# Conventions

## Some Conventional and Some Not So Conventional Wisdom

Siegfried Berninghaus, Werner Güth and Hartmut Kliemt

Abstract

In this paper we consider conventions as regularities in behavior which help to solve coordination problems in a society. These problems can be formalized as non-cooperative games with several equilibria. We know that in such situations serious problems of equilibrium selection arise which cannot be solved by traditional game theoretical reasoning. Conventions seem to be a powerful tool to solve equilibrium selection problems in real world societies. Essentially, two questions will be addressed in this paper: a) Which conventions will emerge in a society? b) How can a society break away from an inferior and reach a superior convention? It turns out that "risk dominance" of a convention plays a crucial role in dealing with both questions and generally in the evolution of conventions.

### 1. Introduction

At least since the times of the British and in particular the Scottish Moralists it has been common practice among social theorist to refer to all social rules as conventions (see (Raphael, D.-D. 1969; Schneider, L. 1967)). The use of the term draws attention to the fact that contrary to the non-mutable laws of nature all human rules of conduct could be otherwise than they in fact are (see on the roots of this distinction in classical thought (Heinimann, F. 1987/1945)). From this point of view even the results of deliberate rule enactment are conventional in the broader sense that people could decide on other rules but we might not want to call the enacted laws of society "conventions".

Besides the quite lose traditional way of using the term "convention" there are more precise technical ones. These clearly exclude enacted laws. But they may diverge in other respects. In the narrow technical sense introduced by Lewis (see (Lewis, David 1969)) conventions are commonly known equilibrium solutions to co-ordination games in which conflict is mostly absent. There have been broader definitions of the term convention that allow at least in a way for conventions as emerging in somewhat more conflictual settings. Such conventions are also solutions of equilibrium selection problems. The evolution of these "conventional solutions" brings about the type of basic social coordination that seems to be at root of all social order (see (Young, H. Peyton 1993; 1998) also (Binmore, Ken 1994)). Indeed an adequate understanding of how the so-called Hobbesian problem of social order (see on this (Parsons, Talcott 1968)) can be solved seems impossible without an adequate account of conventions.[1] Only with such an understanding in hand, we can hope to make some progress towards a new synthesis of concepts, theories and empirical findings in basic social theory.

We will start our discussion with a survey of some theoretical (2.) and experimental findings (3.) on the emergence and maintenance of conventions. We shall then draw some tentative conclusions from the theoretical and empirical findings (4.). Final, more speculative remarks conclude the paper (5.).

## 2. Evolution of conventions: Some basic theoretical results

Let us focus, at least initially, on a simple class of social interaction problems, namely on coordination problems. A coordination problem can be represented by a non-cooperative game with several Nash-equilibria in pure strategies. The equilibria can possibly be Pareto-ranked according to their payoffs (in particular if the game is a symmetric one). In the non-cooperative cordination game the coordination problem emerges because individuals cannot choose an equilibrium unilaterally nor is there an option (like a rule of rule change in the sense of Hart (Hart, Herbert L. A. 1961)) by which individuals could conceivably choose an equilibrium in genuinely collective action.

[1] The essential reasons against solutions of the Hobbesian problem in terms of opportunistically rational choice are presented in Hart, Herbert L. A. 1961. *The Concept of Law*. Oxford: Clarendon Press.). Some reasons why a rational choice solution of the order problem within a folk theorem approach though perhaps possible in principle may nevertheless be inadequate are discussed in Kliemt, Hartmut. 1993. "Constitutional commitments," in *Jahrbuch für Neuere Politische Ökonomie*. Ph. et al. Herder Dorneich ed, pp. 145-73..

## 2.1. A basic example and some basic issues

For the sake of specificity consider a simple *symmetric* 2×2 *coordination game* in normal form in which each player has the strategy set $\Sigma_i$ := {X,Y} ($i$=1,2) and the payoff table2 is given by[1]

Table 1: Payoff table of a symmetric 2×2 coordination game

| | X | Y |
|---|---|---|
| X | $(a,a)$ | $(b,c)$ |
| Y | $(c,b)$ | $(d,d)$ |

where $a > c > 0$, $d > b > 0$, and $a$ $<d$. Clearly, if this game is played one-off, then there are two strict Nash-equilibria $\sigma^*$=(X,X), $\sigma^{**}$=(Y,Y). Obviously, it does not pay for any of the two players to deviate from either $\sigma^*$ or $\sigma^{**}$ unilaterally.

There exists another Nash-equilibrium in mixed strategies which is given by

$$p^* = \frac{d-b}{a-c+d-b} \;;$$

where $p^*$ denotes the probability of choosing strategy X. Since in general it does not make too much sense to conceptualize conventions as based on individual randomization we will not be interested in mixed strategy equilibria of coordination games.

Going along with the "conventional" use of the term "convention", we think of a convention as a regularity in behavior shown by individuals chosen from a typically large population in recurrent inter-individual interactions. The regularity is a convention in the more narrow sense of that term if it is true and common knowledge in the relevant population that

- everyone conforms to the regularity,

[1] Each payoff combination in the payoff table represents the payoff of player 1 (first component) and player 2 (second component) as induced by some particular strategy combination. Due to the symmetry of the game the roles of the players are inter-changeable.

- everyone expects everyone else to conform to the regularity, and
- everyone prefers to act in conformity with the regularity to performing non-conformist actions if the others act in conformity with the regularity.

For the class of 2x2 interactions under consideration here, this informal characterization is sufficiently precise. The definition shows that a convention can be regarded as a strategy configuration in a population (society) in which in pair-wise interactions all members choose the same Nash-equilibrium strategy in a basic 2×2 game. Different conventions guiding individual strategy choices in recurrent interactions correspond to different Nash-equilibria.

In the social interaction of a large group we cannot expect a particular convention to be established instantaneously. Among the many the emergence of conventions should rather be imagined as the result of a gradual process of strategic adaptation. If the process of gradual strategy adaptation is modeled by evolutionary game models two important problems concerning conventions may be answered in a precise way:

1. How do conventions evolve?
2. How can a society switch from bad – payoff dominated – to better conventions?

Obviously, both questions are interrelated. We will analyze the questions from a theoretical point of view first and will then give a brief overview of related experimental results.

## 2.2. Replicator dynamics

The evolution of strategy configurations in large populations can be modeled by so called *replicator dynamics*. Replicator dynamics describe strategy adaptation in continuous time by a simple differential equation. According to this equation those strategies are favored which have higher payoff than the average (beat the average!).

More specifically and precisely, assume that the simple symmetric 2x2 game of table 1 is played repeatedly among players selected from a large population. Before each round of play players are purely randomly matched to play the game of Table 1. A convention exists if either all players tend to choose X or all tend to choose Y and know of each other that they do.

Let us denote by $x(t)$ the proportion of players choosing strategy X at time $t$. Replicator dynamics specify that $x(t)$ evolves according to the differential equation

(RP) $$\frac{\dot{x}(t)}{x(t)} = [(AX(t))_1 - X(t)AX(t)],$$

where $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ denotes the payoff matrix derived from Table 1. The strategy distribution in the population is given by the vector $X(t)=(x(t), 1-x(t))$; where, of course, $x(t)$ denotes the probability that X and $[1-x(t)]$ denotes the probability that Y is chosen on round t by any randomly selected individual from the large player pool. The scalar $X(t)AX(t)$ resulting from multiplying $A$ from the left and the right with the probabilities of using the strategies is the average payoff in the population. $AX(t))_1$, denotes the expected payoff of choosing strategy X or more formally the product of the first row of $A$ with $X(t)$.

