Evolution of Social Behavior in the Global Economy: The Replicator Dynamics with Migration

Suren Basov[1], [2]

May 23, 2002

Abstract

In this paper I consider the e¤ects of migration on the spread and the speed of the propagation of new conventions, technologies, etc. I show that the speed of the propagation increases with the openness of the economy. The application of the model to the equilibrium selection in $2 \pounds 2$ coordination games is also discussed.

1   INTRODUCTION

New ideas, norms, and conventions are born at a speci...c time and at a speci...c location, and only gradually they take hold at the place of their birth and spread over the larger areas. How, exactly, do they evolve and spread? By now there exists a considerable literature on the evolution of social behavior. Foster and Young (1990), Kandori, Mailath, and Rob (KMR) (1993), and Young (1993) use evolutionary models with a persistent randomness to study the long-run behavior in games. They got some strong results on the equilibrium selection. The evolution, however, took place only in time and the spatial dimension was completely ignored.

The papers that introduce some spatial relationship between the players are Anderlini and Ianni (1996), Blume (1993, 1995), Ellison (1993), Ely (1995), Young (1999). These papers assume that the players are connected by some network, and that the behavior adopted by an individual depends on its intrinsic payo¤ and the behavior of her neighbors. The general conclusion of this literature is that the speed of evolution is facilitated by the local nature of the interactions. However, the spatial dimension is modeled in a too stylized fashion, which does not allow us to study the spread of the behavior in the real space. To illustrate the point, suppose you have

a farming community where the neighbors live miles away from each other. Would a new technology originated at one of the farms have traveled faster had they lived next to each other? To give a more modern example, would the globalization, lowering the migration costs, result in a faster diffusion of conventions and customs? How exactly is the speed of diffusion linked to the openness of the global economy? To answer these types of questions, one has to have a model where the evolution takes place in the real physical space rather then in a network.

In this paper I suggest such a model. I consider the evolution of the social behavior when the individuals are able to migrate. The motivation for this is twofold. First, I feel that the migration is an important phenomenon that takes place in the real world. Second, the model will allow me to address some interesting questions, which cannot be even formulated in the previous literature.

A basic model is rather simple. Assume that each member of the population possesses a certain behavioral trait. The trait can be interpreted as a strategy in some game, a determinant of the preferences (a certain "gene" may force you to reciprocate, to have a child out of wedlock, to go to the college, or to commit a certain kind of a crime), or a belief. In the case when

2

a trait is interpreted as a determinant of the preferences, one has to distinguish between the utility payo¤ and the ...tness payo¤. The distinction is well known in the literature on the evolution of preferences. (e. g. Bergstrom 1995, Robson 1996). If a trait is a strategy in a 2 £ 2 coordination game, I will refer to it as a convention. The fraction of the people at a particular location that posses a certain trait is assumed to change due to the di¤erential replication and the migration.

The di¤erential replication is described by the replicator dynamics. For a discussion of the replicator dynamics see, for example, Samuelson (1997). Corradi and Sarin (2000), Schlag (1998), and Samuelson (1997) proposed several behavioral models that give rise to the replicator dynamics. I sketch a slightly modi...ed version of Samuelson's Aspiration and Imitation model in the Appendix 1 to this paper.

The second force that changes the share of the population possessing a certain trait is the migration. The ‡ow of the migrants is assumed to have two components, a strategic component, related to the payo¤s di¤erentials at di¤erent locations, and a random component. Intuitively, a particular trait determines the payo¤s an individual gets only in some of her relations. The decision to migrate, on the other hand, is assumed to be a¤ected by the

3

total payoﬀ the individual can get at the location, which is the sum of the payoﬀs over all the relations. If a trait is important for almost all kinds of the relations (e. g. ability), then the decision to migrate will be strategic with respect to this trait. However, if a trait aﬀects only a small fraction of the relations (e. g. reciprocity) then its contribution to the total utility earned at the particular location is negligible and the migration decision is approximately uncorrelated with this trait.

If a trait is interpreted as a strategy in some game, in the long-run the model described in this paper can be viewed as a model of equilibrium selection. I will show that in 2×2 coordination games it selects the risk-dominant equilibrium, which is similar to the result obtained by Foster and Young (1990), KMR (1993), and Young (1993). It should be stressed that the mathematics of the equilibrium selection in this model is similar to those of Foster and Young. This implies that the selection of the risk-dominant equilibrium depends crucially on the properties of the replicator dynamics and does not generalize to the arbitrary payoﬀ monotone dynamics, which would be the case in the KMR framework.

The model can also be applied to study a speed of the adoption of a new technology. There is a considerable evidence to demonstrate that the

di¤usion of new technologies is a spatial variable. See for example, Thwaites (1982) and Rees, Briggs, Oakey (1984). This conclusion is consistent with my model. Another explanation was suggested by Baptista (2000), who argues that externalities promoting the adoption of new technologies are stronger on the regional level.

The model developed in this paper allows us to go beyond making the long-run predictions. It provides a description of the dynamics of a convention. It can be used to demonstrate that though in the one-dimensional world the local conventions can spread, this is not the case in the two-dimensional world. In order to spread, a convention should initially arise at some non-trivial area.

The model also allows us to study the speed of the propagation of new traits. I show that the traits that a¤ect the payo¤s in many relations spread faster then those that a¤ect few, provided that the rate of di¤erential reproduction is su¢ciently sensitive to the payo¤ di¤erences. I also show that the speed of the propagation of a trait depends on the degree of the openness of the economy. The form of this dependence does not depend on the functional form of the payo¤s. This conclusion is of the particular importance in the view of the globalization of the world economy.

