the following we will determine the conditions for U risk dominating V or vice versa (see Harsanyi and Selten, 1988).

3. The bicentric prior

Formally, the bicentric prior is a mixed strategy vector

$$\text{(III.1)} \qquad p = (p_m, p_0, p_S)$$

which can be derived as follows: For $i = m, 0, S$ let

$$\text{(III.2)} \qquad zU_{-i} + (1 - z)\, V_{-i}$$

denote the correlated behavior of i's coplayers $j \neq i$ according to which they play as prescribed by U with probability z and as prescribed by V with probability

$1-z$ where z is uniformly distributed on $[0,1]$. The bicentric prior $p_i(s_i)$ for strategy s_i of player i is then the probability for s_i being a best reply to this expectation concerning i's coplayers' behavior.

For $i=m$ the payoff for $s_i=B$ exceeds the one for $s_i=C=C(x)$ if

(III.3) $$(1-z)(v-1) > x(v-1-I) + (1-x)(1-z)(v-1)$$

or

(III.3') $$z < \frac{I}{v-1}.$$

From (III.3') follows

(III.4) $$p_m(B) = \frac{I}{v-1}, p_m(C) = \frac{v-1-I}{v-1}.$$

Similarly, one obtains for $i=0$ and $s_0=B$ or $s_0=C=C(y)$ the result

(III.5) $$p_0(B) = \frac{J}{v}, p_0(C) = \frac{v-J}{v}.$$

Finally, for $i=S$ the payoff of T is greater than for N if

(III.6) $$q + (1-q)zy > q[zx + (1-zx)r] + (1-q)[zy + (1-zy)r]$$

or

(III.6') $$z < \frac{q-r}{qx(1-r)-(1-q)yr}$$

yielding

(III.7) $$p_S(T) = \frac{q-r}{qx(1-r)-(1-q)yr}, p_S(N) = \frac{qx(1-r)-(1-q)yr-q+r}{qx(1-r)-(1-q)yr}.$$

4. Best replies to bicentric priors

If player i's coplayers $j \neq i$ rely on their bicentric prior strategy p_j, player i's best reply to p is the pure strategy yielding the higher payoff expectation. For $i = m$ one has

$$\text{(IV.1)} \qquad u_m(C) = x(v-1-I) + (1-x)\,p_S(T)(v-1)$$

and

$$\text{(IV.2)} \qquad u_m(B) = p_S(T)(v-1).$$

The condition $u_m(C) > u_m(B)$ is equivalent to

$$\text{(IV.3)} \qquad \gamma_m := \tfrac{v-1-I}{v-1} > p_S(T).$$

For $i = 0$ the strategy C is better than B if $u_0(C) > u_0(B)$ where

$$\text{(IV.4)} \qquad u_0(C) = y(v-J) + (1-y)\,p_S(T)\,v$$

and

$$\text{(IV.5)} \qquad u_0(B) = p_S(T)\,v.$$

Thus

$$\text{(IV.6)} \qquad \gamma_0 := \tfrac{v-J}{v} > p_S(T)$$

is the condition for $C = C(y)$ being the best reply to the bicentric prior p.

Finally, one obtains

$$\text{(IV.7)} \qquad u_S(T) = q + (1-q)\,p_0(C)\,y$$

and

$$\text{(IV.8)} \qquad u_S(N) = q\left[p_m(C)\,x + (1 - p_m(C)\,x)\,r\right] + (1 - q)\left[p_0(C)\,y + (1 - p_0(C)\,y)\,r\right]$$

so that N is the best reply to p if

$$\text{(IV.9)} \qquad q\,(1 - p_m(C)\,x)\,(1 - r) < (1 - q)\,(1 - p_0(C)\,y)\,r.$$

If the vector of best replies to p is U, then U risk dominates V and vice versa. Thus we have shown

Proposition 1: The solution of the Trust Game with Bank is

(i) U for (IV.3), (IV.6), and (IV.9)

and

(ii) V if all three inequalities in (i) are reversed.