Bearing the preceding in mind we read (RP) as follows:

> The left hand side is the growth rate of the share of strategy X in the overall population. If the average payoff of strategy X is smaller than the average payoff of all strategies then the population share of players choosing strategy X will decline and it will rise otherwise. The size of the growth rate – or how swiftly or slowly decline or rise take place – depends on how much the average payoff accruing to strategy X differs from the average payoff $X(t)AX(t)$.[3]

The differential equation (RP) characterises completely how strategy configurations that correspond to a particular convention in which either all players choose X ($x$=0) or choose Y ($x$=1) may emerge.[4] Reformulating (RP), we finally obtain

---

[3] The dynamics characterized by RP are well-behaved in that they cover the cases of spreading dominant and ruling out dominated strategies.

[4] From a mathematical point of view we consider the simplest version of the replicator dynamics, the 2-dimensional case which can trivially be reduced to a single differential equation. All statements we make refer to this simple dynamical system. The statements have to be qualified appropriately when more general coordination games are taken into account (for general results see, e.g. Taylor, P.D. and L.B. Jonker. 1978. "Evolutionarily Strategies and Game Dynamics." *Mathematical Biosciences*, 40, pp. 145-56.).

(RP') $$\frac{dx}{dt} = \dot{x}(t) = x(t)(x(t)-1)[(d-b)-x(t)(a-c+d-b)].$$

Inspecting this equation the following conclusion can immediately be drawn:

Since both $x^*=1$ and $x^{**}=0$ imply $\dot{x}=0$, both conventions ($x^*=1$ and $x^{**}=0$) form stationary rest points which are also asymptotically stable.

Because of our assumptions we know that the mixed equilibrium fulfils $p^* \in (0, 1)$ and the interval (0, 1) between the two stationary rest points can be divided into a connected sub-set $(0, p^*)$ over which $x(t)<p^*$ declines towards $x^{**}=0$ and into a connected sub-set $(p^*, 1)$ over which $x(t)$ grows towards $x^*=1$. Each of the two conventions has a non-empty domain of attraction whose size depends on the value of $p^*$. The value of $p^*$ depends on the relative size of the payoffs – the losses a-c and d-b of unilateral deviations from any of the strict equilibria – of the specific coordination game under consideration. In our class of coordination games the mixed strategy equilibrium $p^*$ uniquely fixes the best reply structure of the game which assigns to each strategy the set of mixed strategy vectors to which the strategy is a best reply. It follows from (RP') that $p^*$ is also a stationary rest point for the replicator dynamics which is, however, completely unstable. When the population ever should get to the percentage $p^*$ of players selecting X, a slight deviation from $p^*$ suffices to drive the population either to $x^*=1$ or $x^{**}=0$.

In Table 1 we considered a coordination game in which co-ordinating on strategy X is payoff inferior to co-ordinating on strategy Y. However, this does not imply that all players should "rationally" choose strategy Y. Consider for example the following payoff table of a symmetric coordination game.

Table 2: Payoff table of the BKE-game

| | X | Y |
|---|---|---|
| X | (80,80) | (60,10) |
| Y | (10,60) | (90,90) |

Obviously, strategy combination (Y,Y) payoff-dominates the combination (X,X), but it may nevertheless pay for an individual player to choose strategy X. For example, if a player does not trust that her co-player would be guided by considerations of payoff dominance she could argue: If I choose Y and am matched with a player choosing X, then I will “lose” 70 units by missing the payoff of 80 and ending up with 10 instead. However, if I am choosing X and am matched with a player choosing Y my potential payoff loss will amount merely to 90-60=30 units. Therefore, strategy X may be deemed “less risky”.

Harsanyi and Selten (see (Harsanyi, John C. and Reinhard Selten, 1988)) formalized this intuitively appealing idea by what they called “risk dominance. For symmetric 2×2 games one has a simple criterion to check for risk dominance:

> *In a symmetric 2×2 game as presented in table 1, the strategy X is risk dominant iff inequality*
>
> (RD) $(a\text{-}c) > (d\text{-}b)$
>
> *holds; while strategy Y risk dominates X if the reverse inequality applies.*

Obviously, for the game in Table 2 this criterion is satisfied for strategy X.[5]

In understanding the evolution of conventions risk dominance is an important concept. In particular in the coordination games considered here principles of risk dominance and of payoff dominance must be seen together. It is the “tension” between the two[6] that makes the issue of whether a convention and if so which one will emerge in an interaction particularly interesting.

The process in which a convention can emerge is modeled here basically by the dynamics of $x(t)$ or the dynamical development of the percentage of players choosing X. The threshold $p^*$ is related to considerations of risk dominance. If strategy X is risk dominant we have $p^* < 1/2$

---

[5] For a brief discussion of risk dominance including the non-symmetric case see Young, H. Peyton. 1998. *Individual Strategy and Social Structure. An Evolutionary Theory of Institutions*. Princeton: Princeton University Press., chap. 4.1

[6] Transferring the game in Table 2 such that

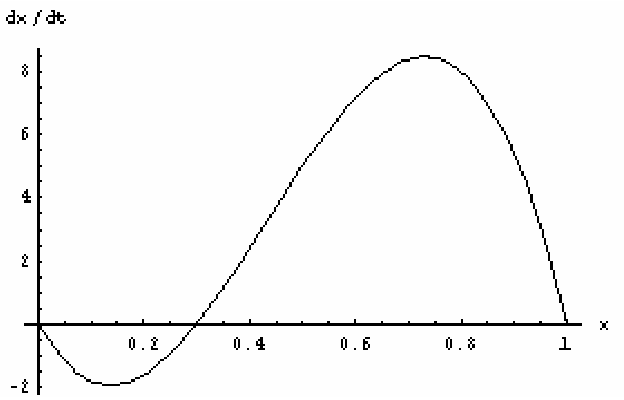
- for player 1 one subtracts his non-equilibrium payoff of 10, resp. 60 in column 0, resp. 1,
- for player 2 one subtracts his non-equilibrium payoff of 10, resp. 60 in column 0, resp. 1,

preserves the mixed strategy equilibrium $p^*$, the best reply structure and thus the risk dominance relation but would render $x^*=0$ as payoff dominant.

and the domain of attraction of the payoff inferior convention is larger than the domain of attraction of the payoff superior convention.

The situation is illustrated by the follwing diagram which shows how the time derivative $\dot{x}\,(t)$ of the percentage of X-players depends on *x(t)* in the coordination game of Table 2.

Figure 1: Dynamics of the coordination game of Table 2




What can we learn from our discussion of this simple example of the evolution of conventions?

1. Risk dominant conventions have a larger basin of attraction and thus a payoff dominated convention may be “more stable”[7] than a convention that payoff dominates it but is risk dominated.

2. That evolutionary replicator dynamics are guaranteed to work towards the emergence of a convention with an undominated outcome only if the payoff dominant convention also is risk dominant.[8]

[7] By this term we mean: If the society “leaves” the risk dominant convention in the sense that few society members deviate from this convention it will be restored rather soon. Or put it the other way round, if we consider the strategy adaptation process starting from a randomly selected point $x(0)$ ⊙ [0,1] there will be a higher chance that the strategy adaptation proess generated by (RP) will end up in the risk dominant convention.