Another interesting conclusion is that in the absence of spatial dependence of the payoffs all the world will eventually adopt the same customs and conventions. Hence, to explain the differences in the national cultures one has to assume that the payoffs to a particular trait differ with the location due, for example, to the difference in physical conditions. Just modelling the social interactions as a game with multiple equilibria is not enough to explain the differences in the long-run behavior across different locations. The dependence of the social conventions on the physical conditions was always appreciated by the historians. Trevelyan (1937), for example, begins his volume of the English history from the description of the geography of the island.

The paper is organized as follows. In Section 2 I introduce a general model of evolution with migration and analyze the properties of the solution. In Section 3 I give some examples. In Section 4 I compare the results of this paper with the results in the literature and discuss some possible extensions. The paper has two Appendices. Appendix 1 contains a review of a version of the Samuelson's Aspiration and Imitation model in the case when the population is hyper...nite. Appendix 2 briefly discusses the construction of the hyperreal numbers from the reals.

2 THE MODEL

In this section I formulate the master equation of my model and study some of its properties. For simplicity, I will concentrate on the case when the trait can take only two values 0 and 1: a strategy in a 2×2 symmetric game, an attitude towards reciprocity (an individual may either return favors or not return them), or an ability in a signalling game. The generalization of the model for the traits that can take more than two values is straightforward. Let $u(x; t)$ be the fraction of the individuals who are located at location x at time t and posses trait 0. I will assume that $x \in \Omega \subseteq R^m$, where Ω is compact and one-connected. In applications I will put $m = 1; 2$, but I will develop the theory for the general case (perhaps, in the anticipation of the space travel). Let $\pi_i(u; x)$ be the payoff to trait i, which may depend on the fraction of the individuals, possessing the trait and on the location. In the absence of migration the evolution of u is assumed to be governed by

$$\frac{\partial u}{\partial t} = u(1 - u)(\pi_0(u; x) - \pi_1(u; x)): \qquad (1)$$

Equation (1) states that the rate of change of the fraction of the individuals possessing a certain trait is proportional to the difference between the payoff

they earn and the payoff earned by the individuals possessing an alternative trait.[1]

To introduce migration consider a compact set $V \subseteq \Re$ with a smooth boundary $\S$. The change of the measure of the individuals possessing a certain trait and located within V occurs due to the differential replication and the migration. Therefore, assuming that the rate of the differential replication is given by the right hand side of (1), one can write

$$\int_V \frac{\partial u}{\partial t} dx = \int_V u(1 - u)(\pi_0(u; x) - \pi_1(u; x))dx + \oint_\S m d\S \qquad (2)$$

Vector m is the net migration flow through $\S$ of the individuals possessing trait 0. Let

$$\Delta(u; x) = \pi_0(u; x) - \pi_1(u; x)$$

In the absence of the strategic motive the net outflow of the migrants from a particular location will be proportional to the size of the population at this location. For example, even assuming the economic conditions are the same

---

[1]Appendix to this paper derives the replicator dynamics from a behavioral model in the context of a 2 × 2 coordination game.

all over the globe, it is likely that the would be more emigrants from China than immigrants into China. Hence, the migration will tend to equalize the fraction of the individuals possessing a certain trait across the locations, i. e. the migration flow will be proportional to the gradient of $u(\cent)$. In general, I will assume that

$$m = {}^\circ({}^\circledR \nabla u + (1_i {}^\circledR)(\acute{A}_u \nabla u + \nabla_x \acute{A})). \qquad (3)$$

Coefficient ${}^\circledR 2 [0; 1]$ measures a random component of the migration flow. If ${}^\circledR = 0$ the decision to migrate is purely strategic, while if ${}^\circledR = 1$ it is totally random. Coefficient ${}^\circ \ge 0$ measures the openness of the economy. Small values of ${}^\circ$ correspond to a closed economy with strict restrictions on the migration.

Expression (3) can be derived from a behavioral model similar to that presented in Appendix 1. One has to assume that the individuals are located in the vertices of a lattice with an in...nitesimal[2] side h and that the probability to move from a vertex of the lattice to a neighboring one in an (in...nitesimal) unit of time is proportional to this unit of time and non-decreasing in the

---

[2] see the Appendix 2 and Albervito, Fenstand, Høegh-Krohn, Lindstrøm (1986) for a discussion of the in...nitesimal numbers and other notions of the nonstandard analysis.

payo¤ di¤erence between the vertexes.

Using the divergence theorem transform the second term on the right hand side of (2)

$$\oint_{\S} m d\S = \int_V div(m) dx: \qquad (4)$$

Since (2) should hold for any compact set with a smooth boundary, one obtains

$$\frac{@u}{@t} = f(u; x) + °div(®r u + (1 ¡ ®)(Á_u r u + r_x Á)); \qquad (5)$$

where $f(u; x) = u(1 ¡ u)Á(u; x)$.

Now let us return to the discussion of (3). To understand the idea of the strategic versus nonstrategic migration assume that at each moment of time and at each location an individual can be called upon to participate in one of K games. Both, the game and the partner are chosen at random. With slight abuse of notation denote $K = f1; ::::; Kg$. Let $\frac14_k(x; t; ¿)$ be the expected payo¤ from participating in game k at location x at time t and possessing trait ¿. For a ...xed $(x; t; ¿)$ it can be viewed as a random variable

10

de...ned on set K. De...ne the expected payo¤ at time t at location x as

$$\text{¦}(x;t;\imath) = \frac{1}{K}\sum_{k=1}^{\text{✖}} \frac14_k(x;t;\imath):$$

It is plausible to assume that the decision to migrate is determined by $\text{¦}(x;t;\imath)$. Let

$$A(\imath;K) = \text{fk } 2 \text{ } K : \frac14_k(x;t;\imath) \text{ depends on } \imath g;$$

and $\#A(\imath;K)$ is the cardinality of set $\#A(\imath;K)$. If

$$\lim_{K!\,1} \frac{\#A(\imath;K)}{K} = 0$$

(trait $\imath$ a¤ects payo¤s payo¤s in a small fraction of games/relations), then asymptotically $\text{¦}(x;t;\imath)$ does not depend on $\imath$. Hence, the decision to migrate is not strategic in $\imath$. In particular, the probability of the migration should be independent of the payo¤ di¤erentials in a particular game affected by $\imath$. In this case $^{\circledR} = 0$ in expression (3). Otherwise, the decision to migrate has a strategic motive and $^{\circledR} 2 (0;1]$.