It will be graphically illustrated below (see Figure VIII) that both cases (i) and (ii) define generic parameter regions. Thus Proposition 1 proves already that two very different results are possible: According to U banks are actively engaged in overcoming the trust problem. How likely bank intervention is actually to occur depends, of course, on x and y, respectively. At least if x is large and y small, bank intervention solves the problem nicely by excluding the untrustworthy mostly from trade (see Figure VIII). But also the contrary is possible, namely that banks exist, but are not asked for help. This typically results when x and y are relatively close (see Figure VIII). In such a case banks are not much better than sellers in singling out unreliable customers. It thus comes as no surprise that they are not asked for help.

5. The linear tracing procedure

For all other best reply vectors to p than those covered by (i) or (ii) in Proposition 1 the definition of risk dominance involves the (linear) tracing procedure (see Harsanyi, 1975, and Harsanyi and Selten, 1988). The tracing procedure relies on a one parameter-family of games according to which each player i expects his coplayers $j \neq i$ to play according to p with probability $1-t$ and according to their actual strategy constellation $s_{-i} = (s_j)_{j \neq i}$ with the complimentary probability t where the parameter t must satisfy $0 \leq t \leq 1$.

For every value t let $s(t)$ denote an equilibrium of the game defined by this value of t and the behavioral expectations as depending on t. Clearly, for $t = 0$ the equilibrium $s(0)$ must be a vector of best replies to p. And clearly for $t = 1$ the equilibrium $s(1)$ is an equilibrium of the original game. The tracing procedure describes a continuous path starting from $s(0)$, i.e. a vector of best replies to p, to $s(1)$, i.e. an equilibrium of the original game. If $s(1) = U$, then U risk dominates V whereas V risk dominates U if $s(1) = V$.

Deriving $s(1)$ for all possible Trust Games with Banks would result in a very complex case distinction. There exist 6 other best reply vectors to p than those covered by Proposition 1 whose tracing paths require further case distinctions of case distinctions, ... We therefore restrict ourselves to one specific numerical constellation, namely

(V.1) $\quad v = 2,\ I = .75,\ J = .5,\ r = .5,\ q = .75,$

leaving just the parameters x and y unspecified. Since in our experimental study we will focus on a constellation which differs from (V.I) only in the components r and q (see (VIII.1) below), we wanted to rely here on a constellation which illustrates the algorithmic aspects in a simple, but typical way. Notice that for (V.1) also inequality (I.8) is satisfied, i.e. the equilibrium U does not imply an expected bank loss. From (V.1) follows

(V.2) $\gamma_m = p_m(C) = .25;\ \gamma_0 = p_0(C) = .75;\ p_S(T) = \frac{2}{3x-y}.$

so that (IV.6) becomes

(V.3) $3x - y > \frac{8}{3}$

whereas conditions (IV.3) and (IV.9) are always violated. Thus there are only two possible best reply vectors to p, namely V in case of

(V.4) $1 > x > y > max\left\{3x - \frac{8}{3}, 0\right\}$

as captured by part (ii) of Proposition 1, and the vector $(B, C(y), T)$ for

(V.5) $1 > x > 3x - \frac{8}{3} > y > 0.$

We thus have to apply the tracing procedure only for the case (V.5).

As illustrated in Figure V condition (V.5) delineates a small parameter region, namely the right lower region with the extreme corner $(x = 1, y = 0)$ representing a banking system with perfect screening ability.

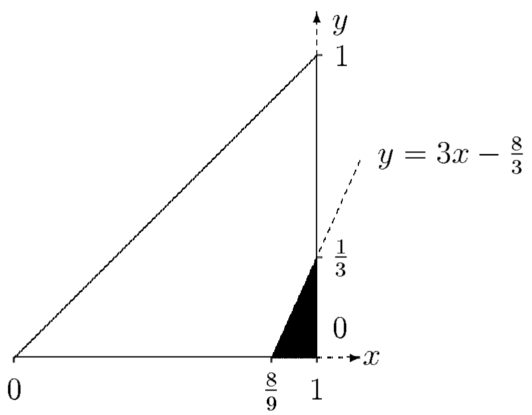

Figure V: The parameter region satisfying inequality (V.5) for the numerical specification (V.1), namely the shaded area