[8] Consider, for example, the 2×2 coordination game characterized by a=80, b=c=0, d=90. In this game the equilibrium strategy configuration (Y,Y) is both risk and payoff dominant.

The preceding elementary example can serve to illustrate some basic analytical results on conventions in the simplest possible form. Adding mathematical complexity we might hope to capture some further aspects of coordination problems of the preceding kind. However, for our present purposes this is not a fruitful route. Such analytical exercises – including much more complicated ones – cannot capture the basic fact that human social interaction is non-homogeneous in the sense that we by our human nature are disposed always to distinguish between the "socially close and the socially remote". Our emotions follow the natural sequence close, closer, closest. Ever since the times of David Hume social philosophers have seen this and argued that the human proclivity to distinguish between the close and the remote both along the time as along the social dimension is a fundamental fact influencing all human social organization. This fact is utilized when the division of labor is extended to the enforcement of norms in human institutions and governance structures (see [Hume, [1739]1978 #630] in particular book III, chap. section on the origin of government). Small groups form the backbone of that process and any ordered large group interaction. Organizing the many in small groups must manage to coordinate on within-group equilibria that as a side-effect bring about the enforcement of norms applying to individuals outside the group. But the small groups are not aliens outside society. In a way rulers and ruled always live in the same society still. Basic conventions guiding behavior will be operative across all the smaller groups of which a large society is composed. But if that is so the question of how the fundamental preference for the near as compared to the more remote and the formation and stability of conventions naturally emerges. – To take into account the effects of the "topology" of social interaction as ingrained in human nature we need to go beyond the simple analytics developed so far to somewhat more complicated ones.

## 2.2. Local interaction structures

That *local interaction structures* can indeed influence the evolution of conventions is intuitively plausible. It has been discussed in a more formal vein in a stimulating paper by Boyer and Orleans (1992). They assume that players are only matched with members of a neighborhood group that is smaller than the total population. For example, in a "nearest neighbor interaction" each player is assumed to have exactly two neighbors with whom he interacts in dyadic co-ordination games. This structure in which each player has precisely one “left” and one “right” neighbor can be illustrated graphically by positioning the players on a circle.

Fig. 2: Nearest neighbor interaction

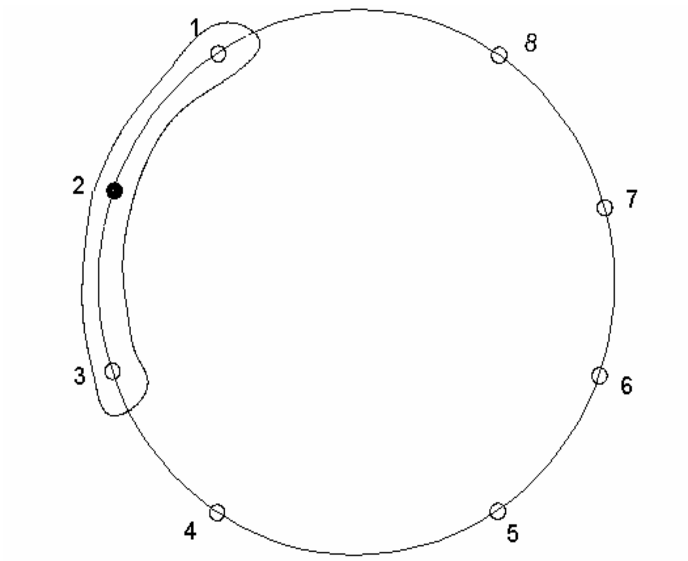



Figure 2 represents a nearest neighbor structure in a population of 8 players. Player 2, for instance, has two direct neighbors (player 1 and 3). Players 1 and 3 again have two neighbors (8 and 2, respectively 2 and 4) etc.

More generally speaking let $N_i$ denote the set of neighbors of player $i$, e.g. $N_2$={1, 3}. We can refer to neighborhood size by $|N_i|$; e.g. in the example in hand $|N_i|$=2 for all i. A somewhat more complex interaction structure emerges by assuming $|N_i|$=4. In this case each player directly interacts exclusively with two right and two left direct neighbors. In figure 2, for example, then player 2’s neighborhood is $N_2$ = {3,4,1,8}. For a group of size $|N|$ all neighborhood structures with $|N_i|$, $|N|> |N_i|$, can form local interaction structures with homogeneous group size.

That local interaction structures matter is intuitively plausible. If, for instance, all members of a finite population of size $N$ play convention Y in the co-ordination game of table 2 then the payoff dominant result emerges. For sufficiently large $N$ this result will prevail under random matching even if a single individual deviates. For a non-deviant player the probability to encounter a player who plays X is then 1/($N$-1). The probability to encounter a player who is still playing strategy Y is ($N$-2)/($N$-1). For a non-deviant player the expected payoff of strategy X is $[1/(N\text{-}1)]*80 + [(N\text{-}2)/(N\text{-}1)]*60$ while the payoff of strategy Y is $[1/(N\text{-}1)]*10 + [(N\text{-}2)/(N\text{-}1)]*90$. With these specific parameter values, we have $[1/(N\text{-}1)]*80 + [(N\text{-}2)/(N\text{-}1)]*60 > [1/(N\text{-}1)]*10 + [(N\text{-}2)/(N\text{-}1)]*90$ for all $N>4$.

Under random matching a single deviation from strategy Y to strategy X will not destabilize the co-ordination equilibrium in the case at hand. However, if we assume that the interaction is characterized by a nearest neighbor interaction structure rather than by random matching the prediction changes. Then the deviating player *j* will interact only in the neighborhood rather than being matched with everybody with the same probability. If every player *j* has precisely two nearest neighbors and player *i*, $i \neq j$, is a nearest neighbor of *j*, then after *j*'s deviation to strategy X player *i* interacts with exactly one player using strategy X and one player using strategy 1. Provided that equi-probability of matching applies in the neighborhood the payoff expectation of player *i* is (80 + 60)/2 if she chooses strategy X and (10 + 90)/2 if she sticks to strategy 1. Since 80 + 60 > 10 + 90 *i* will "join" *j* and switch to strategy X, too. The remaining neighbor of *i* will follow suit in the next period and so on. Obviously, if players, after deviating, do not immediately switch back on the next round of play, the process will in a finite number of steps induce the emergence of the co-ordination equilibrium in which all individuals play strategy X without exception.