The learning process is, however, game speci...c. It may take, for example,

the form speci...ed in Appendix 1. In this case, the change of the share of
the individuals possessing trait ¿ at location x is governed by the replicator
dynamics.

De...ne a function

$$V(u; x) = \int_0^{Z_u} f(z; x)dz;$$

and a functional

$$F(u) = \circ \int_V^Z [h^® r u + (1_i \ ^®)(Á_u r u + r_x Á); r u i_i V(u; x)]dx; \qquad (6)$$

where h¢; ¢i denotes the inner product of two vectors. I will begin the analysis
of dynamics (5) proving the following theorem:

Theorem 1 Let equation (5) hold. Then

$$\frac{dF(u(t))}{dt} \cdot 0. \qquad (7)$$

Proof. Note that equation (5) can be rewritten in the form

$$\frac{@u}{@t} = i \frac{\pm F}{\pm u} \qquad (8)$$

where $\pm F = \pm u$ is the variational derivative of F. Now, using the chain rule and taking into account (8):

$$\frac{dF(u(t))}{dt} = \int_V \frac{\pm F}{\pm u} \frac{@u}{@t} = i \int_V (\frac{\pm F}{\pm u})^2 dt \cdot 0: \qquad (9)$$

∎

Let $f(u; x)$ does not depend on x. Then the immediate Corollary of Theorem 1 is:

Corollary 2 Let u^{π} be a strict local maximum of $V(\updownarrow)$. Then a stationary uniform solution of (5) $u(x; t) = u^{\pi}$ is locally asymtotically stable.

Proof. De...ne Lyapunov function for (5) by

$$L(t) = F(u(t)): \qquad (10)$$

Then the result follows from Theorem 1. ∎

13

Function $V(\cdot)$ can be interpreted as a potential of the system. I will call the local maxima of $V(u)$ the long-run outcomes. Apparently, a long-run outcome is history dependent.

Assume that once in awhile some non-trivial fraction of the population mutates. The question is, which of the local maxima of $V(u)$ will be stable under such a mutation. It turns out that only the global maxima have this property. I will call them the very long-run outcomes.

I will not prove the claim made in the previous paragraph in the full generality (for a proof see Bugaenko et. al., 1993). Instead I will consider in detail a special case, which is of a particular importance in the economic applications.

Assume, $\acute{A}(u;x)$ does not depend on x and has a form

$$\acute{A}(u) = {}^{-}(u \mathbin{\text{¡}} u_\text{¤}); \qquad\qquad (11)$$

where $u_\text{¤} \, 2 \, (0;1)$. This assumption is satis...ed, for example, if the trait is a strategy in a $2 \, \pounds \, 2$ symmetric game. For $^{-} < 0$ there exists a unique stationary solution $u = u_\text{¤} \, 2 \, (0;1)$ and the population in the equilibrium will be heteromorphic. For $^{-} > 0$ the function $V(\cdot)$ achieves its local maxima for

14

the monomorphic populations ($u = 0$ or $u = 1$). If the trait is a strategy in a 2×2 symmetric game, the first case will be realized for a zero-sum game (u_* will correspond to the equilibrium probability of strategy 0), while the second will be realized for a coordination game (u_* will correspond to the equilibrium probability of strategy 0 in the mixed-strategy equilibrium).

Let us consider the case $\delta > 0$. Though both monomorphic states are locally stable, sufficiently big exogenous disturbance (e. g. an invasion of the individuals possessing a different trait) can take the system away from one steady state into another. The switching does not occur simultaneously at all locations, but rather travels along the system in a form of a switching wave. Our next goal is to calculate the speed of the propagation of the switching wave.

Define the total migration coefficient by

$$D = \gamma(\rho + (1 - \rho)\delta): \qquad (12)$$

Then equation (5) takes the form

$$\frac{\partial u}{\partial t} = f(u) + D \Delta u \qquad (13)$$

15

where

$$\mathbb{C} = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}: \tag{14}$$

Let us ...rst consider the case $m = 1$. The switching wave from the state $u = 1$ into the state $u = 0$ moving with the speed c is the solution of equation (13) of the form

$$u = u(x - ct) \tag{15}$$

satisfying the boundary conditions

$$\lim_{y \to 1} u(y) = 0, \quad \lim_{y \to -1} u(y) = 1; \tag{16}$$

where $y = x - ct$. Plugging (15) into (13) one obtains

$$Du'' = -f(u) - cu': \tag{17}$$

Multiplying both sides of (17) by u', integrating from -1 to 1, and taking

16

into account boundary conditions (16):

$$c \int_{i1}^{Z1} [u^0(y)]^2 dy = \frac{-}{12}(1 i 2u_\alpha):$$

(18)

There exists a unique value of c, such that (18) is satis...ed. After performing

the necessary calculations (see Bugaenko, et. al., 1993) one obtains:

$$c = \frac{1}{2} P \overline{D}(1 i 2u_\alpha):$$

(19)

One can verify by a direct calculation that $c > 0$ if and only if $V(0) >$

$V(1)$, hence in the long run the system will adopt a convention with a higher

value of V. This conclusion is general and does not depend on a speci...c

functional form of f which is chosen.

A steady state, which delivers the global maximum to V is called stable,

while the local maxima are called metastable. In the very long-run the system

will move away from a metastable to the stable steady state. If $V(1) = V(0)$

then $c = 0$ and both steady states can coexist in the very long-run. At the

point of switching u will change with a jump.