6. The first switching in case (V.5)

Since all best replies are the only best replies to p, the tracing starting with $s(0) = (B, C(y), T)$ will remain constant initially, i.e. $s(t) = s(0)$ for some interval $0 \leq t \leq t^1$. We refer to t^1 as the first switching time. If at $t = t^1$ customer type 0 would switch over completely to B, then V would be reached and $s(t) = V$ would prevail for all t with $t^1 < t \leq 1$, i.e. V would risk dominate U.

We first show that S and customer type m can never switch first and derive then the value t at which customer type 0 switches, i.e. the first switching time t^1. For $s(t) = s(0) = (B, C(y), T)$ an increase of t will simply increase the seller S's incentive for trust (the move T) since according to $(B, C(x), T)$ his trust is more and more likely to be rewarded. Formally, for $s(t) = (B, C(y), T)$ player S expects at time t

$$\text{(VI.1)} \qquad \begin{aligned} U_S(T) &= t\left[q + (1-q)\,y\right] + (1-t)\,u_S(T) \\ &= \tfrac{3}{4} + \tfrac{3}{16}y + \tfrac{y}{16}t \end{aligned}$$

because of (IV.7) when using T and

$$\text{(VI.2)} \qquad \begin{aligned} U_S(N) &= t\left[qr + (1-q)\,y + (1-q)(1-y)\,r\right] + (1-t)\,u_S(N) \\ &= \tfrac{3}{32}(x+y) + \tfrac{1}{2} - \left(\tfrac{3}{8}x - \tfrac{1}{8}y\right)\tfrac{t}{4} \end{aligned}$$

because of (IV.8) when using N. Clearly, $U_S(T)$ increases with t whereas $U_S(N)$ decreases. Since further by assumption (V.5) one has $u_S(T) > u_S(N)$ this proves that S will never want to switch from T to N as long as $s(t) = s(0) = (B, C(y), T)$ holds.

Customer type m achieves his highest possible payoff by using $s_m = B$ when the seller relies on T. Thus increasing t strengthens his incentive for using B as long as $s(t) = s(0) = (B, C(y), T)$ holds. Formally one obtains

$$\text{(VI.3)} \qquad \begin{aligned} U_m(B) &= t\,(v-1) + (1-t)\,u_m(B) \\ &= \tfrac{2}{3x-y} + \tfrac{3x-y-2}{3x-y}t \end{aligned}$$

because of (IV.2) and

$$\text{(VI.4)} \qquad \begin{aligned} U_m(C) &= t\left[x(v-1-I)+(1-x)(v-1)\right]+(1-t)\,u_m(C) \\ &= \tfrac{x}{4}+(1-x)\tfrac{2}{3x-y}+(1-x)\tfrac{3x-y-2}{3x-y}t \end{aligned}$$

because of (IV.1). Since we are in case (V.5), both $U_m(B)$ and $U_m(C)$, increase with t. Clearly, because of $1 > x$ the derivative of $U_m(B)$ with respect to t is larger than the one for $U_m(C)$. Furthermore, by assumption (V.5) one has $u_m(B) > u_m(C)$, i.e. also customer type m will never want to switch from B to $C(x)$ as long as $s(t) = s(0) = (B, C(y), T)$ holds.