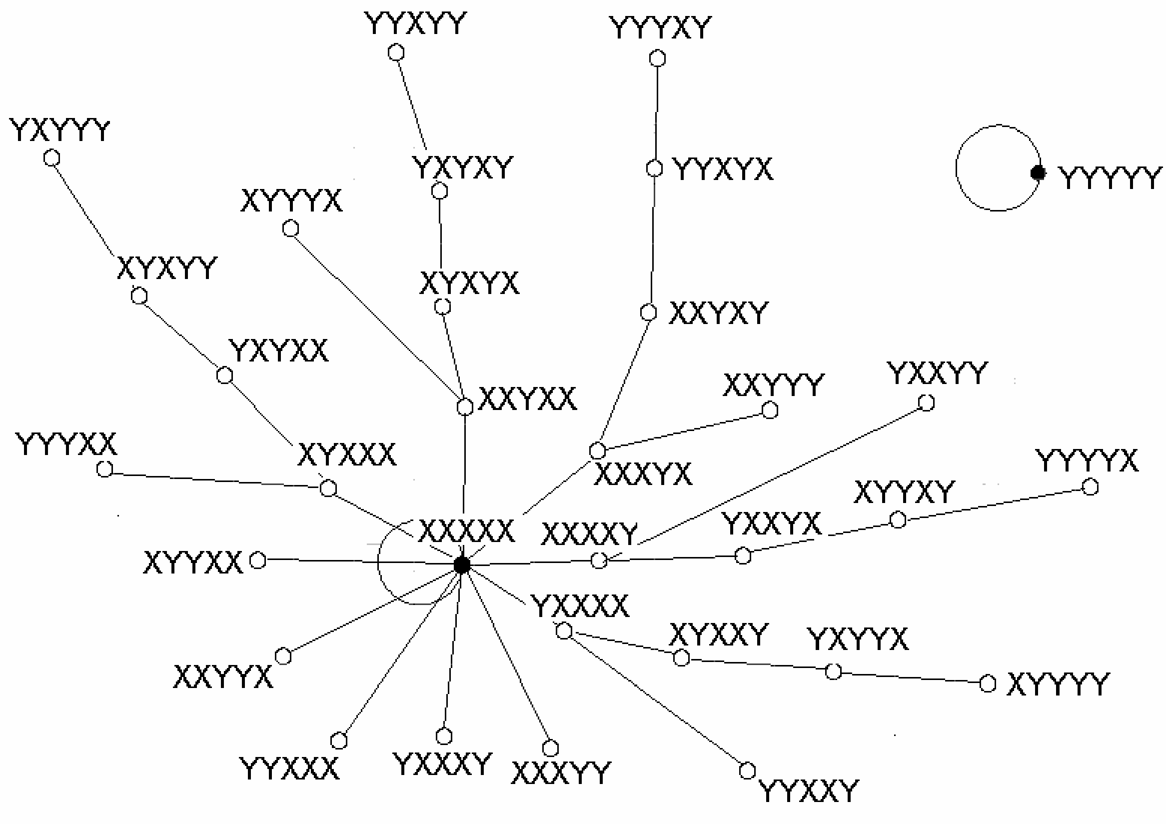
The preceding argument compares the polar extremes of a circular structure in which players interact only with their two nearest neighbors with random matching across the population at large. The comparison shows that a risk dominant convention may drive out a payoff dominant but risk dominated convention through strategic transmission by local interaction structures.

Iteration graphs may be used to study how the distribution of strategy choices unfolds in iterated interactions. For a group *N* of individuals any such distribution can be represented by an $|N|$-tupel listing the strategy choices of the $|N|$ individuals in their natural order.

For illustration consider dyadic interactions as specified by the coordination game presented in table 2 taking place on a circle in a group with $|N|=5$. Individuals have direct intercourse with each other only within their neighborhoods. In the first case assume that for any $i \in N$ the neighborhood is characterized by $|N_i|=2$ while in the second case $|N_i|=4$ applies. The game takes place in rounds. All players behave according to the "local best reply" principle by choosing the strategy that maximizes their payoff against the strategy distribution of their neighbors as observed in the previous period. Under these assumptions, if all start out with playing the payoff dominant convention, then this corresponds to a strategy profile YYYYY.

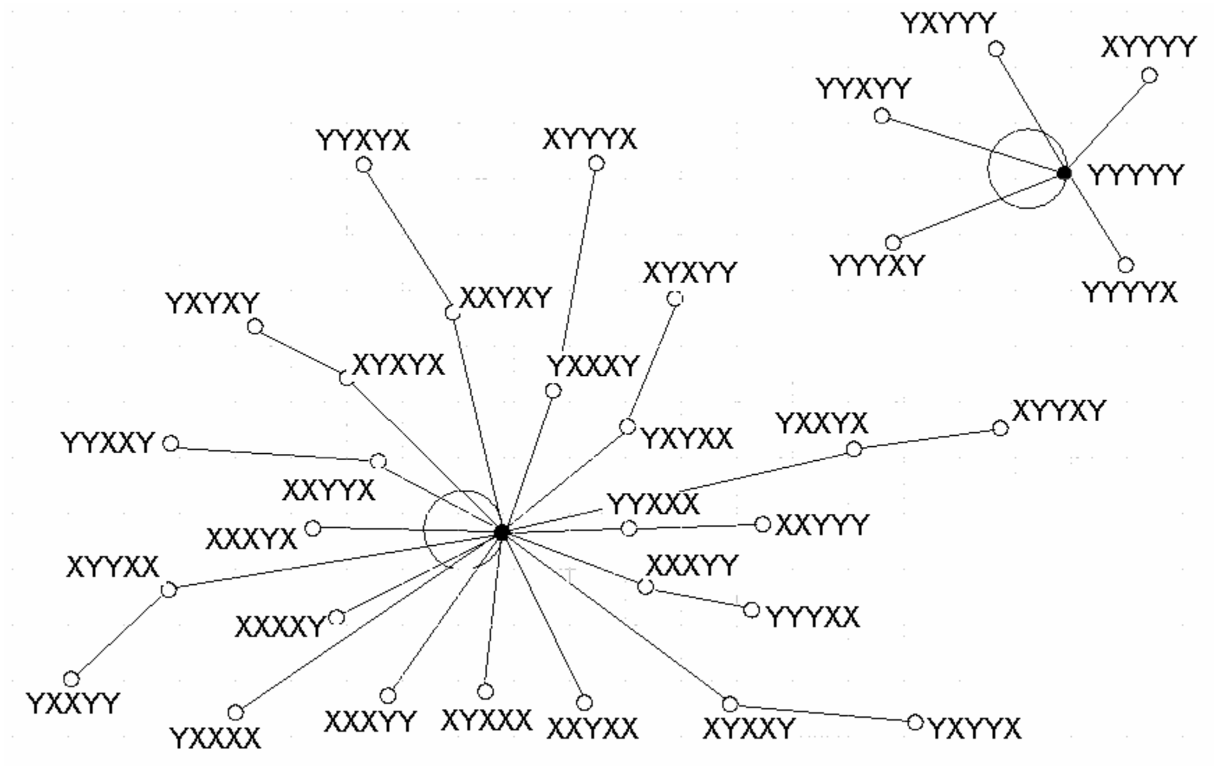
The profile YYYYY will be transformed into YYYYY on each round of play. This result could persist indefinitely in principle if there would be no deviations (trembles or mistakes). But the payoff dominant convention is under a permanent threat of being overthrown by some deviation. For instance YYYYX will successively be transformed into XYYXY then into YXXYX then into XXXXY and end with XXXXX. The profile XXXXX as opposed to YYYYY (the dark dots in Fig. 3) cannot be destabilized by single deviations. In fact nothing short of collective action to switch to YYYYY will do. For, as can be seen quite easily the distribution XXXXX will result under local best reply adaptation in neighborhoods of size 2 for any of the $2^5$-1 possible distributions that contain at least one X as a strategy choice. The next figure illustrates this.

Figure 3: Iteration graph for $|N|$=5, $|N_i|$=2




Obviously the payoff dominant convention is totally unstable if the interaction structure assumed here prevails. As the next example shows this may change, however, if we enlarge the neighborhood size to $|N_i|$=4.

