The Dynamics of Norms and Convention under Local Interactions and Imitation

Working Paper Version 3

SEBASTIAN ILLE

1st November 2011

This article shows that under certain conditions it is highly likely that individuals coordinate on a Pareto superior though risk inferior equilibrium. The tendency to generally favor risk over pay-off dominance as an equilibrium refinement criterion is thus criticised. Results contrast heavily with predictions made by the stochastic stability approach. Restricting the analysis to general 2x2 coordination games, two assumptions are made: Players do not play a best response strategy but simply imitate the most successful player from an individual reference group. In addition, individual pay-offs are also assumed to solely depend on the strategic decisions of a reference group. If these two reference groups are individually identical and small, a Pareto dominant though risk inferior convention will evolve for a large range of parameter values. If an individual interacts with more players and has a larger number of players to whom he compares, a risk dominant but Pareto inferior convention becomes more likely. This effect is enforced if the former reference group is larger than the latter. Simulations are used to support the analytical results.

Keywords: Existence and Stability of Equilibria; Evolutionary Games; Behavior; Simulation Modelling.
Subject Classification: C62, C63, D73, D03, D83.

1 Introduction

The inter-correlation between cultural and economic variables has already been emphasized by Max Weber in his œuvre "Wirtschaft und Gesellschaft" [2007], but also more recently, prominent scholars have stressed the necessity to give culture a proper recognition as an economic determinant [Harrison and Huntington 2000, Huntington and Harrison 2004, Huntington 1997, Welzel and Inglehart 1999, Ades and Di Tella 1996, Bollinger and Hofstede 1987]. "Economic reality is necessarily embedded within broader social relations, culture and institutions, and the real boundaries between the 'economy', and 'society' and 'polity' are fuzzy and unclear."

Electronic copy available at: http://ssrn.com/abstract=2275977

[Hodgson 1996, p. 8]. Agents take decisions subject to cultural constraints defined by the current social norms and conventions. Besides, these decisions define the basis for new norms and conventions, thus altering the rules for future interactions [Bicchieri 2006]. The dynamics of social conventions and norms are therefore of particular economic interest and the influential paper "The Evolution of Convention" by Young [1993], along with the prominent article by Kandori et al. [1993], is of special importance under this perspective. The approach allows to discriminate between the various Nash equilibria that can occur in a game, and defines the *stochastically stable state* (SSS). This equilibrium state determines the convention and hence the strategies that is adhered to by agents in the long-term.

This article will thus compare assumptions and results to this approach, and analyses the generalizability and robustness of the predictions made on its ground, i.e. it asks the question if the general results of the stochastic stability are maintained if some of the fundamental assumptions are changed to fit a more realistic context. Though the adaptive play assumption provides a more realistic setting of play than classic approaches by adding bounded memory, idiosyncratic play and random pairing, other assumptions are still overly restrictive for certain settings that constitute a basis for social interactions. Two major factors are considered.

Though this assumption is relaxed in Young [1998], and Durlauf and Young [2001], Young's framework assumes first that interactions are not local, but global. This is, however, a fairly unrealistic assumption for large and dispersed player populations or if individual perception of conventions is exclusively shaped by the interaction with a reference group (parents, family, friends, colleagues etc.). In contrast, here each player interacts solely with his neighbors on a spatial grid. Second, under more realistic conditions, the determination of a best-response strategy requires superior mental capabilities; even under the assumption of adaptive play. Therefore adaptive (best response) play is substituted by a heuristic that imitates the most successful player.

In these simplified networks with local interactions and imitation, a prevailing convention will not necessarily be defined by the *SSS*. The positive pay-off premium that players earn in a certain equilibrium with respect to the other equilibrium, will affect its likelihood to determine the long-term social convention. In the case of local interaction and imitation, a trade-off between risk and Pareto dominance can thus be observed. This relation is a result of an interesting property of these networks: In contrast to evolutionary games that suppose global interactions and best-response play, it is redundant to assume *a priori* the assortment among players with the same strategy. On the contrary, assortment is an immanent evolving property of this network structure. Consequently, dynamics will differ, since assortment will place more weight on the diagonal elements in the pay-off matrix. In this context, Pareto dominance can prevail over risk-dominance.

The first section considers a symmetric 2×2 coordination game. The second section will generalize the approach to two types of players. The third section will look at the two dimensions of space. In this model individuals have a space which they observe, i.e. an area that defines the set of players that can be imitated, and a space which affects their pay-offs, i.e. the number of other surrounding players that define the individual's pay-off for each strategy in accordance to his strategic choice. In the former sections both were restricted to the adjacent players. Under various conditions, however, both spatial dimensions do not necessarily coincide and can be of various sizes. The equilibrium, to which a population will converge, is affected by the relative sizes of both dimensions of space. The higher the size of the "pay-off space" with respect to the

Electronic copy available at: http://ssrn.com/abstract=2275977

"imitation space", the higher the probability that a population converges to the risk dominant equilibrium.

2 The Evolution in a Spatial Game

In Young's approach one player from each type class (e.g. the "Battle of Sexes" a *man* is paired with a *woman* at random) [1] is drawn at random to participate in the game, implying that the game can be played by any possible combination of players only restricted by the affiliation to a certain player type. The approach further assumes that individuals choose their best response by maximizing the expected pay-off given each available strategy in their strategy set, and the actions sampled from a history of past plays. It is expected that there is a positive probability to sample any of these past plays. The player population is fully connected and interactions are global. Yet, it is more reasonable to assume that most interactions are strictly local, both regarding the pairing and the sampling process (for a critique of *integral* frameworks refer to Potts [2000]).

Furthermore, the determination of a best-response strategy demands each individual to possess the mental capacity to evaluate the exact expected pay-off for each strategy in his strategy set, given an anticipated strategy profile. A player requires thus full knowledge of his and the other individual's strategy set, and his and the other's preferences (thus of the associated pay-offs), along with the assumption of an invariable state space. Obviously these are very strong assumptions for "large" worlds, in which conventions evolve. Individuals tend to choose a strategy based on simplifying heuristics (see Page 2007).[2] "Imitate the best action" is a suitable assumption for explaining strategical choice (a similar rule has been applied in Robson and Vega-Redondo 1996). As Bowles stated:"We know that individual behavioral traits may proliferate in a population when individuals copy successful neighbors. So too may distributive norms, linguistic conventions, or individual behaviors underpinning forms of governance or systems of property rights diffuse or disappear through the emulation of the characteristics of successful groups by members of less successful groups." [Bowles 2006, p.444]. According to this rule, each player adopts for his future play the strategy of the player with maximum pay-off in his reference group. For simplicity it is assumed here that the history of past play, which can be remembered, reduces to the last interaction period.

On the one hand, several articles have addressed the first issue to some degree [Young 1998; 2005, Morris 2000, Lee and Valentinyi 2000, Ellison 1993; 2000], but neglected the second by assuming local interactions with some form of (fictitious) best response play. One the other hand, the literature in evolutionary game theory, using replicator dynamics, has assumed the inverse, i.e. global interactions and strategic choice via imitation (though probabilistic). The approach, elaborated in this article, supposes *both* local interactions and imitation, and generates

[1] Young's assumption can be more generally interpreted as this is only being the case, if a strategy profile is defined by as many strategies as there are types. For example, if only a single type exists in a 2×2 Nash coordination game, obviously 2 players from the same type are drawn and a player can interact with anybody in his population.

[2] The question is whether choice can be *approximated* by rational play. Admittedly, in the model described here, players will be able to coordinate on one of the two equilibria under most conditions and will act like pay-off maximizers. Yet, the difference with respect to such a model is that here players are more likely to coordinate on the pay-off instead of risk dominant equilibrium.

different results than those obtained in the literature previously stated. Only if *both* assumptions (imitation and local interactions) apply, the Pareto dominant equilibrium, instead of the risk dominant equilibrium, will be selected by a population in the long-term for a broad parameter range. The basic dynamics will be illustrated in this section.

Assume the following:[3]

(a) All individuals interact on a toroidal, two dimensional grid, on which they are initially placed at random.

(b) Individuals only interact with their direct neighbors (Moore neighborhood).

(c) Individual pay-offs only depend on the individual's strategy and on the strategies played by his neighbors.

(d) Each individual adopts the strategy of the neighbor with maximum pay-off; if the individual already received not less than the maximum pay-off, he will keep his strategy.

(e) All players update synchronously and once in each period.

(f) Updating is deterministic (no mutations) and the outcome of the game is only defined by the initial conditions, distribution, and the pay-off matrix.

These assumptions will keep the analysis as simple as possible and will enable us to predict the population's evolution without the need to run explicit simulations. Yet, simulations are performed both to support results and to visualize dynamics.

2.1 Symmetric pay-offs

Let N be a finite population of individuals, in which each player is assigned to a unique individual position on a two-dimensional torus-shaped grid, defined by a coordination tuple (x,y) with $x,y \in \mathbf{N}$. Each individual only interacts with his Moore neighborhood. Let $\sim$ be a binary relation on N, such that $i \sim j$ means "i is neighbor of j". For an individual i in a patch with coordinates (x_i, y_i), an individual j, with $j \sim i$, is defined as $\left\{ j : (x_j = x_i + v, y_j = y_i + w) \right\}$, with $v,w = -1,0,1$ and $|v| + |w| \neq 0$. Consequently, it is assumed that the binary relation $\sim$ is irreflexive, symmetric and each player has 8 neighbors surrounding him. Define $\mathfrak{N}(i)$ as the set of neighbors of i, such that $\mathfrak{N}(i) \equiv \{ j : j \sim i \}$. Initially assume that the pay-offs for two strategies $s(i) = A,B$ of player i are given by a symmetric pay-off matrix with a single player type. Hence, it is irrelevant, whether an individual plays as row or column:

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} a,a & b,c \\ c,b & d,d \end{pmatrix} \end{array} \tag{1}$$

[3] This follows in general the assumptions made in other spatial models, such as Nowak and May [1992], Hauert [2001], and Brandt et al. [2003]. Yet to my knowledge coordination games have not been analyzed in this type of literature. Furthermore this literature relies generally on simulation and less on closed form solutions.

with $a > c$ and $d > b$. Define the equilibrium, in which all players choose strategy A, as h_A and the equilibrium, in which all choose B, as h_B. Let A_t be the set of individuals playing strategy A in period t, and B_t the set of individuals playing B in the same period. Further, let $\Theta_t^A(i) = \#\{A_t \cap \mathfrak{N}(i)\}$ and $\Theta_t^B(i) = \#\{B_t \cap \mathfrak{N}(i)\} = 8 - \Theta_t^A(i)$ be the number of strategy A and B playing neighbors of i. The pay-off of player i at time t is thus defined as

$$\pi_t(i) = \begin{cases} \Theta_t^A(i)a + \Theta_t^B(i)b, & \text{if } s_t(i) = A \\ \Theta_t^A(i)c + \Theta_t^B(i)d, & \text{if } s_t(i) = B \end{cases} \tag{2}$$

Let $\Pi_t(i) = \{\pi_t(j) : j \in \mathfrak{N}(i) \cup \{i\}\}$ be the joint set of player i's and his neighbors' pay-offs. Define $\arg(i) = \{s_t(j) | j \in \mathfrak{N}(i) \cup \{i\}, \pi_t(j) = \max \Pi_t(i)\}$. For the following period, this player chooses a strategy s_{t+1} based on an imitation rule.

$$s_{t+1}(i) = \begin{cases} A, & \text{if } \arg(i) = \{A\} \\ B, & \text{if } \arg(i) = \{B\} \\ s_t(i), & \text{if } \arg(i) = \{A, B\} \end{cases} \tag{3}$$

Though the following analysis is local, it enables us to predict the global evolution based on the given pay-off configuration. The Pareto dominant strategy is defined by the Pareto dominant equilibrium. The following results are a direct consequence:

(a) In the case, where a player chooses the Pareto dominant strategy, i.e. the strategy defined by the largest value on the pay-off matrix's main diagonal, his pay-off increases with the number of neighboring players choosing the same strategy. The maximum pay-off for this strategy is obtained by individuals only surrounded by players of the same strategy. This also holds for the Pareto inferior strategy, if the matrix's main diagonal pay-off values are strictly greater than the off-diagonal values.