Although this implies that consumer type 0 is the first to switch, we want to determine t^1 explicitly. Relying on the analogous notation as above one obtains

$$\text{(VI.5)} \qquad \begin{aligned} U_0(C) &= t\left[y(v-J)+(1-y)v\right]+(1-t)\,u_0(C) \\ &= \tfrac{3}{2}y+(1-y)\tfrac{4}{3x-y}+(1-y)\tfrac{6x-2y-4}{3x-y}t \end{aligned}$$

because of (IV.4) and

$$\text{(VI.6)} \qquad \begin{aligned} U_0(B) &= tv+(1-t)\,u_0(B) \\ &= \tfrac{4}{3x-y}+\tfrac{6x-2y-4}{3x-y}t \end{aligned}$$

because of (IV.5). Both payoff expectations increase with t. Although for $t=0$ the payoff expectation $U_0(C)$ is larger than $U_0(B)$ due to $u_0(C) > u_0(B)$ for condition (V.5), the derivative of $U_0(C)$ with respect to t is clearly smaller than the one of $U_0(B)$. The first switching time t^1 is the value of t for which $U_0(C) = U_0(B)$ holds, i.e.

$$\text{(VI.7)} \qquad t^1 = \tfrac{8-9x+3y}{8-12x+4y}$$

which in case of (V.5) satisfies $0 < t^1 < 1$. If one substitutes the boundary equation $y = 3x - \frac{8}{3}$ of case (V.5) into equation (VI.7) one obtains $t^1 = 0$.

7. The continuation after switching in case (V.5)

For $t = t^1$ customer type 0 is indifferent between $s_0 = B$ and $s_0 = C = C(y)$ when $s_m = B$ and $s_S = T$. But not every constellation $\left(B, q_0^{t^1}, T\right)$ with q_0^{t-1} being a mixed strategy of customer type 0 needs to be an equilibrium of the game defined by the parameter value $t = t^1$. More specifically, the question is whether for $t = t^1$ seller S still wants to use T when 0 switches completely from $C = C(y)$ to B. Let q denote the vector (B, q_0, T) and denote also by q_0 the probability by which 0 relies on $C = C(y)$. We investigate when T is still a best reply against q for $t = t^1$. Similarly, as in the previous section one derives

$$\text{(VII.1)} \qquad \begin{aligned} U_S(T) &= t^1 \left[q + (1-q)\, q_0 y\right] + \left(1 - t^1\right) u_S(T) \\ &= \tfrac{3}{4} + \tfrac{3}{16} y \left(1 - t^1\right) + \tfrac{y}{4} t^1 q_0 \end{aligned}$$

because of (IV.7) for seller S using T and

$$\text{(VII.2)} \qquad \begin{aligned} U_S(N) &= t^1 \left[qr + (1-q)\left(q_0 y + (1 - q_0 y)\, r\right)\right] + \left(1 - t^1\right) u_S(N) \\ &= \tfrac{t^1}{2}\left(1 + \tfrac{1}{4} y q_0\right) + \tfrac{1-t^1}{4}\left(\tfrac{3}{8}(x+y) + 2\right) \end{aligned}$$

because of (IV.8). From our previous analysis we know that $U_S(T) > U_S(N)$ for $q_0 = 1$ since S will never switch first. But when switching from $C = C(y)$ to B the mixed strategy parameter q_0 moves from $q_0 = 1$ to $q_0 = 0$. Thus the complete switch requires that the condition

$$\text{(VII.3)} \qquad \tfrac{3}{4} + \tfrac{3}{16} y \left(1 - t^1\right) > \tfrac{t^1}{2} + \tfrac{1-t^1}{4}\left(\tfrac{3}{8}(x+y) + 2\right)$$

is satisfied. In this case $s(1) = V$ holds which proves

Proposition 2: Assume for the Trust Game with Banks the parameter constellation (V.1) and (V.5) and condition (VII.3) where t^1 is defined by (VI.7). Then V risk dominates U.

According to the best reply vector $s(0) = (B, C(y), T)$ to the bicentric prior p the trustworthy buyer m and the seller S always engage in mutually beneficial trade whereas in case of the 0-buyer type this only happens with probability $1-y$. Substituting naive Bayesianism (in the sense of $s(0)$) by strategic thinking, represented by an increase of t, induces then the 0-type buyer to switch from $C(y)$ to B, i.e. to mimic the m-type's behavior. For the parameter region, delineated by the assumptions of Proposition 2, this does suffice to let seller S shy away from trusting. Thus banks may not be actively involved even when the best reply vector $s(0)$ to p does not exclude this possibility, i.e. in case of $s(0) \neq U$.