Figure 4: Iteration graph for $|N|$=5, $|N_i|$=4




As the graph of figure 4 clearly illustrates by increasing the neighborhood size the payoff dominant convention can become “more stable” in the sense that unilateral deviations are restored after one period. Indeed, for quite a large class of co-ordination games increasing the neighborhood size tends to “stabilize” the prevalence of payoff dominant conventions. One should carefully note, though, that such results depend on specific parameter values to some extent and therefore do not straightforwardly imply policy recommendations (for details see Berninghaus and Schwalbe, 1996).

Risk dominance seems to guide the evolution of conventions in a local interaction framework. In this setting the strategy distribution in the population at large is irrelevant for a player who interacts only with $j \in N_i$. A player who chooses according to the local best reply principle will only respond to the proportion of individuals who in his neighborhood play each of the strategies. More specifically let $k_i$ be the number of neighbors who choose the strategy X and $|N_i|$ be the neighborhood size. Clearly under random matching in the neighborhood $i$´s choice should be affected exclusively by the distribution of strategy choices ( $\frac{k_i}{|N_i|}$ ,$(1-\frac{k_i}{|N_i|})$ ) in her direct neighborhood.

Relying on the general form of the co-ordination game as presented in table 1 with parameters *a*, *b*, *c*, *d* we can state that player i will choose strategy X iff

$$\frac{k_i}{|N_i|}a+(1-\frac{k_i}{|N_i|})b>\frac{k_i}{|N_i|}c+(1-\frac{k_i}{|N_i|})d$$

This inequality can be transformed into

$$\frac{k_i}{|N_i|}>\frac{d-b}{a-c+d-b}$$

Suppose now that X is a risk dominant strategy. Then ($a$-$c$) > ($d$-$b$) and the fraction $\frac{d-b}{a-c+d-b}$ on the right hand side of the inequality is smaller than 1/2. Consequently, as long as $\frac{k_i}{|N_i|}\geq 1/2$ there are sufficiently many players around to suggest the choice of the risk dominant strategy X. In a nearest neighbor interaction with $|N_i|=2$ the number of strategy choices need merely fulfill $k_i=1$ to render the choice of the risk rather than the payoff dominant strategy optimal.

The preceding discussion suggests two general conjectures:

1. A risk dominated (relative to the basic dyadic interaction rather than relative to the interaction at large) yet payoff dominant equilibrium can be rendered more stable by decreasing the neighborhood size.

2. If the risk dominant and the payoff dominant equilibrium coincide decreasing neighborhood size will tend to stabilize the payoff dominant equilibrium.

Empirical evidence as gathered from experiments on the formation of conventions supports the conjectures. To this evidence we now turn.

## 3. Evolution of conventions: Experimental Results

How conventions may evolve in practice can be tested by letting individuals play coordination games in a PC-Lab. Van Huyck, Battalio and Beil (1990) demonstrated that risk plays a crucial role in experiments based on a particular coordination game ("weakest link game"). The risk that an individual incurs by behaving according to a convention if coordination fails is – at least in a way – more important for the emergence and stability of conventions in recurrent interactions than the chances of successful coordination; i.e. risk dominance trumps payoff dominance.

Cooper et al. (1992) came to a similar dominance result by conducting experiments in which each player was matched with anonymous partners to play a simple symmetric coordination game.[9] The payoff structure of the basic game of the experiment was such that the payoff dominant convention was risk dominated. The experimental results corroborated the theoretical findings about how risk could influence the evolution of conventions. In none of the matches studied by Cooper et al. both players chose the payoff dominant strategy configuration. In most matches (about 97%) players chose strategies leading to the risk dominant equilibrium.

Moreover, in the experiments participants showed a higher tendency to choose the payoff dominant equilibrium if they were offered an additional *outside option*. The outside option had a payoff higher than the risk dominant equilibrium payoff. Players could choose between the outside option and the coordination game with one still higher payoff dominant outcome. In a kind of forward induction argument rejection of the outside option was obviously treated by the players as a signal that rejecting players intended to coordinate on the payoff dominant equilibrium. As reported by Berninghaus and Ehrhart (2001, section 4), similar results were obtained when subjects had to pay for participating in the coordination game.

The aforementioned experiments offer important and valuable insights. However, the experiments do not deal sufficiently with the central issue of how local interaction structure may affect the formation of conventions. Yet, according to the preceding observations on "human nature" inhomogeneities in the likelihood and frequency of interaction through time should be expected to exert considerable influence on the emergence of conventions. These

[9] The payoff of this game were given by a=800, b=800, c=0, and d=1000. One can easily check that (X,X) is risk dominant but not the payoff dominant Nash-equilibrium.

observations suggest that experiments on the evolution of conventions in a population with a local interaction structure should be conducted.

At the University of Karlsruhe such experiments (see Keser et al., 1998, and Berninghaus et al., 2002) addressing chiefly two questions were performed:

1. Does the local interaction structure matter in practice? Or are the theoretically relevant local interaction structures artificial constructs which have no significance for real world games?
2. Assuming that local interaction does matter, will different interaction structures have a different impact on the evolution of conventions?

The experiment addressing these issues contained several different treatments. In each treatment 8 groups of size *N* were formed. Either a specific local interaction structure (treatment I) or no local interaction structure was imposed on each of the groups (treatment II). All players knew whether they were members of a local interaction structure or not. However, members of a neighborhood, except for the neighborhood size, were not informed about the specific details of the structure. In each period each group member, *i*, had to select a strategy out of the strategy set $\Sigma_i=\{X,Y\}$. On each round of play all players played sequentially with all members of their neighborhood the 2×2 coordination game with the payoff matrix as presented in Table 2 above. The individual payoff of each player per period was the average payoff gained by playing against the neighbors. The game was played 20 times.

The results of treatments I and II, respectively, are particularly relevant for dealing with problem 1: In treatment I population size was fixed at $N$=8, while the local interaction structure was the "nearest neighbor interaction", that is, neighborhood size was equal to two. In treatment II population size was fixed at $N$=3 and no local interaction structure was imposed. Individual payoffs were calculated as in treatment I. We compare the results of choices of strategy X (the risk dominant equilibrium) by presenting two frequency diagrams.