Several properties of (19) are worth noting. First, note that if $^\circledR = 0$

17

then c is proportional to $^-$, while c is proportional to $\overset{\text{p}}{=}$ for $\circledR = 1$. This means that if the migration is strategic the speed of the switching wave is more sensitive to a payo¤ di¤erential then when migration is random. The above result also suggests that if the rate of the di¤erential reproduction is su¢ciently sensitive to the payo¤ di¤erentials ($^- > 1$) the conventions which a¤ect more relations spread more rapidly then the ones which a¤ect fewer relations.

Second, note that the speed of the switching wave is increasing in the degree of openness of the economy and is proportional to the square root of the openness. This conclusion does not depend on the functional form of Á(¢) and can be obtained directly from (5) if one looks for a solution is a form

$$u(x;t) = u(\text{»}) \tag{20}$$
$$\text{»} = \frac{x_{\text{i}} \, ct}{\overset{\text{p}}{=}}: \tag{21}$$

Then (5) will imply

$$\frac{c}{\overset{\text{p}}{=}} u^{0} = {}_{\text{i}} \, f(u) \, {}_{\text{i}} \, \frac{@}{@\text{»}} (\circledR u^{0} + (1 \, {}_{\text{i}} \, \circledR)(\text{Á}_{u} u^{0} + \text{Á}_{\text{»}})): \tag{22}$$

The speed c can be determined from (22) subject to (16). Since neither right

18

hand side of (22) nor (16) contain $°$, solution to (22), (16) will have a form

$$c = {}^P\!\!\!{}_\sigma h(®; [f]; [\acute{A}]). \tag{23}$$

If one hypothesizes that the globalization increases the degree of openness, it would imply that the globalization will increase the speed of the univer-salization of the behavior across the di¤erent locations.

So far, I assumed that the system moves from one steady state to another and computed the speed of the switching wave. Next, I will ask what is the minimal disturbance that will bring a system from one steady state to another. More speci...cally, assume that the globally stable state is $u = 1$, hence

$$\int_0^1 f(u)du > 0: \tag{24}$$

Assume also that at $t = ¡1$ the system is at the state $u = 0$. Then at time

19

$t = 0$ a disturbance centered near location $x = 0$ is created, that is

$$u(x;0) = u_0; \quad 0 < u_0 \cdot 1. \tag{25}$$

$$\frac{@u}{@x}(x;0) = 0 \tag{26}$$

$$\lim_{jxj! \, 1} u(x) = 0: \tag{27}$$

One can interpret (25)-(27) as an invasion of mutants that settle near $x = 0$. Note that I allow for u to decrease arbitrary fast, that is an invasion can be local. The question is, what is the minimal value of u_0 that will cause switching to the state $u = 1$. The minimal disturbance u corresponds to the stationary unstable solution of (5). Under our assumptions on the payo¤ function it reduces to

$$Du^{00} + f(u) = 0 \tag{28}$$

subject to (25)-(27). Integrating (28) with respect to x and using (25)-(27) one gets:

$$\mathbf{Z}^{u_0}_{0} f(u)du = 0: \tag{29}$$

20

Taking into account (24) one obtains $u_0 < 1$. If the integral in (24) is large enough, that is the value of $V(\cent)$ at the global maximum is much bigger than at the local one, u_0 becomes very small and c becomes very big. This implies that in this case a small mutation is sufficient to bring the system from a metastable to the stable set and it travels fast. Vice versa, if $V(\cent)$ achieves approximately the same value at both local maxima, then a large mutation is needed to take the population from a local to the global maximum and it travels slow.

We have seen that the very long-run outcome is generically unique (with exception of special cases when $V(\cent)$ has several global maxima). However, the long-run outcome is history dependent. An interesting observation is that it is determined only by the spacial average of the initial state. Again, I will give a proof only for a special case when $\acute{A}(\cent)$ has form (11) with $^- = 1$ and $- = [0; 1]$. Let expand $u(x; t)$ in a Fourier series

$$u(x; t) = \sum_{n=\text{¡}1}^{\text{X}} u_n(t) \exp(2\text{¼}inx) \qquad (30)$$

21

where i is the imaginary unit ($i^2 = -1$), and $u_n(t)$ is de...ned by

$$u_n(t) = \int_0^1 u(x;t)\exp(-2\pi inx)dx: \tag{31}$$

Plugging (30) into (13) and collecting the terms before the same exponents results in

$$\frac{du_0}{dt} = f(u_0) \tag{32}$$

$$\frac{du_{3m\pm1}}{dt} = -4D\pi^2 u_{3m\pm1} - u_\pi u_{3m\pm1} \tag{33}$$

$$\frac{du_{3m\pm2}}{dt} = -4D\pi^2 u_{3m\pm1} - u_\pi u_{3m\pm2} + (1 + u_\pi)u_{3m\pm1}^2 \tag{34}$$

$$\frac{du_{3m\pm3}}{dt} = -4D\pi^2 u_{3m\pm3} - u_\pi u_{3m\pm3} + u_{3m\pm1}^3 \tag{35}$$

for any $m \in Z$. Equation (32) implies that $u_0(t)$ will converge generically to one of the local maxima of $V(u_0)$. Moreover, the local maximum to which $u_0(t)$ will eventually converge depends only on $u_0(0)$. Equation (33) implies that $u_{3m\pm1}$ converges to zero, but then equations (33) and (34) imply that $u_{3m\pm2}$ and $u_{3m\pm3}$ converge to zero as well. Hence, any initial distribution of the traits converges to a spatially uniform stationary outcome. Since, according to (31), $u_0(t)$ is the spacial average of $u(x;t)$ at time t, the very long-run

22

outcome is determined by the spacial average of the initial distribution only.