If condition (VII.3) is reversed, there would exist a value q_0^* for which $U_S(T)$ equals $U_S(N)$, i.e.

$$\text{(VII.4)} \qquad q_0^* = \left[\tfrac{t^1}{2} + \tfrac{1-t^1}{4}\left(\tfrac{3}{8}(x+y)+2\right) - \tfrac{3}{4} - \tfrac{3}{16}\left(1-t^1\right)y\right] 4/yt^1.$$

The tracing path for $t = t^1$ would then move from $(B, q_0 = 1, T)$ to (B, q_0^*, T) only, i.e. customer type 0 switches, but not completely. The continuation would require an initial decrease of t, i.e. values of $t < t^1$, before t can increase again and finally reach 1. We can, however, prove that under (V.1) and (V.5) this is impossible.

Proposition 3: Conditions (V.1) and (V.5) for the Trust Game with Banks imply inequality (VII.3).

Substitution of t^1, as defined by (VI.7), into (VII.4) yields

$$\text{(VII.5)} \qquad q_0^* = \tfrac{64-96x+32y-12xy+9x^2+3y^2}{4(9x-3y-8)y}$$

Due to (V.4) the denominator of the right hand side of (VII.4) is positive. The numerator, which can be written as a function

$$\text{(VII.6)} \qquad f(x,y) = 32(2+y-3x) + 3\left(3x^2+y^2-4xy\right)$$

of the two screening parameters x and y, is negative in the whole range (V.4).

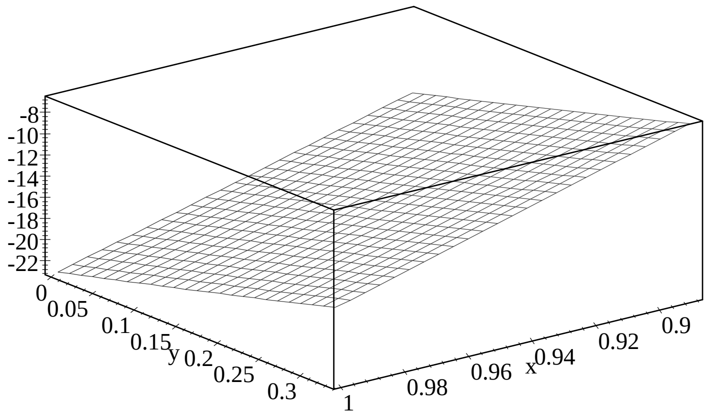

Figure VII: The values of $f(x,y)$ in the rectangle $\frac{8}{9} \leq x \leq 1$ and $0 \leq y \leq \frac{1}{3}$ containing region (V.5)

Due to the continuity of $f(x,y)$ and because of

$$\text{(VII.7)} \qquad f_x(x,y) = -96 + 3(6x - 4y) < 0$$

and

$$\text{(VII.8)} \qquad f_y(x,y) = 32 + 3(2y - 4x) > 0$$

for all (x,y)-constellations, all points (x,y) of region (V.5) satisfy $f(x,y) < f\left(\frac{8}{9}, \frac{1}{3}\right)$ as it is graphically illustrated by Figure VII. Since

$$\text{(VII.9)} \qquad f\left(x = \frac{8}{9}, y = \frac{1}{3}\right) = 32\left(2 - \frac{24}{9} + \frac{1}{3}\right) + 3\left(\frac{192}{81} + \frac{1}{9} - \frac{32}{27}\right) = -\frac{61}{9} < 0$$

this proves Proposition 3.

8. Discussion of results

Let us consider the numerical constellation (V.1) which leaves the screening parameters unspecified. Propositions 2 and 3 prove that V is the solution whenever neither U nor V is the best reply vector to the bicentric prior distribution p, i.e. in the shaded area of Figure V. Since (IV.3) and (IV.9) are violated, only V can result in the non-shaded area of Figure V.