Figure 5: Percentage of X-choices in treatment I

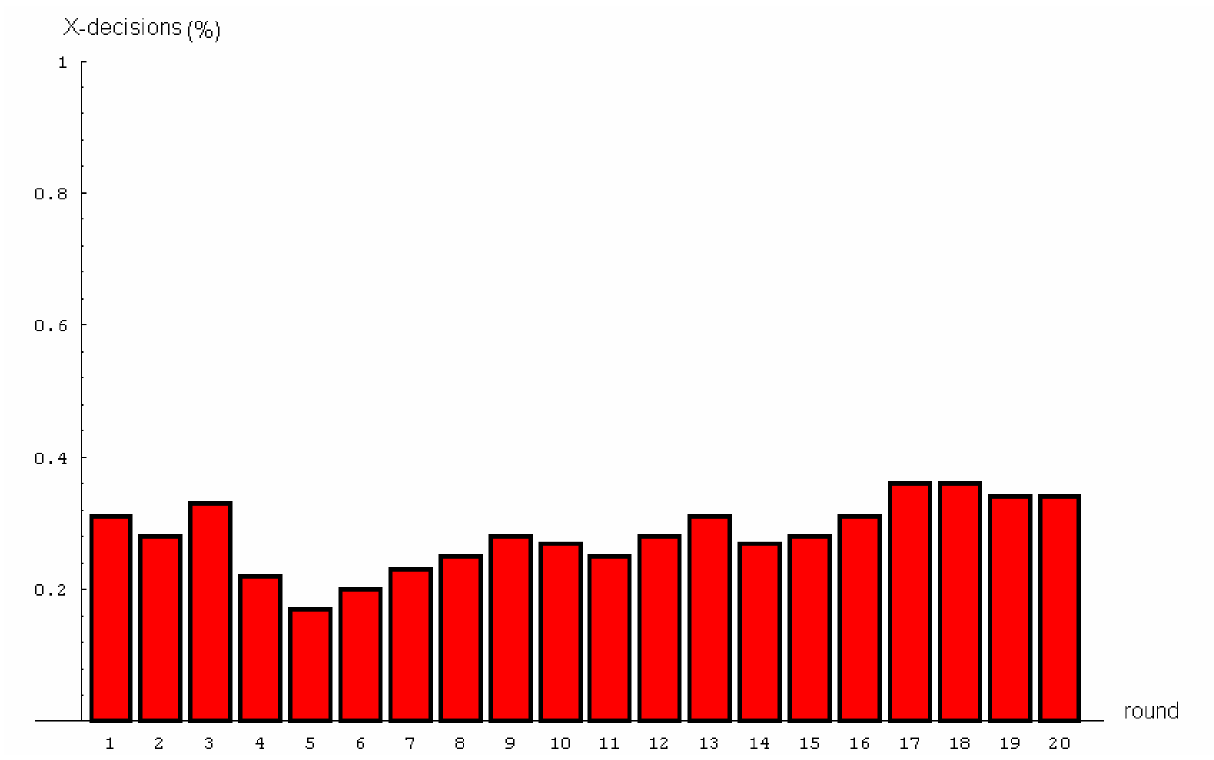
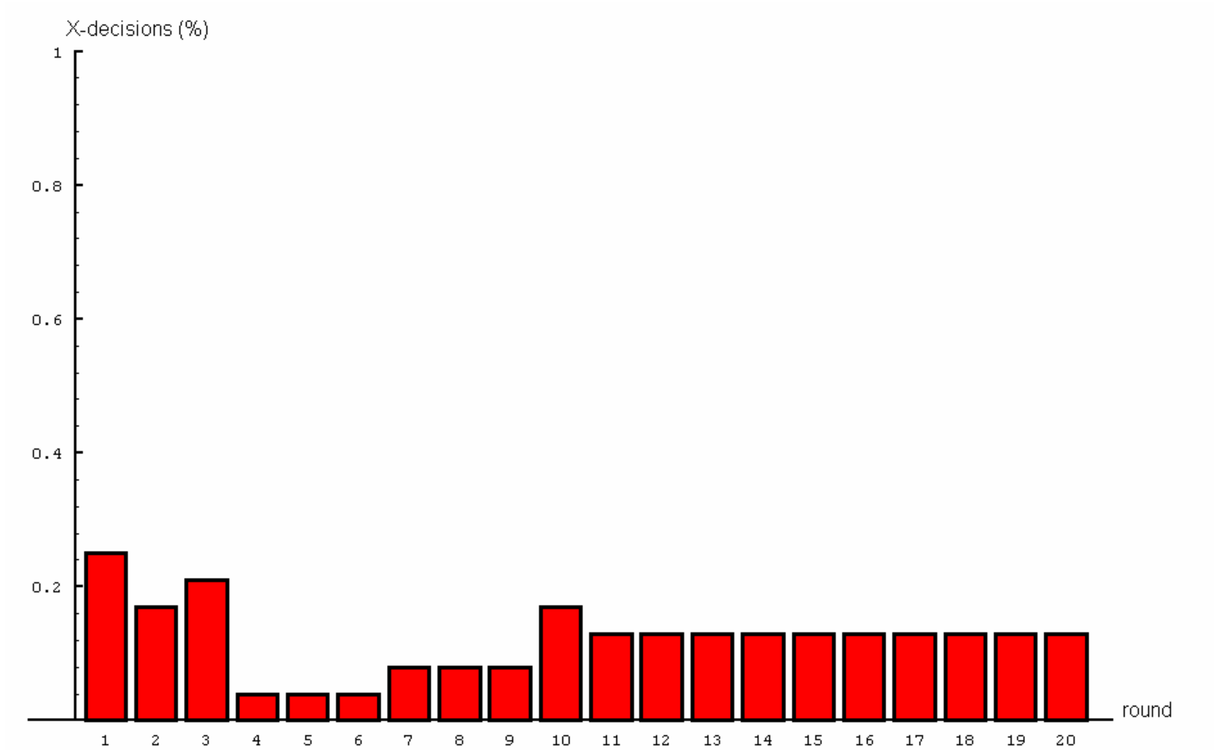



Figure 6: Percentage of 0-choices in treatment II



Although there is no significant difference in the strategy choices over all periods ($\chi^2$ test, $p$=0.196), we see a clear tendency of players in treatment II to choose the payoff dominant convention while under the nearest neighbor local interaction structure of treatment 1 they tended to choose relatively more often the risk dominant strategy. Both interactions were based on triples. In treatment I each player directly interacted with her two direct neighbors (on the circle) only. In treatment II there was no specific neighborhood structure but population size was restricted to $N$=3. The similarity between the two setups brought about by focusing on triplets seems sufficient to support the conclusion that local interaction matters for the evolution of conventions.[10]

After answering the first question affirmatively let us briefly look at experiments addressing the issue of whether or not the specific form of the interaction structure matters. Two further treatments (III and IV, respectively) which were characterized by $N$=16 and $N_i$=4 were analyzed. The local interaction structure imposed on the populations in treatment III was a circular one (similar to the interaction structure in Figure 2 except for each player having *two* left and *two* right neighbors). The interaction structure imposed on the populations in treatment IV was that of a torus on which players were allocated in a two dimensional grid. Each player had exactly four neighbors (so-called v. Neumann neighborhood) who were occupying the four adjacent fields on the torus. The results of treatments III and IV are presented in the following diagrams.

[10] In Keser et al. (1998) it is shown that the difference in behavior between treatments I and II is much more dramatic if the per period payoff is the minimum of what a player gains in all interactions with his neighbors.