(b) Any interior individual, only surrounded by players of the same strategy, has never an incentive to switch, since all players in his neighborhood play the same strategy. Transitions can only occur at borders of clusters.

(c) If an individual, which is completely surrounded by players of his own strategy, plays the Pareto dominant strategy, pay-off is maximal and none of his neighbors will switch to the Pareto inferior strategy.

In order that two equilibria are risk equivalent it must hold that $a - c = d - b$. The pay-off can thus be written as $d = a + \rho$ and $b = c + \rho$. Define the pay-off difference ρ as the "pay-off premium". For $\rho > 0$ equilibrium h_B Pareto dominates h_A.

Definition 1. *A* ***cluster of size*** *r is defined by the highest number of neighbors playing the same strategy in the set of directly connected players with identical strategy, i.e. a cluster of size r is defined as a set of connected players, in which at least one player has r-1 other players with the same strategy in his neighborhood.*

For example, suppose a straight line of players where each player has a neighbor to his left and right with identical strategy, except for the corner elements. Such a straight line will always have a size of 3, since by definition each player has at least one neighbor with 2 identical players, i.e. those, who choose the same strategy, in his neighborhood. Hence, all players in this cluster will compare any player with a different strategy either to $2a+6b$ or to $2d+6c$.[4] The reason for this definition is that the dynamics do not depend on the number of connected players with the same strategy, but on the element with the highest pay-off in the neighborhood. Consequently, the length of such a line of identical players is unimportant. The same reasoning holds for larger clusters. In addition, this definition restricts cluster size to a maximum value of 9, since a player can only have a neighboring player with a maximum of 8 identical neighbors . Based on these assumptions the following proposition is obtained (all proofs are to be found in 6):

Proposition 1. *Given a pay-off matrix as in 1 with two risk equivalent pure Nash equilibria, for any a,b,c,d as long as they satisfy $a-c=d-b$, a population, whose convention is defined by the Pareto inferior strategy A, is successfully invaded by a minimum cluster of size r, choosing the Pareto dominant strategy B, if the pay-off premium satisfies:*

$$\left.\begin{array}{l}\rho > 3(a-c) \text{ and } r \geq 4 \text{ and square} \\ \rho > \quad a-c \ \text{ and } r \geq 5\end{array}\right\} \quad \text{for } a<b$$

$$\rho > \frac{3}{5}(a-c) \text{ and } r \geq 6 \qquad\qquad \text{for } a \geq b$$

If $\rho < \frac{1}{2}(a-c)$ or $\rho < \frac{1}{5}(a-c)$, the population will return to the incumbent convention if the invading cluster is of size 6 or 7, respectively. If none of the conditions is fulfilled, clusters of size 6 and 7 are stable but cannot invade the population.

Hence, a minimum pay-off premium $\rho > \frac{3}{5}(a-c)$ is sufficient for a population to converge to the Pareto dominant strategy of the symmetric coordination game, if a cluster of size 6 or greater can evolve with positive probability. The figure 8 A.) in the appendix shows the result of a set of simulations for $d=4$, $b=3$, $c=a-1$ and a going from 3.35 to 3.55 in steps of 0.01. The values represent the percent of individuals playing strategy A in $t \in (1,50)$, where their initial share is set to 85% in $t=0$. (Remember that initial seeding is random. An initial share of 15% B players generates a cluster of size 6 with positive probability for the given population size.) Thresholds are at their expected values. The population converges to equilibrium h_B for values of a smaller than 3.4 and to equilibrium h_A for values larger than 3.5. Stable mixed equilibria occur for intermediate values (here: one at less than 0.8%, a second at 0.4% strategy B players).

Proposition 2. *Clustering is an evolving property and most clusters of at least one strategy will have a size equal to 9 after an initial period of interaction. In addition, for $b>a$ and $\rho > 7(a-c)$, stable clusters of size $r=1$ can occur, which play the Pareto inferior strategy A. Given the case of $a>b$: Clusters, playing the Pareto dominant strategy B, of size 6 can be stable, if $\frac{1}{2}(a-c) < \rho < \frac{3}{5}a-c$, of size 7, if $\frac{1}{5}(a-c) < \rho < \frac{1}{3}(a-c)$, and of size 8, if $0 < \rho < \frac{1}{7}(a-c)$. Cluster of size 5 are stable iff $a=b$.*

[4]This holds also for the outer elements of the line cluster, if $a,d > b,c$. If this is not the case, these outer element have highest pay-off with $a+7b$ or $7c+d$, if strategy A or B is Pareto inferior.

Consequently, a minimum pay-off premium for the Pareto dominant strategy exist, which is necessary to take over the player population and to determine the long-term stable convention. Since a completely random initial distribution is unstable, clusters collapse, and some or all clusters will attain a size of 9 for at least one strategy. Whether this is the case for one or both strategies depends not only on the parameter values, but also on the initial distribution.

Definition 2. ***A homogeneous initial distribution*** *is defined as a distribution of a player population, in which the average cluster size for all strategies is identical after the first period of interaction. A* ***heterogeneous initial distribution*** *is defined by an initial player population, where average cluster size differ strongly among strategies after the first period of interaction, but the evolution of at least one cluster of size 6 or greater occurs with certainty for any strategy after an initial sequence of interactions.*

Hence, a homogeneous initial distribution defines for example the case, in which all players initially chose one of both strategies completely at random, under the condition that average pay-off for both strategies are sufficiently similar, or alternatively, in which larger agglomerations of players choosing the same strategy exist *a priori* for *both* strategies. In this case, clusters of size 9 exist for both strategies after the initial interaction sequence. A heterogeneous initial distribution can define a situation, in which the entire player population chooses the same strategy except for one miminum mutant cluster of size 6. Alternatively, it describes a situation in which players choose their strategy at random, but average pay-offs are very different, so that the player population collapses into large clusters playing the risk dominant strategy and small clusters with at least one being of size 6 playing the other strategy. Both distributions define the possible extreme cases that will define the boundary conditions for the evolution of a stable Pareto optimal convention.

By proposition 1, for a minimum ρ, minimum cluster size for an invasion is 6. The minimum necessary pay-off premium for highly heterogeneous distributed societies is thus given by $\frac{3}{5}(a-c)$. If this condition holds, the population will be invaded by the Pareto dominant strategy even if there is a sparse distribution of individuals playing the Pareto superior strategy. Also for homogeneously distributed societies, a minimum necessary pay-off premium can be derived. Hence, depending on the degree of homogeneity of the initial distribution, both values define the upper and lower boundary conditions for the lower limit of ρ such that the Pareto dominant equilibrium defines the convention.

Proposition 3. *A population with homogeneous initial distribution will converge to the Pareto superior equilibrium h_B, if the pay-off premium ρ of the Pareto dominant equilibrium is greater than $\frac{1}{7}(a-c)$. If the pay-off premium is smaller, but positive, a player population will consist of clusters playing different strategies.*

Figure 8 B.) shows the result of a set of simulations identical to those in figure 8 A.), but $d=4$, $b=2$ $c=a-2$ and a ranging from 3.6 to 4.4. Furthermore each strategy is initially played by 50% of the population and seeding is completely random. Thresholds are again as expected. The population converges to equilibrium h_B for a smaller than 3.6 and to h_A for values larger or equal to 4.3. The population thus converges to the Pareto optimal convention, except if the pay-off premium is within a marginal perceptible unit.[5]

Based on these results, the question arises of how strongly risk dominance is offset by Pareto dominance. For this assume as before that $d = a + \rho$ and $b = c + \rho$, but also substitute c in matrix 1 by c^*, such that $c^* = c - \mu$. Consequently h_B is Pareto dominant by a value of ρ and h_A is risk dominant by μ. Define this value as the "risk premium". Again, the two cases of initial distribution define the boundary trade-off conditions between risk and Pareto dominance.

Proposition 4. *Given a coordination game as in matrix* (1) *with two equilibria of which h_B pay-off dominates h_A by a pay-off premium of ρ, and h_A risk dominates h_B by a risk premium of μ, the population converges to convention h_B if*

$$\mu < \begin{cases} c - a + 7\rho, & \text{and the initial distribution is homogeneous} \\ c - a + \frac{5}{3}\rho, & \text{and the initial distribution is at least heterogeneous} \end{cases}$$

If the initial population distribution is heterogeneous and $\mu > \frac{2(c-a)+4\rho}{3}$, the population chooses the risk dominant convention. In the case that a population is, however, initially sufficiently homogeneous, the risk dominant strategy only prevails as a convention, if it also pay-off dominates by a value greater than $\frac{a-c}{7}$. Otherwise the population remains in a state of mixed conventions.

For symmetric 2×2 coordination games, this contrasts with Young's original approach. The approach based on imitation does not predict the convention to be solely defined by risk dominance, but dynamics follow a trade-off between risk and Pareto dominance. Furthermore, if the population is initially sufficiently homogeneously distributed, we will observe a Pareto dominant convention basically all of the time. This stems from clustering owing to local interactions and the subsequent emphasize on the matrix's main diagonal's pay-offs.

Figure 8 C.) shows the result of a set of simulations identical to those in figure 8 A.), i.e. given a heterogeneous initial distribution, but c^* is fixed at $c^* = 2$ and a ranges from 3.0 to 3.3. The population converges to the Pareto dominant equilibrium for values of a smaller and equal to 3.24 and converges to the risk dominant equilibrium for $a > 3.29$. For $\frac{1}{4} < \mu < \frac{2}{7}$, the population converges to a mixed equilibrium, with a few square shaped cluster of size 6 or 9 that play the Pareto dominant strategy.

Figure 8 D.) presents the result of a set of simulations identical to those in figure 8 B.), i.e. given a homogeneous initial distribution, but $c^* = 1$ and a ranges from 3.0 to 4.0. The population converges to the Pareto dominant equilibrium for values of a smaller and equal to 3.6, and remains in a mixed equilibrium for larger values.

Figure 8 E.) shows the distribution of the player population for $t \in (0;100)$ and a changing pay-off premium ρ. The pay-off structure is given by $b = 6$, $d = 10$, $c^* = a - 4$, and a takes values from 0 to 20 in unit steps. The effect of the size of ρ on convergence speed towards a single equilibrium is negligible. The same holds for the updating probability.

Figure 8 F.) shows the distribution for pay-offs $a = 3, b = 2, c^* = 1$ and $d = 4$ and different updating probabilities. The updating probability ranges from 1 to 100.

[5] $\frac{1}{7}(a-c)$ defines the marginal perceptible unit, under which no pure equilibrium will occur. Furthermore, the simulation shows how the distribution is affected by the relative average pay-off for a small number of periods, but stabilizes after the initial interaction sequence.

