ON THE GEOGRAPHY OF CONVENTIONS[1]

Andreas Blume

Department of Economics, University of Iowa

Iowa City, IA 52242

E-mail: andreas-blume@uiowa.edu

Ted Temzelides

Department of Economics, University of Iowa

Iowa City, IA 52242

E-mail: tedt@blue.weeg.uiowa.edu

First version: October 1998

This version: August 1999

Abstract

We study a model in which heterogenous boundedly rational agents interact locally in order to play a coordination game. Agents differ in their mobility with mobile agents being able to relocate within a country. The model yields the following predictions: (1) mobile agents always benefit from increased mobility, (2) immobile agents benefit from increased mobility at low levels of mobility, (3) immobile agents lose from increased mobility at high levels of mobility, (4) there is an optimal "country size," (5) "income inequality" is weakly increasing in mobility, (6) if there are arbitrarily small payoff differences between two countries, opening borders causes a "brain drain" effect; in the long run, all mobile agents reside in the favored (former) country and efficiency is attained only in that country.

[1]We thank George Mailath for helpful discussions.

I. Introduction

Economists have long studied the effects of location and mobility on economic decisions by individual consumers and firms. Why is it, for example, that different, often seemingly Pareto ranked, conventions coexist among different locations? What are the effects of regulation that restricts/allows mobility across locations? Answers to these questions are relevant for the study of issues such as crime, unemployment, growth, etc.

Here we study a model where boundedly rational agents are randomly matched repeatedly in order to play a coordination game with tension between risk dominance and Pareto dominance. This is perhaps the simplest stylized framework in which we can formally study the effects of mobility and of exogenous restrictions on mobility on the agents' choices of locations and actions. We assume that agents are heterogeneous, differing in their mobility, where mobility is a function of the number of locations that agents have access to. Mobile agents can move across a finite set of locations the collection of which we shall call a "country." The remaining agents are immobile and cannot move across locations. Our model yields a number of predictions about the effects of different arrangements regarding mobility levels and the corresponding set of long-run outcomes. Mobile agents weakly benefit from increased mobility. If country size is small, i.e. mobility of the mobile agents is restricted to a small number of locations, then the risk-dominant equilibrium obtains at every location. In contrast, if country size is sufficiently large, there are enough mobile players to ensure efficient play at one of the locations and all mobile agents will be at that location. Immobile agents benefit from increased mobility at low levels of mobility. Country size has to be sufficiently large to ensure efficient play at one location. In that case, the immobile agents at that location benefit from the externality generated by the presence of the mobile ones. Immobile agents lose from increased mobility at high levels of mobility. The reason is that any given immobile player is ex ante less likely to benefit from the above mentioned externality if country size is increased further. As measured by the difference in realized payoffs, income inequality is weakly increasing in mobility. In this sense, there is an optimal country size since allowing every mobile agent to move freely reduces the fraction of immobile agents that benefit from the externality. Finally, if there are arbitrarily small payoff differences between two countries that favor one country, opening borders causes a "brain drain" effect; in the long run, all mobile agents reside in the favored (former) country and efficiency is attained only in that country. Note that this is true even if the risk-dominant equilibrium in the favored country is dominated by the Pareto-dominant equilibrium in the less well off country. The result that the mobile agents move to the favored country is *not* a direct consequence of the payoff differential. Rather, it results from the greater robustness of efficient play in the favored country.

There are many possible origins of restrictions on geographical mobility. They could be due to the conventional political or legal constraints on labor mobility across countries. They could be skill based, when skilled workers are employable in a larger number of communities. They could be the result of racial or gender discrimination. Or, they could be the result of strong community ties. Finally, we may think of differential mobility as differential access to communication resources.

Under the conventional political interpretation, one could say that "globalization," i.e., a removal of cross-border restrictions on mobility, favors mobile agents. They seek out opportunities and they form stable clusters, at which the efficient outcome prevails. Those immobile agents who happen to be at the location where such an efficient cluster forms, effectively benefit from an externality generated by the mobile ones. Ex ante, the efficient cluster is equally likely to form at any location that the mobile agents have access to. Moreover, in the long run, the efficient cluster will spend equal amounts of time at any of those locations. In that sense, all the immobile agents benefit from the externality. A precondition for the efficient cluster to form in a country is that there are sufficiently many mobile agents in that country, which in our model is a function of country size. Thus the externality that benefits the immobile agents requires there to be sufficient mobility. Once there is enough mobility to put the externality in place, however, further increases in mobility reduce the frequency with which the externality is encountered by a given immobile agent. In a sense, greater mobility allows mobile agents to emigrate and leave immobile agents behind who on their own are unable to sustain efficient outcomes. This immigration/emigration interpretation can be made somewhat tighter by considering one-way permeability of borders: immobile agents in immigrant (emigrant) regions benefit (lose).