Figure 7: Percentage of X-choices in treatment III

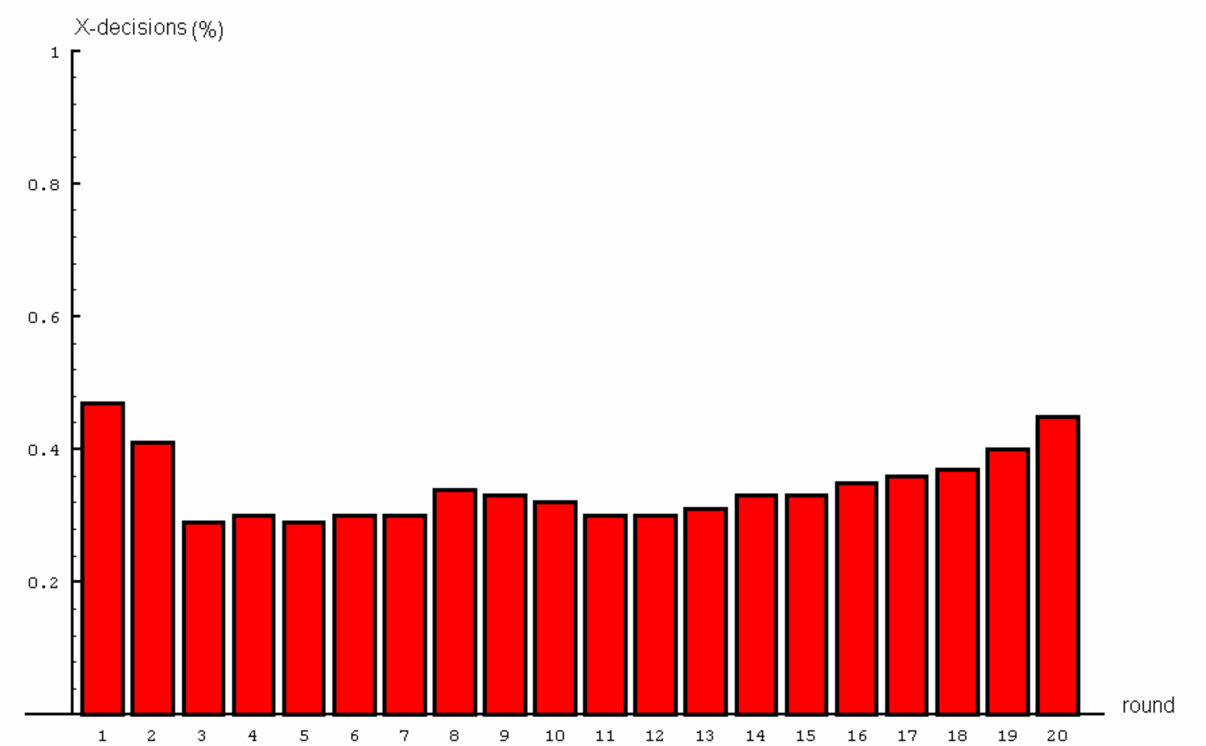
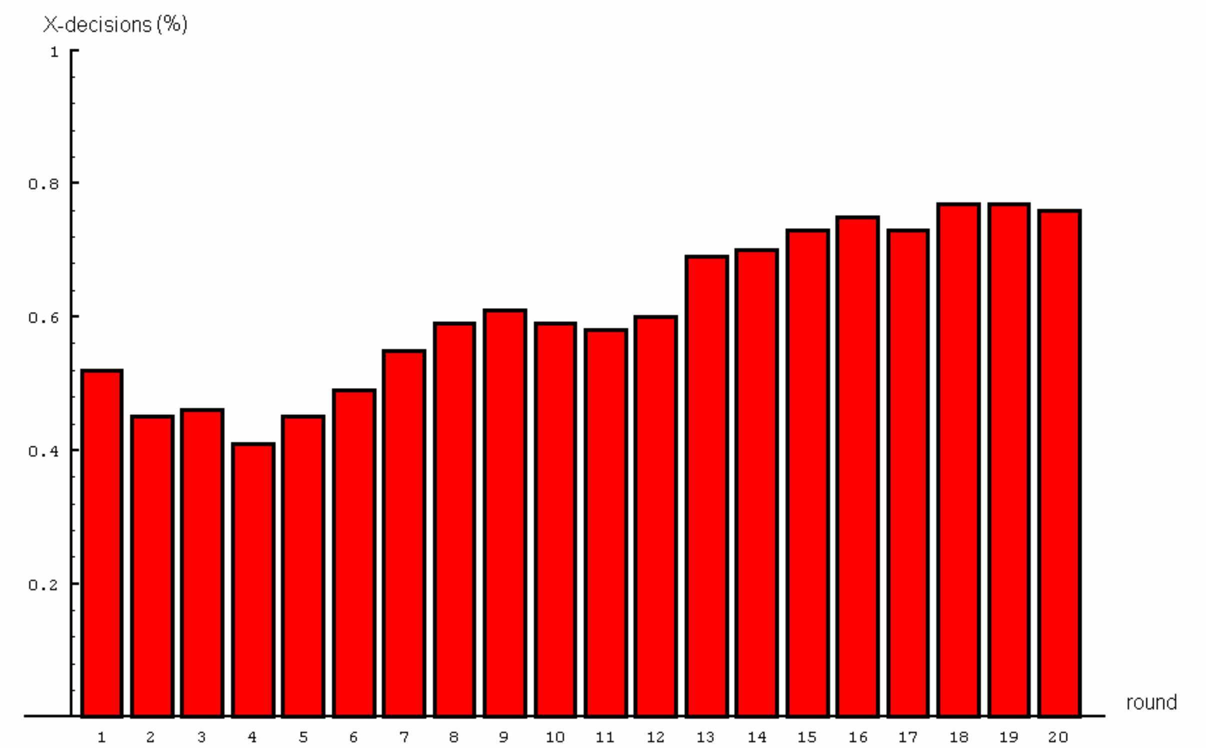



Figure 8: Percentage of X-choices in treatment IV



We see significant differences ($p<0.05$) between treatments III and IV. In the torus interaction structure much more choices of the risk dominant strategy were observed than in the circular interaction structure.

Again the explanation suggested by observed choice behavior in the experiment supports the view that risk dominance is of essential importance to understanding the evolution of conventions. In the torus environment players tended to switch between strategies much more frequently than when playing on the circle. Due to this their co-players were confronted much more frequently with strategy switches. Since there was more variability in the strategy choices in the torus environment players would be inclined to prefer the less risky, that is, the risk dominant strategy in this setting.

In sum the theoretical results about the evolution of conventions in populations with local interaction structure are generally supported by experimental evidence. In theory and practice risk dominance is a guiding principle in the process of selecting conventions. The risk that coordination may fail is in general more important to players than the chance that coordinating on a payoff dominant alternative may succeed.

**4. Conventions, coordination, and stability**

Early modern accounts of conventions starting with Lewis tended to downplay the possibility of conflict in coordination games. Game theoretically inspired approaches to so-called "games of pure coordination" would have it that players more or less automatically should coordinate on strategy choices leading to payoff dominant equilibria. The preceding presentation of theoretical findings and experimental results shows that this is far from being self-evident. The presence of risk along with that of local interaction structures can potentially explain why even in situations in which at first sight conflict seems to be absent, players may tend to coordinate on payoff dominated rather than on payoff undominated outcomes.

Sections 2 and 3 show theoretically and experimentally that even in case of so-called pure coordination games there is no reason for the optimistic assumption that surviving conventions will be efficient. If a certain combination of risk dominance and local interaction structures prevails then, after social evolution has run its course, the outcome will in all likelihood be inefficient. We observe:

- the optimistic view that social adaptation shall lead to efficient results becomes doubtful,
- there may be some scope for political interventions justified solely by efficiency concerns such as to influence the risk or the interaction structure,
- interventions aiming at efficiency may fail to produce a stable result since they may aim at risk dominated solutions,
- local interaction structures do matter and any policy that neglects them may indeed endanger or even wipe out the basis of social stability.

The prevalence of risk dominance as a decisive consideration in the process in which conventions emerge is well in line with the classical philosophical view that social co-operation is not characterized by prisoner's dilemma but rather by assurance and chicken game aspects. Individuals do not feel that they have dominant strategies but rather that they can win or lose depending on how they manage to co-ordinate with others. The risk that co-operation even if intended by all individuals may fail is a very real one. Considerations of risk dominance are relevant in the sense of being prepared for contingencies and are driving individual behavior.