3 General 2 x 2 Coordination Game

The following section analyses the dynamics of general 2×2 coordination games, *ceteris paribus*, in which two player types (row and column) interact with each other. Players choose their initial strategy at random. Whether the homogeneous or heterogeneous case applies, depends on the relative average pay-off of each strategy. The only difference with respect to the assumptions in the previous section is that on each patch two players coexist, one of each type. The general pay-off structure is defined by the pay-off matrix:

$$\begin{array}{ccc} {}^{Type2}_{Type1} & A & B \\ A & \multirow{2}{*}{} a_1, a_2 & b_1, c_2^* \\ B & c_1^*, b_2 & d_1, d_2 \end{array} \tag{4}$$

Define as before $\rho_i = b_i - c_i$ or $\rho_i = d_i - a_i$ and hence $c_i^* = c_i - \mu_i$, for $i = 1, 2$. It must also hold that $a_i > c_i^*$ and $d_i > b_i$.

For two player types x and y with $x, y = 1, 2$ and $x \neq y$, define for a player i of type x the set of neighbors of his own type as $\mathfrak{N}_x(i)$ and the set of neighbors of the other type as $\mathfrak{N}_y(i)$. Further, let $A_{t,y}$ define the set of players of type y that play strategy A in period t, and accordingly $B_{t,y}$ as the set of players of type y playing B. Correspondingly, define $\Theta^A_{t,y}(i) = \#\{A_{t,y} \cap \mathfrak{N}_y(i)\}$ and $\Theta^B_{t,y}(i) = \#\{B_{t,y} \cap \mathfrak{N}_y(i)\}$ as the number of strategy A and B playing neighbors of i that are of type y. The pay-off of player i in time t is[6]

$$\pi_{t,x}(i) = \begin{cases} \Theta^A_{t,y}(i)a + \Theta^B_{t,y}(i)b, & \text{if } s_t(i) = A \\ \Theta^A_{t,y}(i)c^* + \Theta^B_{t,y}(i)d, & \text{if } s_t(i) = B \end{cases} \tag{5}$$

Analog to the former section, let us also define $\Pi_{t,x}(i) = \{\pi_{t,x}(j) : j \in \mathfrak{N}_x(i) \cup \{i\}\}$ as the joint set of agent i's pay-off and the pay-offs of his neighbors of the same type, and $\arg(i) = \{s_t(j) | j \in \mathfrak{N}_x(i) \cup \{i\}, \pi_t(j) = \max \Pi_{t,x}(i)\}$. The imitation rule is then determined by condition 3 on page 5 in the former section.

The complexity of the analysis increases through the augmentation of possible parameter combinations and the interdependence of the two player types' strategy choices. Yet, the fundamental dynamics are defined by only a few conditions similar to what has been obtained for the single type case. Only 2×6 conditions have to be analyzed in the general game. To derive this and as a first step, remember that the conditions for the pay-off structure, in addition to the imitation principle and the local interaction generate three useful characteristics.

(a) In general, the player choosing the Pareto dominant strategy with respect to his type benefits from the relative abundance of players in his neighborhood that belong to the *other* player type and are playing the same strategy. If $a_i > b_i$ and $d_i > c_i^*$, $\forall i = 1, 2$, this also holds for the strategy that is not Pareto dominant.

[6] Notice that following the former assumptions and for reasons of symmetry, a player does not interact with the other player type on his own patch. He thus has still 8 neighbors.

(b) A strategy change will only occur at the edges of clusters. This property must also hold for mixed clusters, i.e. where both player types play different strategies on the same patches in a neighborhood.

(c) During the initial sequence of interaction, the strategy distribution on the grid will be heavily determined by the relative average pay-off of each strategy, since it is more likely that, during this process, a player adopts the strategy that has higher average pay-off if players initially choose their strategy at random with equal probability.

The following proposition greatly simplifies the future analysis:

Proposition 5. *In a general 2×2 coordination game with asymmetric pay-offs and all other conditions being as before, the strategy distributions of the two player types will coincide after a brief sequence of interactions.*

In order to better understand the idea behind the proof, consider the case in which for both player types the same strategy has higher average pay-off. After the first period of interaction, larger clusters appear, playing this strategy. They surround smaller clusters that are either mixed, i.e. where each type on one patch plays a different strategy, or those that are uniform and play the strategy with lower average pay-off. The reason why mixed clusters will not sustain is the following: The players on the edge of the mixed cluster will always choose the strategy of the pure cluster in their neighborhood. The schematic in figure 1 helps to understand the underlying dynamic. It shows two clusters, each with two layers symbolizing the inter-acting player types. The pure cluster is stylized on the left side of the figure, the mixed cluster on the right. The higher layer illustrates type 1, the lower type 2. Assume without loss of generality and with respect to the mixed cluster that type 1 plays strategy A and type 2 plays B, and the pure cluster only plays A. By assumption, type 1 players choose the same strategy in both clusters. Since imitation is horizontal, i.e. between players of the same type, this type cannot imitate any other strategy. Consequently, players of type 2 will only interact with players choosing A. Since by definition $a_i > c_i^*$ and $d_i > b_i$, a player of type 2 in the pure cluster has always higher pay-off than a player of the same type in the mixed cluster. Hence, type 2 players at the edges of the mixed cluster will switch to strategy A. The dynamics are independent of what type plays which strategy and whether the strategy is risk or Pareto dominant (since the assumptions have ruled out strictly dominant strategies). Mixed clusters therefore vanish and only small strings of mixed

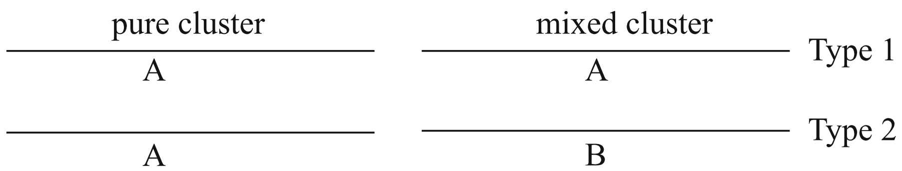

Figure 1: Dynamics of mixed clusters.

clusters of a maximum width of 3 at the borders of uniform clusters remain. (i.e. the *external* of each cluster and the player in between).

Similarly, in the case, where the strategy with higher average pay-off is different for both player types, larger mixed clusters will surround uniform clusters, but any mixed cluster will vanish in the subsequent periods. Hence, the strategy distribution on the grid will coincide for both player types after an initial sequence of interactions, though transition to this state will be faster in the first case than in the second. The following figure 2 illustrates this behavior for two cases with identical initial distribution and the subsequent 3 periods of interaction.

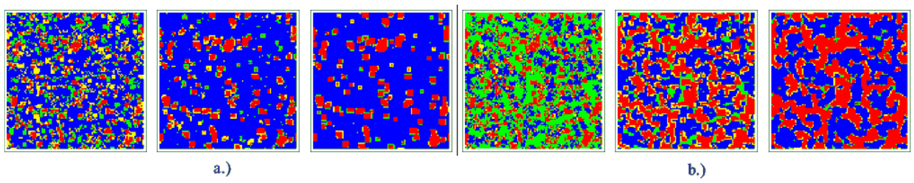

Figure 2: Strategic distribution for 3 periods – a.) $a_1 = a_2 = 6$, $b_1 = b_2 = 6$, $c_1^* = c_2^* = 0$, $d_1 = d_2 = 8$; b.) $a_1 = d_2 = 6$, $b_1 = c_2^* = 6$, $c_1^* = b_2 = 0$, $d_1 = a_2 = 8$; using colour coding: blue: $s_i = A$, red: $s_i = B$, green: $s_1 = A, s_2 = B$, yellow: $s_1 = B, s_2 = A$.

This leads to the following proposition

Proposition 6. *Given a pay-off matrix as in matrix* (4) *and* $a_i, d_i > b_i, c_i^*$, *for each type* $i = 1, 2$, *the convergence speed of the player population towards equilibrium* h_A *is determined by the largest integer* $\lceil \eta_A \rceil_i$ *less than* $\frac{-8\rho_i}{a_i - c_i - \rho_i}$. *Equivalently the largest integer* $\lceil \eta_B \rceil_i$ *less than* $\frac{8\rho_i}{a_i - c_i + \rho_i + \mu_i}$ *defines the convergence speed to equilibrium* h_B.

If the population is initially sufficiently homogeneous, the population converges to h_B *if* $\max_i \{\lceil \eta_A \rceil_i\} < \max_i \{\lceil \eta_B \rceil_i\}$ *and to* h_A *if* $\max_i \{\lceil \eta_A \rceil_i\} > \max_i \{\lceil \eta_B \rceil_i\}$. *Otherwise both strategies persist in the long-term.*

Since the strategies for both player types are congruently distributed after an initial sequence of interactions, most type specific complexities are eliminated. Simply speaking, subsequent to a transition period, large uniform clusters with cluster size 9 occur after an initial sequence of interaction. Each of these clusters pushes towards a convention and the one pushing strongest will eventually prevail. This is, however, only the case if average pay-off for both strategies are not too different and the distribution is initially sufficiently homogeneous.[7] In general, the edges of clusters will be either horizontal, vertical or diagonal. In such a case, the clusters' edges can thus be generalized to the following three types:

From this figure we observe that, though various parameter combinations could lead to strategic changes, only six conditions for each strategy influence the dynamics of the entire population in the long-term:

[7]Remember that even if players choose a strategy at random with equal probability before the first interaction, too diverse average pay-offs will inhibit the evolution of clusters, constituted by players of the risk inferior strategy, that have sufficient size to overtake the player population.

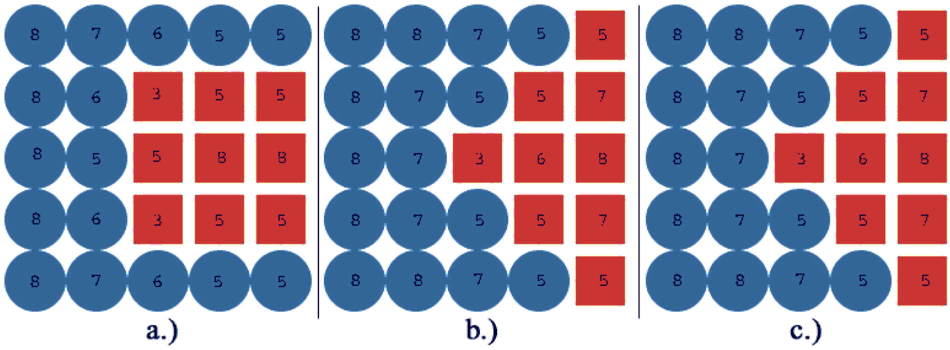

Figure 3: The three variants of cluster edges – numbers indicate the number of players with the same strategy in the individuals neighborhood, cluster are supposed to continue beyond the figure's frame .

$$\begin{array}{lll|ll}
 & \multicolumn{2}{l|}{\text{For } s_i = A \text{ to overtake } s_i = B} & \multicolumn{2}{l}{\text{For } s_i = B \text{ to overtake } s_i = A} \\
\text{I.} & \eta_A = 1: & 7a_i + b_i > 8d_i & \eta_B = 1: & 7d_i + \ c_i^* > 8a_i \\
\text{II.} & \eta_A = 2: & 6a_i + 2b_i > 8d_i & \eta_B = 2: & 6d_i + 2c_i^* > 8a_i \\
\text{III.} & \eta_A = 3: & 5a_i + 3b_i > 8d_i & \eta_B = 3: & 5d_i + 3c_i^* > 8a_i \\
\text{IV.} & \eta_A = 4: & 4a_i + 4b_i > 8d_i & \eta_B = 4: & 4d_i + 4c_i^* > 8a_i \\
\text{V.} & \eta_A = 5: & 3a_i + 5b_i > 8d_i & \eta_B = 5: & 3d_i + 5c_i^* > 8a_i \\
\text{VI.} & \eta_A = 6: & 2a_i + 6b_i > 8d_i & \eta_B = 6: & 2d_i + 6c_i^* > 8a_i
\end{array} \qquad (6)$$

for $i = 1, 2$. These conditions are equivalent to the results in proposition 6, and the condition I. and III. are identical to those found in proposition 4 for the single type case. The first condition implies that clusters can be overtaken along diagonal edges. It will turn an inlying cluster (red) as in a.) into an inlying cluster as in b.). If both conditions in I. are fulfilled, i.e. one strategy for each player type, these corner elements will continuously switch between strategies. Condition III. applies to the horizontal and vertical cluster's edges. It will not affect the players surrounding the corner elements of inlying clusters (see b.) and c.)). Hence, an inlying quadrangular cluster will expand and will incrementally turn the horizontal or vertical edge into a diagonal edge. Since condition III. includes condition I., the cluster will also continue to expand along these diagonal edges. Furthermore, under condition III. any inlying cluster can be invaded. The remaining conditions have a minor effect on the convergence speed than the aforementioned. Condition IV. and greater only concern the growth along the corner elements.