Note that the long-run outcome is spatially uniform, provided payo¤s do not depend on x. This implies that for customs and conventions to di¤er across the locations (for example, for national cultures to exist) in the long-run one has to postulate that payo¤s for a particular trait di¤er across the locations. For example, the traits that facilitate cooperation (e. g. reciprocity) may have more value in a severe climate.

Now let us brie‡y consider the case m = 2. Introduce the polar coordinates (r; Á) on the plane by:

$$x_1 = r \cos Á$$

$$x_2 = r \sin Á;$$

where r is the distance from point x to the center of the switching wave and Á is the polar angle. Symmetry suggests that u = u(r). Let c(R) be the velocity of the circular switching wave with radius R. Then, repeating the derivation that lead to equation (17), one obtains:

$$¡ (c(R) + \frac{D}{r})u^0 = f(u) + Du^{00}:$$

If the boundary that separates areas with different conventions is thin, then $c(R)$ is approximately given by

$$c(R) = c - \frac{D}{R};\qquad(36)$$

where c is the velocity of the switching wave in the one-dimensional system with the same payoff function. For the details, see Markstein (1951). Note, that for

$$R < \frac{D}{c}\qquad(37)$$

the velocity becomes negative. It means that a mutation that occurred within an area contained in the circle of radius D/c will not survive. In particular, a local mutation will not survive. Hence, to cause the switching from a metastable to the stable state, the mutants should not only be sufficiently large in numbers at the particular location, but should also occupy a sufficiently spread area. This phenomenon does not occur in the one-dimensional systems, however it occurs in all dimensions starting from two. It will imply that a custom of even a very populated city will not spread, while a custom of a less populated country or state may.

24

3 SOME EXAMPLES

In this section I am going to give some examples of the application of the general model developed in Section 2.

Example 1 Consider a 2×2 coordination game

$$
\begin{array}{ccc}
 & A & B \\
A & 1;1 & -b;a \\
B & a;-b & 0;0
\end{array}
$$

where $a < 1$ and $b > 0$. The game has two pure strategy equilibria $(A;A)$ and $(B;B)$ and a mixed-strategy equilibrium in which strategy A is played with probability $b/(b + 1 - a)$. For $a + b < 1$ the equilibrium $(A;A)$ is both risk and Pareto dominant, while for $a + b > 1$ the equilibrium $(B;B)$ is risk dominant, while $(A;A)$ is Pareto dominant.

Let $u(x;t)$ be the fraction of the individuals that play strategy A at location x. In the absence of the migration the evolution of u is governed by the replicator dynamics

$$\frac{du}{dt} = u(1 - u)((1 - a + b)u - b). \tag{38}$$

It has three steady states $u = 1$, $u = 0$, and $u = b/(b + 1 - a)$. In the

25

...rst two all the population is coordinating on a pure strategy equilibrium. These steady states are asymptotically stable. The third one corresponds to a heteromorphic population with the share of the population playing strategy A equal to its equilibrium probability, and is unstable. Since both pure strategy equilibria are asymptotically stable the predictions of the replicator dynamics are history dependent. Kandori, Mailath, Rob (1993), and Young (1993) introduced mutations and showed that as the mutation rate goes to zero the risk-dominant equilibrium is selected.

With an exception of a one-time mutants invasion, I will not allow for any mutations taking place on the dynamic path. I will, however, allow for the migration. Assuming that the migration is totally random, I will show that in the very long-run the risk-dominant equilibrium is selected and calculate the speed of the switching wave.

Assuming $m = 1$ equation (5) takes the form

$$\frac{\partial u}{\partial t} = u(1 - u)((1 - a + b)u - b) + D\frac{\partial^2 u}{\partial x^2} \qquad (39)$$

$$D = \gamma(\alpha + (1 - \alpha)(1 - a + b)): \qquad (40)$$

Assume the system switches from $u = 0$ to $u = 1$. Then from (19) the speed

26

of the switching wave is given by

$$c = \frac{1}{2}\sqrt{\frac{D}{(1-a+b)}}(1-(a+b));\qquad(41)$$

Note that $c > 0$ if $a + b < 1$, thus the system in this case indeed switches to $(A;A)$, while if $a + b > 1$ it switches to $(B;B)$. Hence, in the very long-run it switches to the risk-dominant equilibrium. Note, that c is proportional to $1-(a+b)$, that is stronger one equilibrium dominates the other in terms of risk, faster the switch.

Example 2 This example is slightly modi...ed from Basov (2001). Consider a world that consists of a continuum of workers and a ...nite number of ...rms. The ...rms are assumed to be pro...t maximizers. The workers can be of two types: self-interested or reciprocal. The type is ...xed for the life, which consists of two periods.

The ...rms do not observe the type of each worker, but know the distribution of types in the population. They can o¤er two types of contracts: incentive contracts and trust contracts. Given an incentive contract, both types of workers react identically by choosing the optimal e¤ort, which generates zero expected pro...ts for the ...rms and zero expected utility for the

workers. Given a trust contract, a self-interested worker shirks, generating expected monetary payoff U_2 to herself and expected profits B to the firm; and a trustworthy worker exerts an effort, generating expected monetary payoff U_1 to herself and expected profits A to the firm. One can rationalize this behavior assuming that the utility of a worker equals her current monetary payoff, plus the utility from the reciprocation if and only if the worker is reciprocal, and assuming that the time discount factor is zero. If the utility from the reciprocation for the reciprocal workers is large enough the strategies described above are optimal. Assume that $A > 0 > B$ and $U_2 > U_1 > 0$. After the end of the first period the worker may be fired. The probability of being fired is p_1^F if the worker did not shirk, and p_2^F if she did. Assume $1 > p_2^F > p_1^F > 0$, and that all firms observe whether the worker was fired in the first period before offering the second period contract. Finally, define

$$\Delta U = (2 - p_1^F)U_1 - (2 - p_2^F)U_2;$$

and assume that $0 < \Delta U < p_2^F U_2 - p_1^F U_1$.