In our view, the preceding account of coordination games with multiple strict equilibria suggests that "public relation management", "appeals to the public", "moral suasion" and the like may be perfectly respectable measures of economic policy or even therapy. Whereas usually economic policy is either directed at institutional aspects (in the sense of mechanism or constitutional design, in German "Ordnungspolitik") or at changing parameters, e.g. by monetary or tax policy or by anticyclic government expenditures, "equilibrium selection policy or therapy" does not aim at changing the rules but at (re)coordinating expectations such that society switches from one strict to another more preferred equilibrium. Imagine, for instance, that consumers "save too much" that firms reduce employment and investment since they expect low demand in the near future. Refocusing expectations such that the future looks brighter might induce consumers to spend more since they do not expect (more) unemployment in turn avoiding the need of cutting employment and investment activity.

These thoughts may not be entirely new but thinking more clearly about conventions shows that they deserve to be taken seriously. Since evolution and behavioral adaptation may not lead to efficient conventions there is room for "equilibrium selection policy or therapy". In view of this we see too little theoretical and political discourse discussing the premises, the difficulties and the practical aspects of such a form of policy. If "moral suasion" can play a role in politics besides direct interventions, institutional or other, it needs to be analyzed more thoroughly and should eventually be incorporated in the social science account of economic policy. Not "the economist as a preacher" should be the aim here but rather the "economics of preaching".

**5. Final remarks**

There may be no need to have concerted action for guiding individual conduct in inter-individual encounters. Individuals who know the customs of the land know how to get along with one another even if they have not met and will in all likelihood not meet again. Even if they do live in local interaction structures of relatively close-knit societies of individuals who interact with each other over extended periods of time while having only few outside contacts conventions may cohere between neighborhoods. In particular there can always be some "overlapping convention" that connects neighborhoods of individuals who do not directly interact with each other. Whenever individuals who belong to the same such "social nexus" interact with individuals who are not from their own neighborhood they can still coordinate.

The large scale interactions as typically emergent in modern legal orders depend on the presence of the legal staff, on the court system, the bureaucratically organised police etc. However, we need not and should not think exclusively in terms of the state and of government here. Rules and the (customary) "law of the land" may emerge as institutions that are not backed by the monopoly power of the state (see for instance the examples in (Klein, Daniel B., 1997) and the discussion in (Benson, Bruce 1990)). Still, there will always be the organizing small group structure that makes viable ordered large-scale interaction.

In all governance structures the organizational backbone of a small group structure will become visible at closer inspection. It is here that the traditional sense of the term "convention" comes up again. As mentioned before, traditionally the concept of a convention was more or less referring to the fact that something could conceivably be otherwise. The restriction to phenomena that could emerge as "consequences of human action but not of human design" which seems to characterize the modern notion of a convention is much more special than the classical concept. Of course, the classics of social theory were not unaware of the role of unintended consequences. Quite to the contrary, the classics of social theory made much of the insight that social order may emerge without deliberate enactment of the basic rules of such an order. Still they were not entirely clear about what kinds of constraints would be imposed on the type of order that could emerge spontaneously in an invisible hand process rather than being deliberately enacted by a hidden hand. The preceding analysis of conventions which puts them into the twin perspective of risk dominance and local interaction structures can add considerably to classical theory here and is of interest in itself for modern theory as well as modern social practice.

Of course, social conventions are often defined more literally. One may, for instance, refer to the social norm of distributive justice (requiring rewards to be proportionally shared according to individual contributions, see, for instance, Homans, 1961) as a convention. Such a norm may indeed be stable in the sense of a strict equilibrium if every deviation from it will be detected and harshly punished with sufficiently high probability. If that holds even other ways of distributing rewards, e.g. by sharing equally instead of proportionally, could be also justified as a strict equilibrium outcome. Altogether, the selection of the social norm how to split rewards would thus appear as a coordination as theoretically and experimentally discussed above.

A rigorous analysis of social norms as conventions would, however, to capture more explicitly of how norm deviations can be detected, who and why engages in punishing then when detected, what happens if one refrains from punishing norm deviators etc. The difficulties of such norm enforcing games became even more involved when considering the local interaction aspect as in our analysis above. This explains why we have concentrated on rather simple games with multiple strict and payoff ranked equilibria. We nevertheless hope that at least some insights will carry over to more complex game structures like those when trying to study the evolution of social norms and testing their reliability in experiments (see Roth, 1995, for a survey of some illustrative results).

## References


Benson, Bruce. 1990. *The Enterprise of Law. Justice Without the State*. San Francisco: Pacific Research Institute for Public Policy.

Berninghaus, S.K., Ehrhart, K.M. 2001, Information and Efficiency in Coordination Games: Recent Experimental Results, in: *Economics Essays*, G. Debreu, W. Neuefeind und W. Trockel eds, Springer Verlag, pp. 19-39.

Berninghaus, S.K., Ehrhart, K.M., und C. Keser 2002, Conventions and Local Interaction Structures: Experimental Evidence, erscheint in: *Games and Economic Behavior.*

Berninghaus, S.K., and U. Schwalbe 1996, Conventions, local interaction and automata networks. *Journal of Evolutionary Economics*, 6, pp. 313-324.

Binmore, Ken. 1994. *Game Theory and Social Contract Volume I - Playing Fair*. Cambridge, London: The MIT Press.

Boyer R., and A. Orléans 1992, How do conventions evolve? *Journal of Evolutionary Economics*, 2, pp. 165-177.

Cooper, R.W., D.V. De Jong, R. Forsythe, and T.W. Ross (1992), Forward induction in coordination games. *Economics Letters*, 40, pp. 167-172.

Harsanyi, John C. and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass.: MIT Press.

Hart, Herbert L. A. 1961. *The Concept of Law*. Oxford: Clarendon Press.

Heinimann, F. 1987/1945. *Nomos und Physis*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Homans, G.C. 1961, Social Behavior: Its Elementary Forms. *Wiley*, New York.

Keser, C., K.-M. Ehrhart, and S.K. Berninghaus 1998, Coordination and local interaction: experimental evidence. *Economics Letters*, 58, pp. 269-275.

Klein, Daniel B. ed. 1997. *Reputation*. Ann Arbor: The University of Michigan Press.

Kliemt, Hartmut. 1993. "Constitutional commitments," in *Jahrbuch für Neuere Politische Ökonomie*. Ph. et al. Herder Dorneich ed, pp. 145-73.

Lewis, David. 1969. *Convention*. Cambridge, Mass.: Harvard.

Parsons, Talcott. 1968. *Utilitarianism. Sociological Thought.* New York und London.

Raphael, D.-D. ed. 1969. *British Moralists*. Oxford: Oxford University Press.

Roth, A.E. 1995, Introduction to experimental economics. In: *The Handbook of Experimental Economics*, eds. J.K. Kagel and A.E. Roth. Princeton University Press, Princeton, pp. 3-109.

Schneider, L. ed. 1967. *The Scottish Moralists on Human Nature and Society.* Chicago und London.

Taylor, P.D. and L.B. Jonker. 1978. "Evolutionarily Strategies and Game Dynamics." *Mathematical Biosciences*, 40, pp. 145-56.

Van Huyck, J.B., R.C. Battalio, and R.O. Beil 1990, Tacit coordination Games, strategic uncertainty, and coordination failure. *American Economic Review*, 80, pp. 234-249.

Young, H. Peyton. 1993. "The Evolution of Conventions." *Econometrica*, 61:1, pp. 57-84.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure. An Evolutionary Theory of Institutions*. Princeton: Princeton University Press.