The easiest asymmetric coordination game is a game of "common interest". This denotes a pay-off structure, in which the same strategy is Pareto dominant for both types. According to the former notation, either $a_i > d_i$ for both $i = 1$ and $i = 2$ (or the inverse). If at least one player type fulfills at least the first condition given by $\eta_A = 1$ ($\eta_B = 1$), the population either converges to the convention defined by h_A (h_B), or ends up in a mixed equilibrium with interior rectangular shaped clusters playing A that cannot expand (see figure 3.A.) on page 12). As we have seen, this depends on the initial random distribution and the average pay-off of each strategy. If at least one player type meets condition $\eta_A = 3$ ($\eta_B = 3$), the convention is defined by h_A (h_B) given that a minimum invading cluster evolves. Since both player types either prefer equilibrium h_A or equilibrium h_B, the convergence by one player type towards an equilibrium is not counteracted by the other player type. Though convergence speed is irrelevant for the final distribution and convention, the more conditions are fulfilled by one or both types, the faster the population converges to its Pareto dominant equilibrium.

If the individuals find themselves in a "conflict game", in which the Pareto dominant equilibria are not identical for both player types, the convention is defined by proposition 6. The strategy that fulfills the higher condition in equation 6 will define the convention. In the following I will simulate a player population to confirm that population dynamics behave according to the previous results.

The most convenient way to test for the correctness of these results is to fix pay-off parameters of one player type at the different levels at which the constraints in equation 6 can be fulfilled, e.g. ranging from none to all six. The dynamics for each parameter of the other player type are then simulated with respect to each of these levels. As a basis for the analysis assume the following pay-off matrix:

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \quad A \qquad\qquad\qquad\qquad B \\ \begin{pmatrix} a_1 = 3.5,\ a_2 = 3.5 & b_1 = 3,\ c_2^* = 2 \\ c_1^* = 2,\ b_2 = 3 & d_1 = 4,\ d_2 = () \end{pmatrix} \end{array} \tag{7}$$

In order to account for the various variables, at which the conditions can be fulfilled, d_2 is set to one of 6 different values, namely $d_2 = 3.16; 3.22; 3.28; 3.34; 3.41; 3.47$ in each simulation run. This implies values of convergence speed given by $\eta_A = 5; 4; 3; 2; 1; 0$ and that player type 2 converges to h_A except for $d_2 = 3.47$ and thus $\eta_A = 0$.

For each of these values, one parameter of player type 1 is analyzed by a set of simulations. In the first set of simulations the parameters of player type 2 are set to the values in matrix 7 and $d_2 = 3.16$. For the first simulation of this set, one parameter of player type 1 is fixed at its lowest value, at which it does not fulfill any condition (or all, depending on the parameter) of the first column in 6. His other pay-off parameters are set to the values as in matrix 7 (if not stated otherwise). The initial distribution is set to 50 : 50 (if not stated otherwise), with completely random seeding. After the system has been simulated for a fixed number of periods, the parameter of player type 1 is changed by an increment and the system is simulated again. Using the same initial distribution renders the results directly comparable. Simulations are repeated until the parameter reached a maximum value, at which all condition (or none) of the first column in 6 are fulfilled. Hence, player type 1 will progressively (or regressively) converge to equilibrium h_B, whereas player type 2 will converge to equilibrium h_A at the speed determined by the value of d_2. After the value of player 1's parameter has reached its maximum value, the set of simulations is repeated for each of the remaining values of d_2. At the beginning of each set of simulations, the population is "seeded" anew. Each remaining parameter of player type 1 is analyzed in the same way, obtaining 6 sets of simulations for each parameter of player type 1, thus 24 sets of simulations in total. The figures show the proportion of type 1 players choosing strategy A. Since the distribution for both types concurs after the initial periods, it suffices to graph only one player type, as before.

Figure 9 on page 25 shows the result for parameter a_1. In order to maintain the assumption that $a_i > b_i$, the value of b_1 is adapted accordingly and set to $b_1 = 2.3$. This change is made only for the simulations concerning a_1. a_1 takes value from 2.375 to 3.875 in increments of 0.25. If both types have the same convergence speed, the population converges to a mixed equilibrium, where the strategy distribution is determined during the initial periods of interaction, i.e. by the average pay-off and the random initial distribution. In order to compensate this effect (since for $\rho_A = \rho_B$, a_1 is relatively large in the later simulations), the initial distribution was set to 55% strategy B players in the last 4 simulations. The predicted threshold values from proposition 6 for the parameters are as shown in table 1.

Table 1: Convergence speed for each player type in the simulations

$\eta_A < \eta_B \rightarrow h_B$, $\eta_A > \eta_B \rightarrow h_A$							
η_A	0	1	2	3	4	5	
d_2	3.47	3.41	3.34	3.28	3.22	3.16	
η_B	0	1	2	3	4	5	6
$a_1 <$	4	3.75	3.5	3.25	3	2.75	2.5
$c_1 >$		0	2	$\frac{8}{3}$	3	3.2	$\frac{10}{3}$
$d_1 >$	3.5	$\frac{26}{7}$	4	4.4	5	6	8

Figure 10 on page 25 provide the results for parameter b_1 taking value from -4 to 3 in increments of 1. Looking at 6 shows that b_1 is extraneous in the constrains of the right column. b_1 only affects the average pay-off and thus the number of strategy A players after the initial sequence of interactions. It thereby has an indirect impact on the time required to converge to an equilibrium, since it modifies the initial sequence after which the convergence to the convention occurs. $\eta_B = 1$ satisfies $8a_1 = 2c_1^* + 6d_1$ only as equality and it only holds that $8a_1 < c_1^* + 7d_1$. Consequently, convergence to h_B is only observable for $d_2 = 3.47$. For $d = 3.41$ the convergence speed for both player types is identical and thus stable mixed equilibria occur. In the case of $d = 3.34$ convergence to h_A is slow since condition II. is met with equality by player type 2. Similarly the simulation for $b_1 = 3$ and $d = 3.47$ approaches h_B slowly as the number of strategy A players is high after the initial sequence of interactions.

c_1^* adopts value from -0.1 to 3.5 in increments of 0.2 in the simulations in figure 11 on page 26. Since c_1 can take small critical values, simulations are conducted with an initial distribution of 55% strategy B players in the first two simulations and 60% strategy B players in the later simulation, in order to avoid that cluster size of strategy B is too small after the initial sequence of interactions. The threshold values for this parameter are again to be found in table 1.

The final set of figures 12 on page 26 shows the dynamics for d_1. d_1 takes values from 3.2 to 8.48 in increments of 0.33. In order to compensate for the "average pay-off effect" the share of initial strategy B player was set to 60% in the last 4 simulations. All simulations behaved according to the predictions made in table 1.

By proposition 6, the population can converge to different equilibria, though the level of risk and Pareto dominance are equal. Figure 4shows the share of strategy A players for a set of simulations, given $b_1 = 2$, $c_1^* = 0$, $a_2 = 5$, $b_2 = 0$, $c_2^* = 2$ and $d_2 = 4$. $a_1 \in (2.0, 9.0; 0.1)$ and

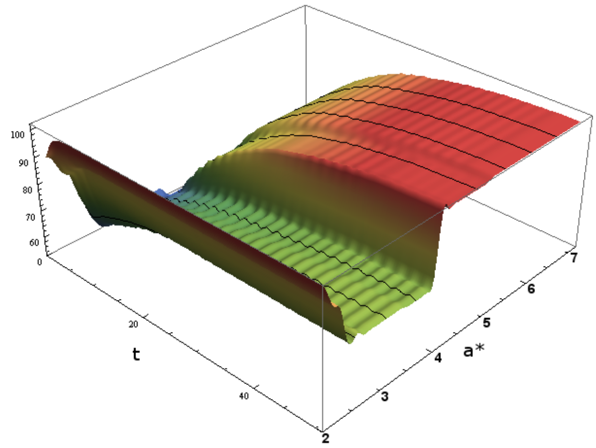

Figure 4: Percentage of strategy A players; Set of simulations with $a_1 \in (2.0, 7.0; 0.1)$ and $d_1 = a_1 + 1$: $|\rho| = 1$ and $|\mu| = 1$.

$d_1 = a_1 + 1$ and $t \in (1, 50)$. There are 70 simulations, for all $|\rho| = 1$ and $|\mu| = 1$. Strategy A is risk and strategy B Pareto superior for type 1, and the inverse for type 2. Player 2 moves to h_A with a speed $\eta_A = 1$. h_A is convention for $a_1 > 7$, and h_B is convention for $a_1 < 3$, for values in between no distinct convention evolves. Note that by proposition 6 and equations 6, the conditions, determining the dynamics, are unaffected by linear transformations of the pay-offs. In addition, such a transformation will have no effect on the relative average pay-off.

4 The Effect of Space - Planting Late in Palanpur

So far it has been assumed that only the 8 surrounding neighbors are considered both for the calculation of pay-offs and for imitation. There are several questions, which can be raised: Do the derived properties still hold, if the space (representing the reference group), which affects an individual's decision, increases? What happens, if the space (reference group) considered for the calculation of pay-offs differs from the one used for imitation? Though the analysis of the former two sections can be expanded to larger spaces, this will not be done in the scope of this article. This section will only provide general results without defining the clear conditions for each size of space.

Consider the following illustrative example taken from Bowles [2006]. In Palanpur, a backward village in India, peasants use to sow their crops at a later date than would be maximizing their expected yields. This results from a coordination failure. If a single farmer decided to sow early, seeds would be quickly eaten by birds and the harvest would be lost. The more farmers should agree to plant early, the less the loss for an individual "early seeder", since seeds lost to birds would be shared by the entire group of "early planters". Mechanism design, contracting, and implementation theory has dedicated much work to answer, which institutions are necessary to achieve the desired shift towards the Pareto optimal equilibrium. The question of why a population ended up in a Pareto inferior equilibrium is, however, more interesting in this context.