Let $u(x; t)$ be the share of the reciprocal workers at location x and time t. Assume and that their fitness payoff equals the undiscounted sum of their

monetary payo¤s. Then the evolution of the share of the reciprocal workers at a particular location is governed by:

$$\frac{@u}{@t} = u(1 \text{ ¡ } u)\text{Á}(u) + {}^\circ div({}^\circledR r\, u + (1 \text{ ¡ } {}^\circledR)\text{Á}_u r\, u); \qquad (42)$$

where

$$
\text{Á}(u) =
\begin{cases}
0, & \text{if } u \in [0; \overline{u}] \cup \{u^\text{¤}\}; \\
\text{¢}U, & \text{if } \overline{u} < u < u^\text{¤}; \\
2(U_1 \text{ ¡ } U_2), & \text{if } u^\text{¤} < u \cdot 1;
\end{cases}
$$

where $\overline{u} = jBj = (A + jBj)$, $u^\text{¤} = jBj\, p_2^F = (Ap_1^F + jBj\, p_2^F)$.

Note that while the replicator dynamics has a continuum of steady states, the only strict maximum of V (u) and hence, the unique asymptotically stable state is $u^\text{¤}$. The result holds irrespectively of m and $^\circledR$.

Basov (2001) obtained a similar result considering noisy replicator dynamics a la Foster and Young (1990). See also Basov (2001) for a discussion of the properties of this solution. However, the behavioral assumptions necessary to get the noisy replicator dynamics are not trivial. For a discussion,

29

see Corradi and Sarin (2000). In my view, the equilibrium selection argument based on the migration is more persuasive.

The speed of switching wave in this case cannot be calculated analytically, but for small $U_1 - U_2$ it will be proportional to $V(u^\pi) - V(\overline{u})$.

4 DISCUSSION AND CONCLUSIONS

In this paper I developed a model of the replicator dynamics with migration. The model can be used to study the equilibrium selection in games. For 2×2 coordination games the selection criterion coincides with the one obtained in Foster and Young (1990), KMR (1993), and Young (1993). The model also allows us to raise some new questions. For example, it allows us to calculate the speed of the propagation of a new behavior. That is, it allows to ask, how long will it take for some custom conceived at California to reach Boston. Note, that this question is different from how long will it take for all the population to adopt a new custom. The last question received a considerable attention in the literature (see, for example, Young 1999 and Ellison 2000), while this one is, to my knowledge, new.

I was also able to demonstrate that in two-dimensional world, for a new

convention to spread it is not only necessary, that sufficient number of people adopt it (it has a big size), but they should also be sufficiently spread. For example, if a ten-million city, say New York, adopts a new driving convention, it is unlikely that such a convention would survive. However, if a ten-million country or state adopts one it may survive for a while. Should the future practice show that it is intrinsically better (for example, causes fewer accidents) it can spread world-while. The distinction between the size and the spread is impossible to draw in a network model. It makes perfect sense, however, within the model of this paper.

Why the spread matters for if there is more than one spacial dimension, but does not in a one-dimensional world? The driving convention example suggests a simple intuition. If a city adopts a new driving convention in a two-dimensional world everybody else travelling between two different locations can simply avoid driving through the city. As long as the city is not too spread this will not be too costly. This is, however, impossible in the one-dimensional world.

Another interesting conclusion is that local differences in behavior (national cultures) cannot be explained simple by postulated that the social interaction is adequately captured by a game with multiple equilibria. For

31

customs and conventions to diﬀer across the locations (for example, for national cultures to exist) in the long-run one has to postulate that payoﬀs for a particular trait diﬀer across the locations. This justiﬁes the interest of the historians to the physical conditions that dominated a nation's history.

One might wonder, how the results of this paper generalize, if there are more than two traits. In this case an equation similar to (5) still holds, but it might be impossible to construct the functional (6). The main diﬃculty is that $f(u; x)$ in general need not be a gradient of some function. If it is, all the results of Section 2 go through, otherwise the analysis becomes more complicated. One can still write the analog of equation (17), use it to ﬁnd the velocity of the switching wave c, and make an equilibrium selection argument based on the direction of c. However, the selected equilibrium will no more be the global maximum of some function, and c can be found only numerically. However, some important conclusions do generalize to this case. The speed of the spread of the conventions is still determined endogenously and is proportional to the square root of openness, and for a convention to survive in at least two spatial dimensions it should have a minimum spread.

References

Albeverio, S., J. E. Fenstad, R. Høegh-Krohn, and T. Lindstrøm (1986): Nonstandard Methods in Stochastic Analysis and Mathematical Physics, Academic Press, Orlando, FL.

Bergstrom, T. C. (1995): "On Evolution of Altruistic Rules for Siblings," American Economic Review, 85, 58-81.

Anderlini, L. and A. Ianni. (1996): "Path Dependence and Learning from Neighbors," Games and Economic Behavior 13, 141-177.

Baptista, R. (2000): "Do Innovations Di¤use Faster with Geographical Clusters," International Journal of Industrial Organization 18, 515-535.

Basov, S. (2001): "An Evolutionary Model of Reciprocity," The University of Melbourne, Department of Economics, Research Paper # 812.

Bugaenko, L. T., M. G. Kuzmin, and L. S. Polak (1993): High Energy Chemistry, E. Horwood, New York.

Blume, L. (1993): "The Statistical Mechanics of Strategic Interaction," Games and Economic Behavior 5, 387-423.

Blume, L. (1995): "The Statistical Mechanics of Best Response Strategy Revision," Games and Economic Behavior 11, 111-145.

Corradi, V., and R. Sarin. (2000). "Continuous Approximation of Stochastic

Evolutionary Game Dynamics," Journal of Economic Theory, 94, 163-191.

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," Econometrica 61, 1047-1071.