Notice, however, that the solution V in the non-shaded and the solution V in the shaded area of Figure V result from different selection principles. In the non-shaded area it is mainly the definition of the bicentric prior p since the best replies to this prior expectation p is already the solution V. In the shaded area the solution relies, however, on the concept of the (piecewise linear) tracing path modelling a continuous transition from best replies to the bicentric prior p to a final equilibrium of the original game.

Based on our results we also want to demonstrate that a solution U is possible where, due to our limited results, it must be a best reply solution U to the bicentric prior p. A numerical constellation allowing to satisfy the requirements (IV.3), (IV.6), and (IV.8) for such a solution (see part (i) of Proposition 1) is

(VIII.1) $\qquad v = 2,\ I = .75,\ J = .5,\ r = .45,\ q = .5.$

Inserting (VIII.1) into these conditions yields

(VIII.2) $\qquad 1 > \frac{8}{11x-9y}$ or $y < \frac{11}{9}x - \frac{8}{9}$

as (IV.3),

(VIII.3) $\qquad 3 > \frac{8}{11x-9y}$ or $y < \frac{11}{9}x - \frac{8}{27}$

as (IV.6), and

(VIII.4) $\qquad \frac{9}{11} > \frac{4-x}{4-3y}$ or $y < \frac{11}{27}x - \frac{8}{27}$

as (IV.9). In Figure VIII the $U(p)$-region is the parameter region in the x, y-triangle where U is the vector of best replies to p and thus the solution, whereas

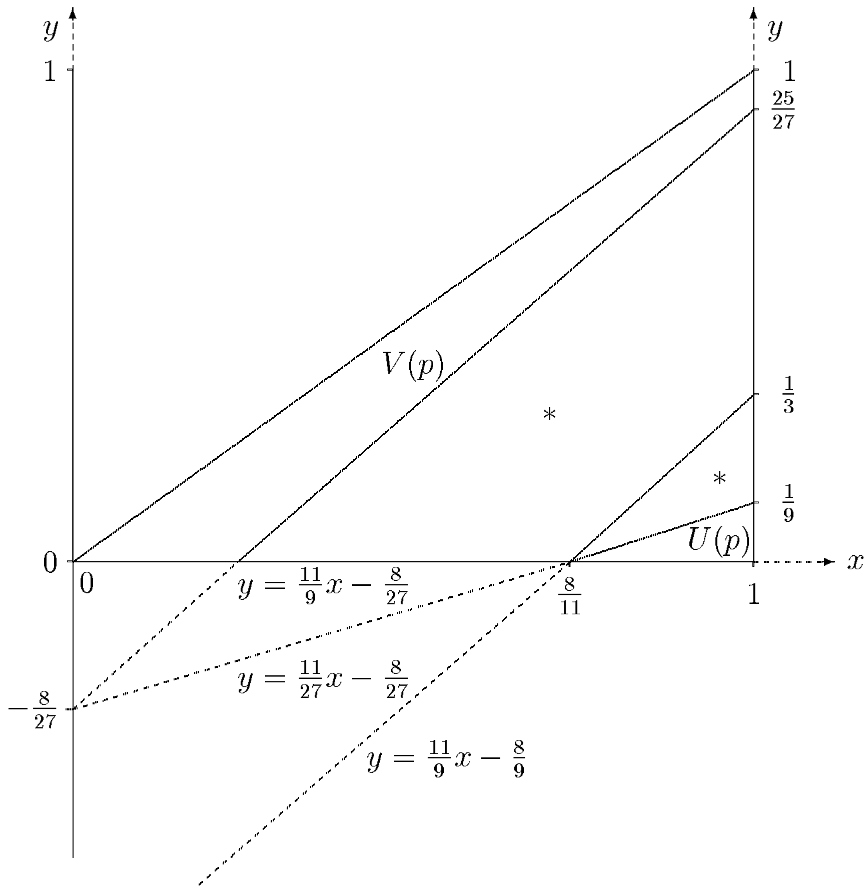

Figure VIII: The $U(p)$ – and $V(p)$ – region for the numerical constellation (VIII.1) where U, respectively V, is the best reply to the bicentric prior p and thus the solution according to Proposition 1 (in the $*$ – regions neither U nor V is the best reply so that one has to apply the tracing procedure)

in $V(p)$-region this holds for V. In the residual parameter regions "$*$" neither U nor V is the best reply vector to the bicentric prior p so that one would have to apply again the tracing procedure.