The above analysis has shown that in the case of imitation and local interactions a population converges to the Pareto dominant equilibrium, if the Pareto dominant strategy is played by a minimal number of individuals. Yet, Palanpur is a case, in which a population got trapped at the risk dominant equilibrium. It can be assumed that initially both strategies (*plant early, plant late*) have been played with strictly positive probability. Thus, it seems that the standard solution of this stag-hunt game (i.e. the 50:50 distribution lies in the basin of attraction of the risk dominant equilibrium) is more appropriate in this case than the answer I have presented so far in subsection 2.1. This conclusion is, however, precipitative. To illustrate this, assume the following pay-off matrix (also taken from Bowles 2006) for the game:

$$\begin{array}{c} \\ \textbf{Early} \\ \textbf{Late} \end{array} \begin{array}{c} \begin{array}{cc} \textbf{Early} & \textbf{Late} \end{array} \\ \begin{pmatrix} 4,4 & 0,3 \\ 3,0 & 2,2 \end{pmatrix} \end{array} \tag{8}$$

A first answer is that in Palanpur, the interacting population is very small. As a consequence, a low, but positive probability exists that the initial distribution of early and late planters was such that a stable cluster of early planters could not evolve and take over the whole population. Under random initial choice, it is much more likely that the Pareto dominant convention or a mixed stable equilibrium evolves than the risk dominant equilibrium in pure strategies.[8] Hence, a possible but less convincing answer is that peasants in Palanpur were simply unlucky.

A second answer is that peasants do not only consider neighbors at a distance of 1, i.e. their 8 surrounding neighbors on the grid, but have a much larger space, which defines the peasants they are interacting with.

Definition 3. *The **imitation radius** is defined by the largest Chebyshev distance between a player and a member of the set of observable neighbors who he can imitate. The **pay-off radius** is similarly defined as the largest Chebyshev distance between a player and a member of the set of neighbors that affect his pay-off.*

The radii thus define the minimum number of steps a “king” requires to move from the player to his farthest neighbor (in the set of the observable or pay-off affecting players) on the “chess board” grid. In the former sections both radii have been assumed equal to 1, i.e. an individual has only considered the adjacent 8 players. As the radii increase, it is observed that the population converges more likely to the risk dominant equilibrium. This is caused by a rapid decline of those individuals playing the Pareto dominant strategy in the first periods. Figure 5.A.) illustrates this behavior for a“ large” population (10.000 individuals). If the initial fall in the number of Pareto dominant players is higher than a certain threshold, the population will not converge to the Pareto optimal equilibrium, since a cluster of minimum size does not evolve. As the pay-off radius increases, pay-offs converge to the expected pay-offs in the first period in a random distributed population, and less weight is placed on the diagonal elements of the pay-off matrix. Since by definition, the risk dominant strategy has a higher average pay-off than the Pareto dominant, though risk inferior strategy, individuals in homogeneously distributed areas, i.e. where

[8] In addition, a small mutation rate seems also plausible in that society, and an invading cluster should occur with positive probability.

no large clusters of one player type exist, will adopt the risk dominant strategy. Individuals will adopt the Pareto dominant strategy only in neighborhoods, in which a sufficient number of individuals playing the Pareto dominant strategy is originally agglomerated. For a random initial distribution, these agglomerations are more likely to occur the larger the population size.

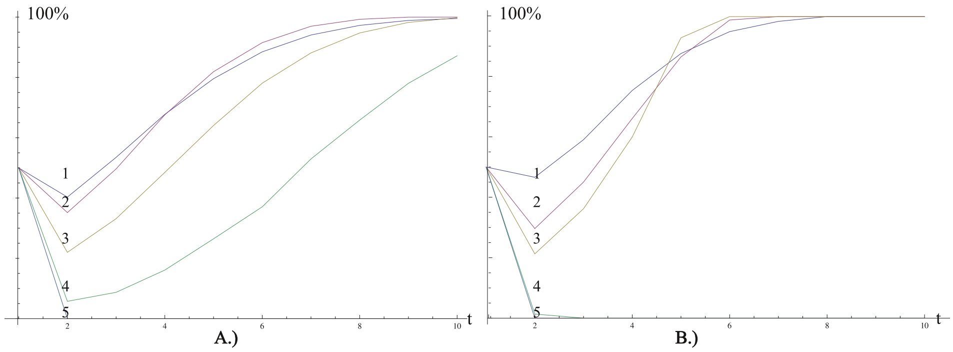

Figure 5: Convergence for two different societies: both figures show the number of *early seeders* A.) 10.000 B.) 441 - Radius from 1 to 5 - the higher the radius the higher the initial decrease in individuals playing the Pareto dominant strategy.

Furthermore, note that minimum sustainable cluster size depends on the imitation radius under consideration. The surrounding clusters, observed after the first interaction sequence, increase with the imitation radius, and thus, minimum cluster size for the Pareto optimal strategy also has to increase in order to be sustainable.[9] This occurs with decreasing probability. Hence, small societies tend towards the risk dominant equilibrium at smaller radii in comparison to larger societies. Figure 5.B. shows the dynamics for a small population.

The large population converges to the Pareto dominant equilibrium for radii smaller than 5, the small population converges only for radii smaller than 4. Figure 13 on page 27 shows how both populations appear after the first period of interaction. This implies, however, that an individual needs to consider 120 neighbors in the large population; in the much smaller population still 80 neighbors to cause the population to converge to the risk dominant equilibrium. In addition, changing the initial distribution only slightly in favor of the Pareto dominant equilibrium requires even higher radii (e.g. changing initial distribution to 58% of Pareto dominant players would require a radius of 7 for both populations in order to converge to the risk dominant equilibrium).

In the case of the Palanpur peasants, it seems plausible that the imitation space, which peasants observe, is much smaller. If that were not the case, peasants could easily implement the Pareto dominant equilibrium by observing all peasants in the village and collectively impose a fine on anyone sowing late. It is therefore more reasonable to assume that the imitation radius is relatively small with respect to the pay-off radius. The space that defines the individual's next period's strategy is defined by those fields, on which the peasant can observe the last yield. Most probably these are the fields surrounding his own. The pay-off radius is, however, defined by the

[9]For the given pay-off, minimum sustainable cluster size is: radius 2 = 14, radius 3=30, radius 4=48 in contrast to 6 for radius=1.

birds' hunting ground. It is highly probable that this radius is much larger. Consequently, the imitation radius is smaller than the pay-off radius.

Simulation shows that, *ceteris paribus*, the convergence towards the risk dominant equilibrium occurs more likely the higher the discrepancy between the imitation and pay-off radius. The following figure 6 illustrates the simulated results for the same initial distributions as the figure before. Simulations have been conducted for pay-off radii from 1 to 5 and imitation radii from 1 to 3, where the first is always greater than the second. The large population of 10.000 individuals only converges entirely to the Pareto dominant equilibrium for an imitation radius of 2 and pay-off radii of 3 or 4 (very slowly after 10 periods), or an imitation radius of 3 and a pay-off radius of 4. Imitation radius of only 1 and pay-off radii of 2 and 3 converge to a mixed equilibrium (approximately 95% and 0.5%). All other combinations converge to the risk dominant equilibrium. The small population of 441 individuals converges only to the Pareto dominant equilibrium, if imitation radius is 2 and pay-off radius 3. The pair of imitation radius of 1 and a pay-off radius 2 converges to the mixed equilibrium, at which approximately 20% of the players choose the Pareto dominant strategy. All other pairs converge again to the risk dominant equilibrium.

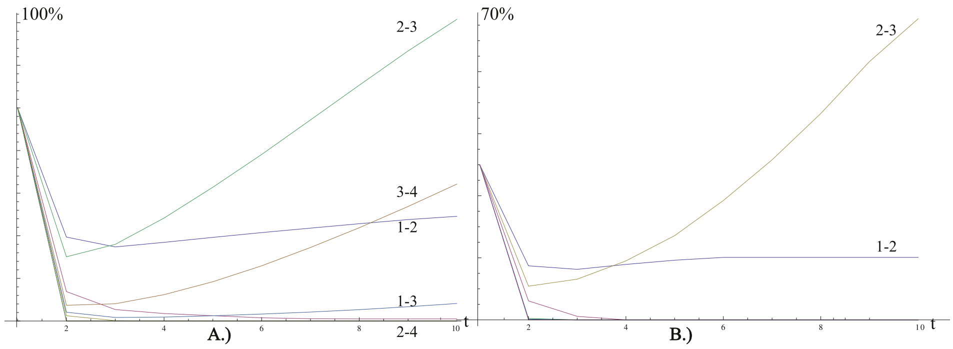

Figure 6: Convergence for two different societies: A.) 10.000 B.) 441. , first value refers to the imitation radius, second value to the pay-off radius

The effect is explained as follows: Increasing the pay-off radius benefits relatively the risk dominant players. A large imitation radius, conversely, increases the spatial effect of a large agglomeration of Pareto dominant players on the neighboring players' strategy choice for the next period. If individuals compare pay-offs only highly locally, the effect of such a large cluster on its surrounding is negligible, leading to the observations:

Observation 1. *Large populations are more likely to converge to the Pareto dominant equilibrium than smaller populations.*

Observation 2. *For small imitation radii with respect to the pay-off radii, a population will converge to the risk dominant equilibrium with high probability.*

This implies a positive relation between the scope, with which individual choice affects other players' pay-offs, and individual short-sightedness on the one hand, and the probability of convergence towards the risk dominant convention on the other hand. If externalities are far reaching, individuals tend to choose the risk dominant strategy. This is exactly the case of Palanpur. Figure 7 shows the structure of the mixed stable equilibria of the former set of simulations.

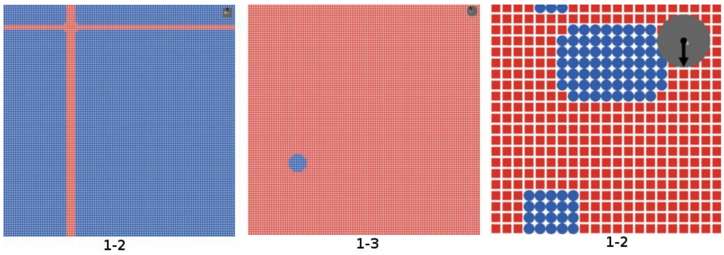

Figure 7: Stable Radius Ratios: first number indicates radius of imitation, second number the pay-off radius.

5 Conclusion

Although the assumptions that have been made at the outset are indeed very idealized, a straightforward interpretation of the abstract findings displays an intricate and intuitive set of results that governs the evolution of social conventions. In the setting with imitation driven strategy choice and local interactions, a shift in conventions can only be triggered by a "minimal group" of individuals that completely adheres to the alternative convention. This must hold for all interacting social levels (i.e. strata - defined by the player type) in this group. The driving force is thereby the type of player, which benefits most from a conventional shift. This contrasts with approaches that favor the risk dominant equilibrium, such as models relying on stochastic stability. In these models the determinant force may be defined by the player type, which loses most from a conventional shift (for details of this critique refer to Bowles 2006). In the model, elaborated in this article, the likelihood of the evolution of a certain convention is defined by a non-linear relationship between pay-off (Pareto) dominance and risk dominance. In addition, for certain "pay-off" constellations several stable conventions evolve and coexist, yet, each convention is respected by a larger group of individuals. Hence, the joint assumption of imitation and local interaction can lead to long-term conventions that are not defined as the stochastically stable states. The approach described in this article thus challenges the assumption of risk dominance (or a minimum stochastic potential) as a general discriminative criterion. Given the results obtained here and assuming a low mutation rate, table 2 roughly summarizes the appropriate discriminatory criterion for the two equilibria in evolutionary 2×2 coordination games.