Ely, J. (1995): "Local Conventions," mimeo, University of California, Berkeley.

Foster, D., and P. Young. (1990): "Stochastic Evolutionary Game Dynamics," Theoretical Population Biology, 38, 219-232.

Kandori, M., G. Mailath, and R. Rob (1993): "Learning, Mutation and Long Run Equilibria in Games," Econometrica, 61, 29-56.

Keisler, H. J.(1978): An In...nitesimal Approach to Stochastic Analysis, mimeo, MIT.

Markstein, G. H. (1951): "Experimental and Theoretical Studies of Flame Front Stability," Journal of Aeronautic Science, 18, 199-201.

Rees, J., R. Briggs, and R. P. Oakey. (1984): "The Adoption of New Technologies in the American Machinery Industry," Regional Studies, 18, 489-504.

Robson, A. J. (1996): "The Evolution of Attitudes to Risk: Lottery Tickets and Relative Wealth," Games and Economic Behavior, 14, 190-207

Samuelson, L. (1997): Evolutionary Games and Equilibrium Selection, MIT Press.

Schlag, K. H. (1998): "Why Imitate and if So, How? A Boundedly Rational Approach to Multiarmed Bandits," Journal of Economic Theory, 78: 130-56.

Thwaites, A. (1982) "Some Evidence of Regional variations in the Di¤usion of new Industrial Products and Processes within British Manufacturing Industry," Regional Studies, 16, 371-381.

Trevelyan, G. M. (1937): History of England, Longmans, Green and Co.

Vulikh, B. Z. (1963): Introduction to Functional Analysis for Scientists and Technologists, Pergamon Press.

Young, P. (1993). "The Evolution of Conventions," Econometrica, 61, 57-84.

Young, P. (1999): "Di¤usion in Social Networks," Center on Social and Economic Dynamics, Working Paper #2.

APPENDIX 1

In this Appendix I present a slightly modified version of the Aspiration and Imitation model (Samuelson 1997, Ch. 3). The only deviation from the original model is that I assume that the population is hyperfinite rather then finite. I show that in this the share of the population following a particular strategy is governed by the replicator dynamics. The result is exact in contrast with the finite population case where it is obtained only in the limit. For a definition of the hyperfinite numbers and the other notions of the nonstandard analysis see, for example, Albeverio, Fenstad, Høegh-Krohn, and Lindstrøm (1986) and Appendix 2 to this paper.

The nonstandard analysis allows for the existence of positive infinitely small and finite infinitely large numbers. Intuitively, a hyperfinite number is an infinite natural number. The value of using the nonstandard techniques is that they allows us to have an exact law of large numbers without running into the measurability problems (Keisler, 1987).

Consider the following symmetric 2 £ 2 game

$$
\begin{array}{c|cc}
 & X & Y \\
\hline
X & A; A & C; B \\
Y & B; C & D; D
\end{array}:
$$

Let there be a single population containing N players, where N is hyper...-nite.[3] Time is divided into discrete intervals of an in...nitesimal length $\iota > 0$. In each period a player is characterized by a pure strategy X or Y she is currently playing. If a player with strategy S_1 (S_1¡ strategist) is matched against the player with a strategy S_2 (S_2¡ strategist) then the realized payo¤ is the sum of the payo¤ shown in the matrix above and a realization R of a random variable $\mathbb{R}$ with zero mean. This random variable captures the factors outside the model that might a¤ect the players' payo¤s.

If a player plays strategy X in a population where proportion k of her opponents play X, then her expected payo¤ is given by

$$\tfrac{1}{4}_X(k) = kA + (1 ¡ k)C:$$

---

[3]Since the external cardinality of any hyper...nite set is continuum (Albeverio, Fenstad, Høegh-Krohn, and Lindstrøm 1986) we are actually speaking about a population with a continuum of individuals.

The expected payoﬀ to a Y_i strategist in the same population is

$$\mu_Y(k) = kB + (1 - k)D.$$

In each period of length ι, each player takes a draw from an independently distributed Bernoulli random variable. Assume that a draw "learn" is produced with probability ι. If the player receives a learn draw, she abandons her current strategy if and only if her realized payoﬀ is below her aspiration level, $\mathbb{C}$. If player i has abandoned her strategy, she must now choose a new strategy. For this purpose she randomly selects a member j of the population and imitates player's j strategy. This rule determines a Markov process.

Let $z(t)$ denote the proportion of X_i strategists in the population at time t. To proceed further I need the following deﬁnition

Deﬁnition 1 Let t be a standard real number. The derivative of $z(t)$ is deﬁned as

$$\frac{dz}{dt} = St(\frac{z(t + \iota) - z(t)}{\iota}); \qquad (A.1)$$

where $\iota \neq 0$ is inﬁnitely small.

Here $St(\cent)$ is the standard part operator. For each hyperreal number its standard part is the unique standard (usual) real number such that it di¤ers from the hyperreal in question by an in…nitely small quantity. One can show that if $z(t)$ is di¤erentiable in the usual sense the right hand side of (A.1) does not depend on ι and the derivative de…ned in this way coincides with the classical derivative.