Figure VIII illustrates that extreme situations in the sense of x and y being close (the $V(p)$-region) or very different (the $U(p)$-region) can be solved by more basic ideas like naive Bayesianism in the sense of $s(0)$. For the intermediate situations (the $*$-region) such basic ideas cannot resolve the coordination problem since $s(0)$ is no equilibrium yet and thus does not survive the introduction of strategic thinking, represented by positive values of t. Here we have demonstrated how more sophisticated concepts like the (linear) tracing procedure allow to single out a unique solution even in the intermediate $*$-region.

One may argue that, although the Trust Game with Banks of Figure I is rather special and simple, our results are still incomplete. Our main excuse is, of course, that a complete solution even of such a simple model would have to rely on a complex case distinction which would go beyond what can be easily represented in an article. One can easily design an algorithm by which one can solve generic parameter constellations like (V.1) and (VIII.1) numerically (for degenerate constellations computational problems like tracing a continuous curve numerically may result). But presenting all numerical results still would overburden the picture since the (too) complex case distinction remains.

A general theory of equilibrium selection is not necessarily one that can be easily applied to a specific example. Its great advantage is that it relies on rationality requirements which make sense regardless of the application at hand. Ad hoc-solution concepts for specific classes of games, e.g. those for signaling games like the intuitive criterion (Cho and Kreps, 1987), which may be easier to apply, should be implied by such a general theory whenever they make sense. What is intuitive or rational in a specific class of games should finally reflect what generally is thought to be intuitive or rational. Theorems stating that for certain subclasses of games easily applied ad hoc-concepts imply the same results as general theories of equilibrium selection would help. But up to now the subclasses are far too special (see Harsanyi and Selten, 1988, and Güth and Kalkofen, 1989).

In our view, we need to develop general theories of equilibrium selection and to apply them to interesting special classes of games even when the results remain incomplete. Other important economic concepts like competitive or efficient allocations are also difficult to derive and nevertheless still dominate our understanding of market economies. Similarly , it may be important to know that there are general theories of equilibrium selection which provide unique recommendations how to play in any type of game. One such instance is, for example, when one wants to derive a unique benchmark solution for an experimental study of a numerically specified game like the Trust Game with Banks. Actually we plan to study experimentally four different (x, y)-constellations satisfying $x + y = 1$ for the numerical constellation (VIII.1).

References

[1] Cho, I. and D. M. Kreps (1987): Signaling games and stable equilibria, *Quarterly Journal of Economics*, 102, 179 - 221.

[2] Güth, W. (1998): Do banks crowd in or out business ethics? - An indirect evolutionary analysis, *Discussion Paper* 109, Humboldt-University of Berlin.

[3] Güth, W. and B. Kalkofen (1989): Unique solutions for strategic games - Equilibrium selection based on resistance avoidance, *Lecture Notes in Economics and Mathematical Systems*, 328.

[4] Güth, W. and H. Kliemt (1994): Competition or co-operation - On the evolutionary economics of trust exploitation and moral attitudes, *Metroeconomica*, 45, 155 - 187.

[5] Harsanyi, J. C. (1975): The tracing procedure: A bayesian approach to defining a solution for n-person noncooperative games, *International Journal of Game Theory, 4, 61 - 94.*

[6] Harsanyi, J. C. and R. Selten (1988): *A general theory of equilibrium selection in games*, Cambridge/MA: M.I.T. Press.

[7] Hildebrand, B. (1864): Natural-, Geld- und Kreditwirtschaft, *Jahrbücher für Nationalökonomie und Statistik*, Bd. 2, 1 - 24.

[8] Schils, R. (1998): Comment on: Specific institutional aspects of international cooperation - A game theoretic account, *Working Paper*, University of Bonn.