We further observe that the minimum number of individuals in a group that is required for a transition towards the Pareto dominant equilibrium is determined by the size of an individual's reference group and the degree with which a society is connected. In the case of very small external effects of an individual's strategic choice, Pareto efficient conventions evolve more likely, since fewer like-minded players are necessary for a propagation of the Pareto dominant strategy. In such cases, risk dominant but Pareto inefficient conventions can only persist, if they are incumbent conventions. If a conventional behavior is not yet defined, a population will

Table 2: Criterion for the most likely long-term convention

	imitation of most successful player	versions of best-response play
strictly local interactions	Pareto/Pay-off Dominance (Risk dominance only if high risk)	Risk Dominance
global interactions	Risk Dominance	Risk Dominance

either converge to the Pareto dominant convention or, if both conventions offer almost equal pay-offs, the population will be in a stable state, in which both conventions co-exist. This will not be the case for large external effects, since a positive correlation exists between the size of the externalities and the reference group on the one hand, and the likelihood with which a risk dominant convention evolves. If individuals experience large scope external effects, they will almost certainly follow the risk dominant convention. This effect is reinforced in small secluded societies.

The issue at hand is that, in this model, a population will never adopt a strategy as a convention that is both Pareto and risk inferior. Although intuitive at first sight, it might be doubted if this is generally the case. There existed and still exist examples of societies that adopted persistent interaction patterns that are inferior (for examples see Edgerton 2004). One explanation is that adopting certain social conventions and norms could exhibit a path dependent strategy set, implying that not only full knowledge but also full accessibility to the individual's strategy set cannot generally be assumed. Behavioral patterns and social customs might dictate a bearing that inhibits the evolution towards other equilibria and thus the adoption of certain strategies. Hence, an evolutionary process might turn out to be a blind alley. As Nelson states: "[...][B]eliefs about what is feasible, and what is appropriate, often play a major role in the evolution of institutions." Since a conventional game is assumed with a fixed strategy set for each player, from which he can chose freely under the constraint of the imitation principle, the approach cannot take account of this circumstance. Nevertheless, I believe that a theory that tries to shed light on the evolution of social conventions should allow for such kind of path dependency that social evolution adheres to.

Although the spatial approach neglects the effect of temporal choice constrains that are imposed by the adoption of a specific strategy profiles on the individual of a society, the approach can be directly expanded to incorporate this effect by adding a third dimension on the spatial grid to the model. This third dimension is not a spatial interaction constraint in this context, but represents the choice constraint that a strategy exhibits on the individual for one period of time. The number of periods is identical to the number of restrictive elements in a row along the third dimension and each player finds himself on a unique 2 dimensional plane in each period. The plane will be defined by the strategy choice, which a player has made in the previous period. Hence, his "choice path" is represented by the spatial depth. In consequence, if a player chooses a certain strategy, he is restricted only to a subset of subsequent alternative strategies. A population or cluster can thus end up in a evolutionary dead end, and an evolutionary path once chosen may lead inevitably to an "inferior convention". This extension might provide valuable material for further research.

In addition, fictitious (or adaptive) play and pure imitation are two extreme representations of the heuristics that individuals apply to choose a strategy. For further research, it is appealing to see which results can be obtained by mixed heuristics, and if a threshold can be found defining which degree of imitation will still maintain a positive probability to access the Pareto dominant though risk inferior equilibrium. Since the group of possible learning algorithms is much larger than only those two described herein, a broader analysis might also be of interest. Furthermore, an expansion of the approach to more than two strategies seems also promising for future research, not only with respect to the *survival* of strategies given certain parameter combinations, but also in regard to the spatial patterns that can evolve.

Acknowledgments

Financial support by the University of Siena is gratefully acknowledged. I greatly benefited from discussions with members of the GTI at the GREQAM, Leonardo Boncinelli, Davide Pirino, and especially from those wth Alan Kirman and Simona Settepanella. I am also grateful to an anonymous referee and Mervat Madkour for helpful comments and suggestions.

6 Appendix

Recall the aforementioned conditions $a > c$ and $d > b$, and that h_B pay-off dominates h_A by a positive constant ρ. Define a player that is entirely surrounded by neighbors playing the same strategy like him as an *internal*, otherwise we define him as an *external*. By the definition of cluster size r there exists an *internal* iff $r = 9$. Denote by $\mathfrak{e}_{\mathfrak{C}}$ and $\mathfrak{i}_{\mathfrak{C}}$an external and internal in $\mathfrak{C}$, respectively.

Proof of proposition 1: Assume that the player population is currently in h_A and that there exists a mutant cluster $\mathfrak{C}$ of size r playing strategy B and invading the player population. We observe that, for $\mathfrak{C}$ of size $r = 9$ that $\pi_t^B(\mathfrak{i}_{\mathfrak{C}}) = 6d$ and the maximum pay-off of an external in $\mathfrak{C}$ is $\pi_t^B(\mathfrak{e}_{\mathfrak{C}}) \geq 3c + 5d$. Focusing on the case in which $r < 9$, we redefine pay-offs as $\pi_t^A(i) = \mathfrak{F}_t^A(i)(a-b) + 8b$, if $i \in \mathfrak{I}_t^A$ and $\pi_t^B(i) = \mathfrak{F}_t^A(i)(c-d) + 8d$, if $i \in \mathfrak{I}_t^B$. We observe that $\frac{\partial \pi_t^B(i)}{\partial \mathfrak{F}_t^A} = c - d < 0$. Also note that no player in $\mathfrak{C}$ is *internal*. For any player $l \in \mathfrak{C}$ denote by k a player, s.t. $\pi_t^B(k) = \max\{\pi_t(h) : h \in (\mathfrak{N}(l) \cap \mathfrak{C}) \cup \{l\}\}$. The size of the invading cluster diminishes if l switches, i.e. iff

$$\max\{\pi_t(h) : h \in (\mathfrak{N}(l) \setminus \mathfrak{C})\} > \pi_t^B(k) \tag{9}$$

Since the relation between cluster size r and the pay-off of player k is positive, we have in this case

$$\mathfrak{F}_t^A(k) = (9-r) \text{ and } \pi_t^B(k) = (9-r)c + (r-1)d \tag{10}$$

Two cases occur: either $a \geq b$ or $b > a$.

1. case $a \geq b$: In this case it holds that $\frac{\partial \pi_t^A(i)}{\partial \mathfrak{F}_t^A} = a - b > 0$. Since cluster size of the incumbent strategy is 9, all players not in cluster $\mathfrak{C}$ have at least one *internal* neighbor with pay-off $\pi_t^A(\mathfrak{i}_{\setminus\mathfrak{C}}) = 8a$. For the proliferation of an invading cluster of strategy B players it must thus hold for some external player that $\pi_t^B(\text{external of } \mathfrak{C}) > 8a$ and by equation 10, we obtain that the condition for proliferation is defined by $8a < (9-r)c + (r-1)d$. Since $d = a + \rho$ and $b = c + \rho$, this condition becomes $(2r-10)\rho > (9-r)(a-b)$. Given $a \geq b$, this condition is violated for $r \leq 5$. In addition, a cluster of size smaller than 5 can never be sustained, i.e. never resists an invasion by the incumbent strategy, as $\pi_t^B(k) \leq 5c + 3d < 5a + 3$, where the latter is the smallest pay-off of an external not in $\mathfrak{C}$. Given the results for $\mathfrak{C}$ of size $r = 9$, we observe that the condition for cluster size 6 is sufficient and necessary for larger cluster sizes to expand.

2. *case $a < b$:* In this case it holds that $\frac{\partial \pi_t^{\zeta}(i)}{\partial \mathfrak{F}_t^A} < 0$, for $\zeta = A, B$. In other words, the pay-off inferior strategy benefits from the abundance of individuals playing the pay-off dominant strategy in the neighborhood. Therefore only players neighboring the invading cluster have highest pay-off. The condition for sustaining and proliferation by the incumbent strategy are thus identical, both are determined by equation 9. It is proven by considering the geometric structure of each case that an invading clusters of size smaller than 4 cannot persist. For a cluster of size 3 to proliferate it must hold $6c + 2d > 5a + 3b$, which is a contradiction of $\rho > 0$. A cluster of size 4 can only prevail if its structure is such that all its its players have the same pay-off (a square). In this case it must hold that $5c + 3d > 6a + 2b$ and hence $\rho > 3(a-c)$. If it is not square shaped we require $5c + 3d > 5a + 3b$, a contradiction of assumption $a > c$. For cluster size $r = 5$ the condition is $4c + 4d > 5a + 3b$ and hence $\rho > a - c$. Any larger mutant cluster always resists invasion, since $3c + 5d > 5a + 3b$. □

Proof of proposition 2: For two clusters $\mathfrak{C}_A$ and $\mathfrak{C}_B$, playing strategy A and B respectively, to be neighboring there are at least two players $n \in \mathfrak{C}_A$ and $m \in \mathfrak{C}_B$ with $n \sim m$. Define a player j, such that $\pi_t(j) = \max\{\pi_t(i) : i \in \mathfrak{C}_A\}$ and player k, s.t. $\pi_t(k) = \max\{\pi_t(i) : i \in \mathfrak{C}_B\}$. This implies that player j has highest pay-off in cluster A and k in cluster B.

First concentrate on clusters of $r < 9$. It must be that j, k are external. With positive probability either $j \sim k$, or $l \sim j, k$ for some player l. For none of the players to switch strategy it must be $\pi_t(j) = \pi_t(k)$. The pay-offs of both players can be rewritten as $a(\mathfrak{r}_A - 1) + b(9 - \mathfrak{r}_A) = c(9 - r_B) + d(r_B - 1)$. Notice that r_B defines the size of $\mathfrak{C}_B$, and $\mathfrak{r}_A = r_A$, i.e. the size of $\mathfrak{C}_A$, if $a \geq b$ or $r_A < 3$. If $a < b$ and $r_A \geq 3$, then $\mathfrak{r}_A$ does not necessarily coincide with the size of $\mathfrak{C}_A$, since the pay-off function refers to the player in $\mathfrak{C}_A$, which is least connected to players of the same strategy. Solving the equation shows that for some values of ρ, a set of value pairs $(\mathfrak{r}_A, r_B)$ exists for which the equation is fulfilled.[10] Define such a set of pairs for a given ρ as $R(\rho) = \{(\mathfrak{r}_A, r_B) : \pi_t(j, \mathfrak{r}_A) = \pi_t(k, r_B)\}$. For any two stable neighboring clusters C_A and C_B and a given ρ, it must hold that their value pair $(\mathfrak{r}_A, r_B) \in R(\rho)$. This occurs with zero probability for all such neighboring clusters under the condition of initial random distribution and a large population. At least one cluster collapses triggering the instability of others and at least one strategy thus develops clusters of size 9 with positive probability.