To derive the equation governing the evolution of $z(t)$ …rst note that the probability a player abandons her strategy is given by

$$g(\mu) = F(\cent_{\textellipsis}\ \mu),\tag{A.2}$$

where μ is the expected payo¤ her strategy gives and $F(\cent)$ is the c.d.f. of random variable R. I will assume that R is determined uniformly on $[_{\textellipsis}\ !\ ;\ !\ ]$, where for some large …nite $!$. Let $^{\circ} = 1{=}N$. Then at each moment of time $z(t)\ 2\ f0;\ ^{\circ};\ 2^{\circ};\ :::;\ 1g$. The state of the population is characterized by the value of $z(t)$. The change in the share of $X_{\textellipsis}$ strategists during a single period can be written as number of $Y_{\textellipsis}$ strategists becoming $X_{\textellipsis}$ strategists, F_{in}, minus the number of $X_{\textellipsis}$ strategists becoming one strategists, F_{out}. A $Y_{\textellipsis}$ strategist becoming an $X_{\textellipsis}$ strategist if three thinks happen: she receives

39

a learn draw, she has to decide to abandon her current strategy, and she should select an X_i strategist from the population. Taking into account that the law of large numbers holds exactly in the hyper...nite setting (Keisler, 1987) the in‡ow of X_i strategists can be written as:

$$F_{in} = z(1 - z)N \cdot g(\pi_X(z)):$$

Following a similar logic

$$F_{out} = z(1 - z)N \cdot g(\pi_Y(z)):$$

Hence,

$$N(z(t + \iota) - z(t)) = z(1 - z)N \cdot (g(\pi_X(z) - g(\pi_Y(z)):$$

Using (A.2) and De...nition 1

$$\frac{dz}{dt} = \frac{z(1 - z)}{2!}(\pi_X(z) - \pi_Y(z)):$$

Rescaling time to eliminate the 2! we obtain the replicator dynamics.

APPENDIX 2

It is not the aim of this Appendix to develop a complete theory of hyper-real numbers. An interested reader can address herself to Albeverio, Fenstad, Høegh-Krohn, and Lindstrøm (1986). Here I want just to persuade the reader that the hyperreal numbers are no more mysterious than the reals and hence, all the calculations based on the nonstandard analysis are perfectly sound.

One of the ways to construct the reals R from the rationals Q is to add to Q new points to represent limits of the fundamental sequences of real numbers (see, for example, Vulikh 1963). To get the hyperreals from the reals one has to follow the same procedure but be more accurate in identifying sequences. Intuitively, we care not only about the limit, but also about the rate of convergence and the asymptotic properties.

We would certainly like to identify two sequences that agree everywhere but on a ...nite set of indexes. Such sequences clearly have the same asymptotic properties and it will be desirable to have them in the same equivalence class. Call the corresponding equivalence relation %. It would be a natural attempt to identify hyperreals with the equivalences classes of %. However, if we identify a hyperreal number with the set of equivalence classes of relation % then the set of hyperreals will contain the divisors of zero. Consider, for

41

example, two sequences $\{x_n\}_{n=0}^{\infty}$ and $\{y_n\}_{n=0}^{\infty}$. Assume that $x_n = 1$ if n is even and $x_n = 0$ if n is odd. The reverse is true for $\{y_n\}_{n=0}^{\infty}$. Neither of these sequences is in the same equivalence class as $\{z_n\}_{n=0}^{\infty}$, where $z_n = 0$ for any n. However, $x_n y_n = 0$. To avoid this problem we have to redefine the equivalence relation in such a way that either $\{x_n\}_{n=0}^{\infty}$ or $\{y_n\}_{n=0}^{\infty}$ becomes equivalent to zero. That is, not only the sequences that differ on the finite set of indexes are equivalent, but so are some sequences that differ on an infinite set of indexes.

Formally, let N be the set of natural numbers.

Definition 2 $F \subseteq 2^N$ is called a filter if

1. $N \in F$, $\emptyset \notin F$.

2. $A_1 \in F$, $A_2 \in F$ implies $A_1 \setminus A_2 \in F$.

3. $A \in F$, $A \subseteq B$ implies $B \in F$.

It is easy to check that if F is the family of all subsets of N with a finite complement then F is a filter. Call this filter F_0.

Call two sequences $\{u_n\}_{n=0}^{\infty}$ and $\{v_n\}_{n=0}^{\infty}$ equivalent with respect to F or $F-$ equivalent if

$$\{n \in N : u_n = v_n\} \in F:$$

42

It is easy to see that neither $\{x_n\}_{n=0}^{\infty}$ nor $\{y_n\}_{n=0}^{\infty}$ is not F_0-equivalent to $\{z_n\}_{n=0}^{\infty}$. The problem hence, is to find an extension F_u of filter F_0 in such a way that at least one of these sequences is equivalent to $\{z_n\}_{n=0}^{\infty}$ with respect to F_u.

Definition 3 A filter F_u is called a free ultrafilter if

1. F_u contains no finite sets.

2. For any $E \subseteq N$ either $E \in F_u$ or $N{=}E \in F_u$.

It can be shown (see, Albeverio, Fenstad, Høegh-Krohn, and Lindstrøm 1986) that there exists a free ultrafilter F_u such that $F_0 \subseteq F_u$. Clearly, either $\{x_n\}_{n=0}^{\infty}$ or $\{y_n\}_{n=0}^{\infty}$ should be equivalent to zero with respect to such an ultrafilter, since according to the second part of definition it should contain either odd or even numbers.

The ultrafilter extending F_0 is not unique. Fix such an ultrafilter and call two sequences of reals equivalent if they are F_u-equivalent. The set of equivalence classes are called hyperreals. One can define all usual arithmetic operations and the order relation on this set. The process is similar to the extension of the arithmetic operations and the order relation from the rationals to the reals using the construction discussed in the beginning of this Appendix.

It is easy to see that this construction allows for the existence of the positive in...nitely small reals[4]. Consider, for example, the equivalence class of the sequence $\{u_n\}_{n=0}^1$, where

$$u_n = \frac{1}{n+1}:$$

Call the corresponding hyperreal ". Then " $\neq$ 0. Indeed, for " to be equal to zero it is necessary for set $Z = \{n \in N : u_n = 0\}$ to be in...nite, since F_u is free. But $Z = \,;$. One can check, however, that k" < 1 for any $k \in N$. Naturally, 1=" > k for any $k \in N$, namely 1=" is in...nitely large.

---

[4]A positive number $\pm$ is called in...nitely small if $m\pm < 1$ for any $m \in N$.