From proposition 1 we know that for $a \geq b$ and a stable cluster of size r_B, playing strategy B,

[10] Yet, some of these pairs are geometrically impossible, e.g. in the case, where $\mathfrak{r}_A = 1$ and $r_B = 2$ and $\rho = \frac{3}{5}(a-c)$, both clusters have identical highest pay-off, but only a cluster of size 4 or larger can fully surround a cluster size 1. 11 feasible pairs remain after ruling out the geometrically impossible pairs. These are $(1,8)$ if $\rho = 7(a-c)$; $(7,4)$ if $\rho = 3(a-c)$; $(2,7), (8,3)$ if $\rho = 5(a-c)$; $(8,7)$ if $\rho = \frac{(a-c)}{5}$; $(7,6), (4,3)$ if $\rho = \frac{(a-c)}{3}$; $(2,6)$ if $\rho = 2(a-c)$; $(3,6)$ if $\rho = (a-c)$; $(8,6)$ if $\rho = \frac{(a-c)}{2}$; while $(5,5)$ is stable for all values of ρ.

surrounded by cluster of size $\mathfrak{r}_A = r_A = 9$ playing A, it must hold $8a \geq d(r_B - 1) + c(9 - r_B) \geq 7a + b$, from which we obtain the second part of the proposition. For $b > a$, $\frac{\partial \pi_t^A(i)}{\partial \mathfrak{F}_t^A} < 0$ and the cluster's maximum pay-off player j is always *external*, if he plays A, for which it must hold $8d \geq a(\mathfrak{r}_A - 1) + b(9 - \mathfrak{r}_A) \geq 7d + c$, which only holds for $r_A = 1$, whence the second result of the proposition. By proposition 1, if $b > a$ no cluster of size $r_B < 9$, playing the pay-off dominant strategy, is stable. □

Proof of proposition 3: The initial distribution is by assumption homogeneous. By proposition 2 some clusters will have a size of 9 with positive probability after some initial period of interaction. Consequently, for two such neighboring clusters $\mathfrak{C}_1$ and $\mathfrak{C}_2$ of size 9, and an external player $\mathfrak{e}_{\mathfrak{C}1} \in \mathfrak{C}_1$, define again two maximum players for each cluster, i.e player k and j, such that $\pi_t(k) = \max\{\pi_t(i) : i \in (\mathfrak{N}(\mathfrak{e}_{\mathfrak{C}1}) \cap \mathfrak{C}_1) \cup \{\mathfrak{e}_{\mathfrak{C}1}\}\}$ and player j, such that $\pi_t(j) = \max\{\pi_t(i) : i \in \mathfrak{N}(\mathfrak{e}_{\mathfrak{C}1}) \cap \mathfrak{C}_2\}$.

Consider the case $a \geq b$. By definition k is internal and j is external, and for cluster size 9 it must hold either $\pi_t(k) = 8a$, if $s_t(\mathfrak{e}_{\mathfrak{C}1}) = A$ or $\pi_t(k) = 8d$, if $s_t(\mathfrak{e}_{\mathfrak{C}1}) = B$. Since $j = external$, his maximum pay-off is either $\pi_t(j) = c + 7d$, if $s_t(\mathfrak{e}_{\mathfrak{C}1}) = A$ or $\pi_t(j) = 7a + b$, if $s_t(\mathfrak{e}_{\mathfrak{C}1}) = B$. For $\mathfrak{e}_{\mathfrak{C}1}$ to switch strategy, it must hold $\pi_t(k) < \pi_t(j)$. Since $\rho > 0$ only $c + 7d > 8a$ occurs without contradiction. In the case $a < b$, k is external if $s_t(\mathfrak{e}_{\mathfrak{C}1}) = A$ with maximum pay-off $\pi_t(k) = 7b + a < 7d + c$. Thus h_A cannot prevail by assumption $a > d$. □

Proof of proposition 4: This is a direct consequence of the former proofs: Given that h_B pay-off dominates h_A, by the former propositions, in the case of a heterogeneous initial distribution the pay-off dominant strategy overtakes if $3\hat{c} + 5d > 8a$ and the risk dominant strategy prevails if $7a + b > 3\hat{c} + 5d$. For a homogeneous initial distribution the constrains are $\hat{c} + 7d > 8a$ and $7a + b > 8d$. The condition for the risk dominant strategy to prevail in a homogeneous distributed population is thus $\rho < |\frac{c-a}{7}|$, where the latter is the marginal perceptible unit for a pay-off dominant strategy to invade a population. □

Recall that $a_i > c_i$ and $d_i > b_i$.

Proof of proposition 5: Consider two clusters $\mathfrak{P}$ and $\mathfrak{M}$, where the former is pure and the latter mixed. Define the set of players of type x as X and the set of type y players as Y. Assume without loss of generality that $s_t(j) = A, \forall j \in \mathfrak{P}$, and for cluster $\mathfrak{M}$ that $s_t(j) = A, \forall j \in (\mathfrak{M} \cap X)$ and $s_t(j) = B, \forall j \in (\mathfrak{M} \cap Y)$. Since strategical change occurs only at borders of clusters, we need only to consider an external player $\mathfrak{e}$ such that $(\mathfrak{N}(\mathfrak{e}) \cap \mathfrak{P}) \neq \emptyset$ and $(\mathfrak{N}(\mathfrak{e}) \cap \mathfrak{M}) \neq \emptyset$.

First, assume that $\mathfrak{e} \in (\mathfrak{M} \cap X)$. Thus $s_t(\mathfrak{e}) = A$, but also $s_t(i) = A, \forall i \in \mathfrak{N}_x(\mathfrak{e})$. Consequently, $s_{t+1}(\mathfrak{e}) = A$. The same holds if $\mathfrak{e} \in (\mathfrak{P} \cap X)$. Now assume that $\mathfrak{e} \in (\mathfrak{M} \cap Y)$, thus $s_t(\mathfrak{e}) = B$. Since, $s_t(i) = A, \forall i \in (\mathfrak{P} \cup \mathfrak{M}) \cap X$, it follows that $\pi_t(f) = 8a_i, \forall f \in (\mathfrak{P} \cap \mathfrak{N}_y(\mathfrak{e}))$ and $\pi_t(g) = 8\hat{c}_i, \forall g \in (\mathfrak{M} \cap \mathfrak{N}_y(\mathfrak{e}))$. As $a_i > \hat{c}_i$, it follows $s_{t+1}(\mathfrak{e}) = A$. The same, for $\mathfrak{e} \in (\mathfrak{P} \cap Y)$. Equivalent results are obtained for $s_t(h) = B, \forall h \in \mathfrak{P}$, since $\pi_t(f) = 8d_i > 8b_i = \pi_t(g)$. Consequently, an external always chooses the strategy that is played by the pure cluster in his neighborhood if he has not done so before. □

Proof of proposition 6: We make the additional assumption that $a_i > b_i$. By proposition 5 and proposition 2, it follows for an initially homogeneous distribution that, after an initial sequence of interactions, large pure clusters $\mathfrak{P}_A$ and $\mathfrak{P}_B$ of size $r_A, r_B = 9$ play strategy A and B, respectively. Further by proposition 5, player types can be neglected with respect to the dynamics and type specific subscripts are left out when not required.

Assume an external player $\mathfrak{e}_A \in \mathfrak{P}_A$ and a external player $\mathfrak{e}_B \in \mathfrak{P}_B$ with $\mathfrak{e}_A \sim \mathfrak{e}_B$. In order to cause $\mathfrak{e}_B$ to change strategy, it must be that $\pi_t(\mathfrak{e}_A) > \pi_t(\mathfrak{i}_B)$, given internal player $\mathfrak{i}_B \in \mathfrak{P}_B$ with $\mathfrak{i}_B \sim \mathfrak{e}_B$. Since $\mathfrak{i}_B$ is internal, it follows that $\pi_t(\mathfrak{i}_B) = 8d_i$. Define $\eta_A = \#\left(\mathfrak{I}_t^B \cap \mathfrak{N}(\mathfrak{e}_A)\right)$. In general the pay-off of $\mathfrak{e}_A$ is then given by $\pi_t(\mathfrak{e}_A) = (8 - \eta_A)a_i + \eta_A b_i$, leading to condition $(8 - \eta_A)a_i + \eta_A b_i > 8d_i$. Similarly, define $\eta_B = \#\left(\mathfrak{I}_t^A \cap \mathfrak{N}(\mathfrak{e}_B)\right)$. There exists an internal player $\mathfrak{i}_A \in \mathfrak{P}_A$ with $\mathfrak{i}_A \sim \mathfrak{e}_A$. Consequently $\pi_t(\mathfrak{i}_A) = 8a_i$ and to trigger a strategy switch of player $\mathfrak{e}_A$ it must hold $(8 - \eta_B)d_i + \eta_B \hat{c}_i > 8a_i$.

For $i = 1, 2$ define $\max_i \{\lceil \eta_A \rceil_i\}$ and $\max_i \{\lceil \eta_B \rceil_i\}$ as the largest integer that fulfills each condition, respectively, given the type specific parameters values. Since the probability that $\pi_t(\mathfrak{e}_A) > \pi_t(\mathfrak{i}_B)$ is proportional to $\max_i \{\lceil \eta_A \rceil_i\}$ and the probability that $\pi_t(\mathfrak{e}_B) > \pi_t(\mathfrak{i}_A)$ is proportional to $\max_i \{\lceil \eta_B \rceil_i\}$, both values determine the likelihood by which a cluster expands and thus the speed, at which each type pushes the population towards the corresponding equilibrium. □

7 Simulations

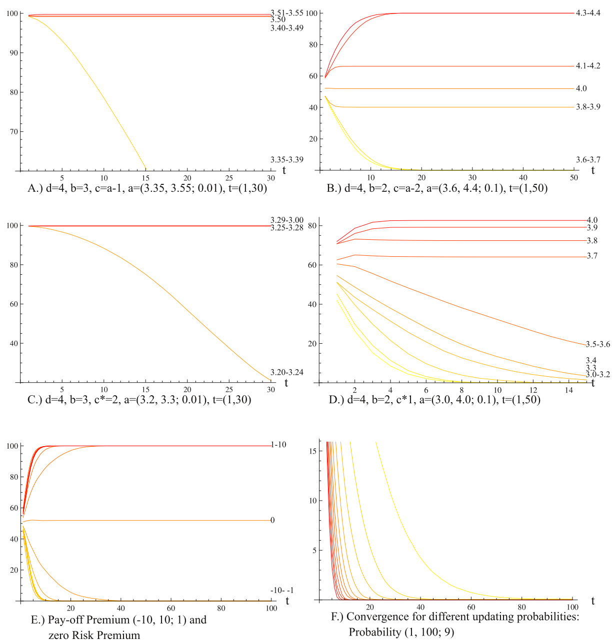

Figure 8: All graphs show the share of A players for the given pay-off value and period t of interaction.

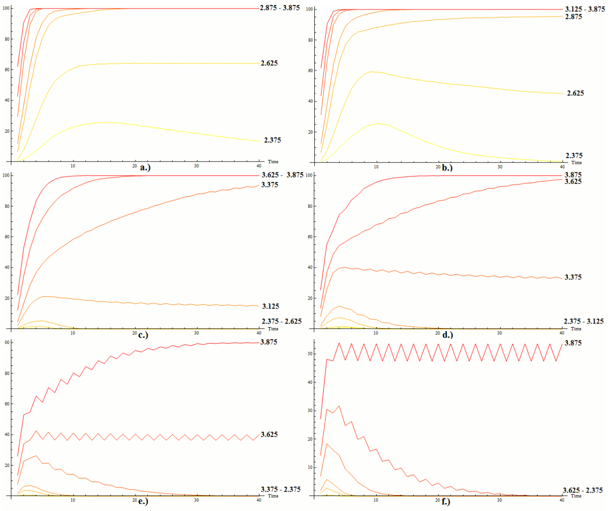

Figure 9: Simulation for $a_1 \in (2.375, 3.875; 0.25)$ A.) $\eta_A = 5$; B.) $\eta_A = 4$; C.) $\eta_A = 3$; D.) $\eta_A = 2$; E.) $\eta_A = 1$; F.) $\eta_A = 0$; in C.)-F.) initial share of $B = 55\%$.

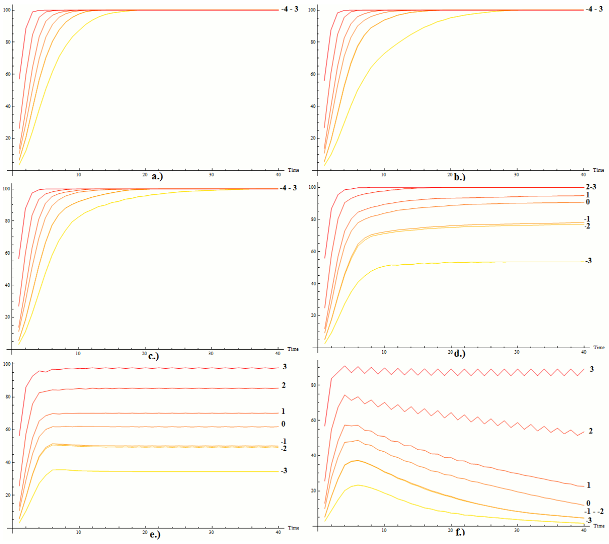

Figure 10: Simulation for $b_1 \in (-4, 3; 1)$ A.) $\eta_A = 5$; B.) $\eta_A = 4$; C.) $\eta_A = 3$; D.) $\eta_A = 2$; E.) $\eta_A = 1$; F.) $\eta_A = 0$.

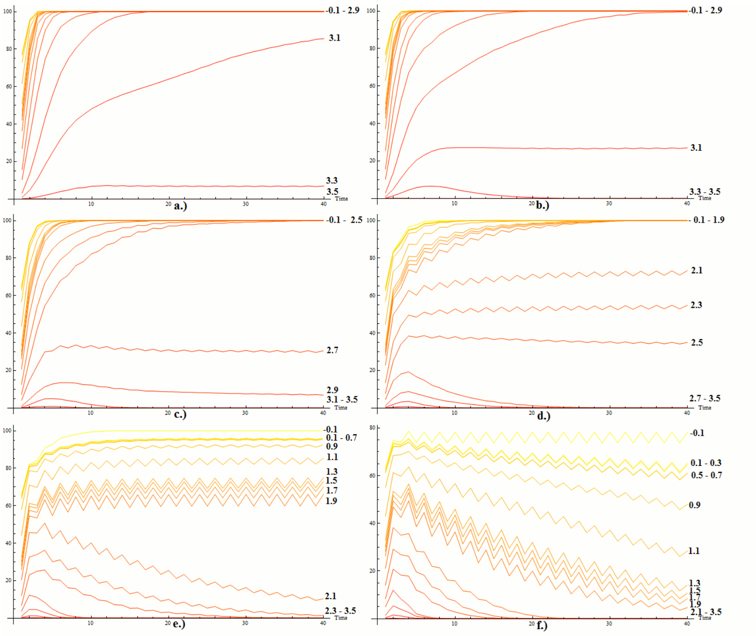

Figure 11: Simulation for $c_1^* \in (-0.1, 3.5; 0.2)$ A.) $\eta_A = 5$; B.) $\eta_A = 4$; C.) $\eta_A = 3$; D.) $\eta_A = 2$; E.) $\eta_A = 1$; F.) $\eta_A = 0$; A.)-B.) 55% strategy B players, C.)-F.) 60% strategy B players.

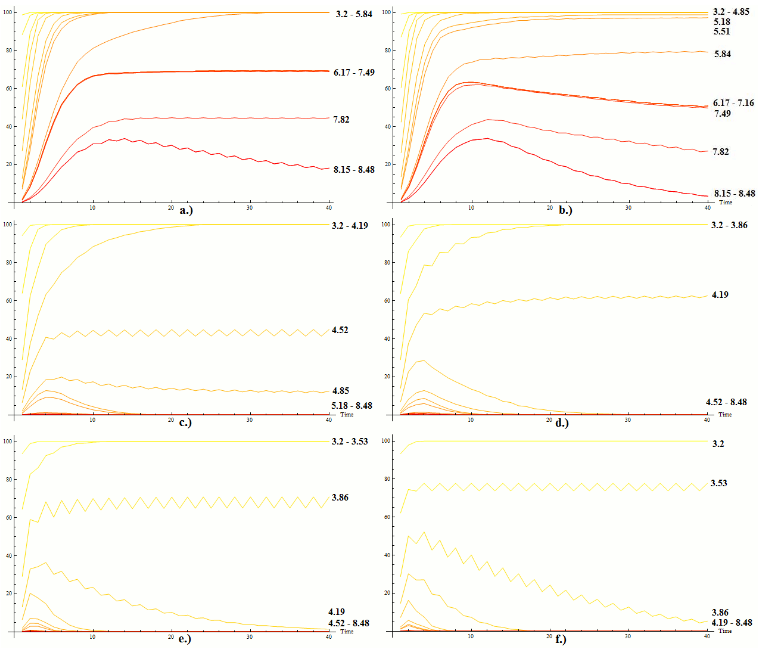

Figure 12: Simulation for $d_1 \in (3.2, 8.48; 0.33)$ A.) $\eta_A = 5$; B.) $\eta_A = 4$; C.) $\eta_A = 3$; D.) $\eta_A = 2$; E.) $\eta_A = 1$; F.) $\eta_A = 0$; C.)-F.) 60% strategy B players.

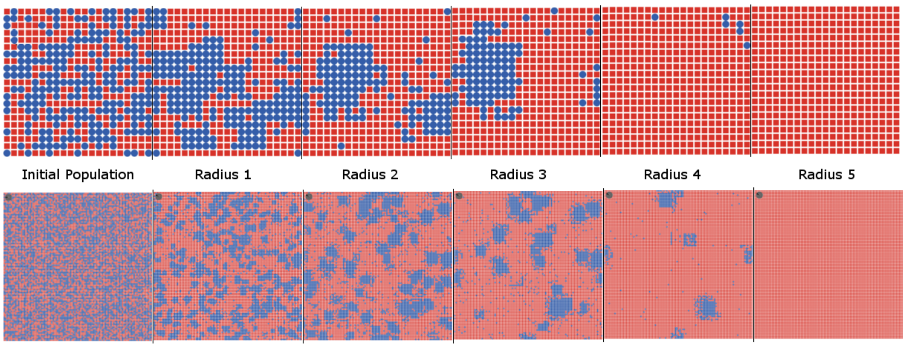

Figure 13: Two populations (10.000 and 441 individuals) after 1 interaction period.

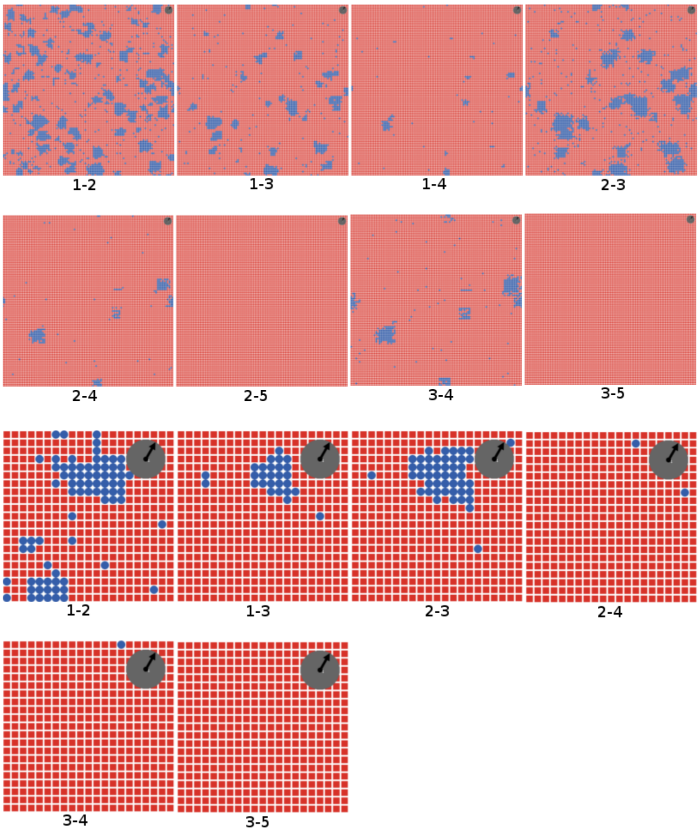

Figure 14: Two different societies (10.000 and 441 individuals), distribution after one period of interaction - first number indicates radius of imitation, second number pay-off radius.

References

Ades, Alberto and Di Tella, Rafael [1996] The causes and consequences of corruption. *IDS Bulletin*, **27**, 35–49.

Bicchieri, Cristina [2006] The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press.

Bollinger, Daniel and Hofstede, Geert [1987] Les différences culturelles dans le management. Les Editions d'Organisation.

Bowles, Samuel [2006] Microeconomics: Behavior, Institutions, and Evolution (The Roundtable Series in Behavioral Economics). Princeton University Press.

Brandt, Hannelore, Hauert, Christoph and Sigmund, Karl [2003] Punishment and reputation in spatial public goods games. *Proceedings of the Royal Society of London*, **270**, 1099–1104.

Durlauf, Steven N. and Young, H. Peyton [2001] Social Dynamics. The MIT Press.

Edgerton, Robert B. [2004] Traditionelle Überzeugungen und Praktiken: Gibt es bessere oder schlechtere?, Goldmann Verlag. 189–210.

Ellison, Glenn [1993] Learning, local interaction, and coordination. *Econometrica*, **61**, 1047–1071.

Ellison, Glenn [2000] Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *The Review of Economic Studies*, **67**, 17–45.

Harrison, Lawrence E. and Huntington, Samuel P. [2000] Culture Matters: How Values Shape Human Progress. Basic Books.

Hauert, Christoph [2001] Fundamental clusters in spatial 2×2 games. *Proceedings of the Royal Society of London*, **268**, 761–769.

Hodgson, Geoffrey [1996] Economics and Evolution: Bringing Life Back into Economics. University of Michigan Press.

Huntington, Samuel P. [1997] The Clash of Civilizations and the Remaking of World Order. Touchstone Press, 1 edition.

Huntington, Samuel P. and Harrison, Lawrence E. [2004] Streit um Werte: Wie Kulturen den Fortschritt prgen. Goldmann Verlag.

Kandori, Michihiro, Mailath, George J. and Rob, Rafael [1993] Learning, mutation, and long run equilibria in games. *Econometrica*, **61**, 29–56.

Lee, In Ho and Valentinyi, Ákos [2000] Noisy contagion without mutation. *The Review of Economic Studies*, **67**, 47–56.

Morris, Stephen [2000] Contagion. *The Review of Economic Studies*, **67**, 57–78.

Nowak, Martin A. and May, Robert M. [1992] Evolutionary games and spatial chaos. *Nature*, **359**, 826–829.

Page, Scott E. [2007] The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton University Press.

Potts, Jason [2000] The New Evolutionary Microeconomics: Complexity, Competence, and Adaptive Behaviour. Edward Elgar Publishing Ltd.

Robson, Arthur and Vega-Redondo, Fernando [1996] Efficient equilibrium selection in evolutionary games with random matching. *Journal of Economic Theory*, **70**, 65–92.

Weber, Max [2007] Wirtschaft und Gesellschaft. area Verlag.

Welzel, Christian and Inglehart, Ronald [1999] Analyzing Democratic Change and Stability: A Human Development Theory of Democracy. Wissenschaftszentrum Berlin für Sozialforschung GmbH (WZB).

Young, H. Peyton [1993] The evolution of conventions. *Econometrica*, **61**, 57–84.

Young, H. Peyton [2005] Diffusion of innovations in social networks, Oxford University Press.

Young, Peyton H. [1998] Individual Strategy and Social Structure. Princeton University Press.