Our model predicts that greater mobility shifts the "income distribution" in favor of the mobile agents. In light of the different origins of geographic mobility, this can result in a bias in favor of skilled agents, against agents who suffer from discrimination, against agents with cultural biases, or against agents with limited access to communication resources.

The important role of externalities for the consequences of migration has been recognized in the literature on migration; see for example the survey by Greenwood [1975]. Greenwood dates this recognition to the application of the theory of human capital to the analysis of the international brain drain. That literature points of the selectivity of migration, the fact that younger, better educated workers are more likely to migrate. As a result, one might expect receiving regions to benefit and sending regions to lose from migration. We get similar effects in our model without selectivity, as the mobile agents at a given location have the same strategic options and the same payoffs from their available actions as the immobile agents. Rather than selectivity of migration, it is the mobility of the migrating agents itself that confers benefits on the receiving region or country. The reason is that mobile players move to locations at which efficient equilibrium play arises by chance. This increases the number of agents at that location and thus makes it less likely that agents' mistakes will upset the efficient equilibrium at that location.

More recently Akerlof [1997] has noted the pervasive role of externalities in social decisions, as studied in the social interactions literature (see the references in Akerlof). He notes that empirically, group effects have been found to be important (see the references in Akerlof) and, in order to overcome identification problems in the econometric work on this issue, presents additional ethnographic and biographical evidence. He observes that group loyalty may limit mobility and that group conformity may impede social advancement.

Our paper uses techniques from evolutionary game theory. In particular we shall often make use of the concepts of the *radius*, *coradius*, and *modified coradius* of basins of attraction of a dynamical system, as introduced by Ellison (1995) who built on Kandori, Mailath, and Rob (K-M-R,1993) and Young (1993) who, in turn, used techniques in Freidlin and Wentzell (1984). We use an adaptation of these techniques in one of our proofs that might prove useful in other applications. Roughly, the radius of a basin of attraction of a set is the cheapest way out of the set, and the coradius is the cheapest way into the set from the worst starting state. Usually one uses these concepts to prove stochastic stability of a set by showing that for that set the radius exceeds the coradius (or, if necessary, a modified version of the coradius). We use this argument repeatedly. We show for two different sets that the radius exceeds the (modified) coradius and conclude that the stochastically stable set must be in the intersection of those two sets.[2]

The role of mobility in local interaction models has been explored by Bhaskar and Vega-Redondo (1996), Ely (1996), and Oechsler (1997). These papers show that unlike in random matching models, where the risk-dominant equilibrium is selected, local interaction favors the Pareto-dominant equilibrium. The consequences of restricted mobility in these models have been investigated by Anwar (1996) and Dieckmann (1999). Like the present model, Anwar and Dieckmann demonstrate that restricted mobility may lead to coexistence of different conventions. One of our innovations is to study heterogeneity in mobility and the distributional effects of such mobility.

II. The Model

There is one continent and k locations. A *country* is a collection of locations. Bigger countries are larger sets of locations. As part of the exercise, we will vary the number (size) of countries but we always assume that all countries are of equal size. The total population of the continent is

[2] Ellison (1995) studies a simple stylized example by using a similar argument.

kM agents. Countries are symmetric and it is assumed that they do not overlap.[3] Without loss of generality assume that each location has initially M agents. A number, n, of the M agents are immobile. The remaining $m = M - n$, are mobile. The agents are randomly matched in pairs to play a 2x2 symmetric coordination game. Action profile (s_1, s_1) or (G) corresponds to the payoff dominant equilibrium and action profile (s_2, s_2) or (B) corresponds to the risk dominant (and in this case Pareto inferior) one.

In each period, each agent chooses a location and an action. The choice of action is subject to noise while the choice of location is not.[4] Of course, the choice of location is trivial for the immobile agents. In each period, mobile agents will move to the location that gives them the highest payoff among all locations available to them and given the state in the previous period. If indifferent, mobile agents choose a location at random. The mixed equilibrium is given by $(xs_1, (1 - x)s_2)$, where $x \in (\frac{1}{2}, 1)$. Thus, (s_2, s_2) is the risk dominant action choice.

Let - ^{P} be the state space, where P denotes the set of countries with generic element ρ, and let - be the restriction of the state space to a single country. A (restriction of the) state in - is $\omega = (z, f)$, where z is a vector describing the fraction of s_1 players in each location (this does not distinguish between mobile and immobile agents but this is without loss of generality), and f is a vector describing the number of mobile agents in each location. The immobile agents stay put. Regarding choice of location of the mobile agents, the underlying best reply dynamic is assumed.[5] Let k^ρ stand for the location(s) that offers the highest expected payoff in country ρ. Given that the present state is ω, the distribution of mobile agents across locations in country ρ is given by $b^\rho(\omega) \in \Delta(k^\rho)$, where $\Delta(k^\rho)$ is the set of all possible distributions of mobile agents across the k^ρ locations. We assume that each element of $\Delta(k^\rho)$ has positive probability.

Let $u(s_\gamma, k, \omega)$ be the expected payoff of action s_γ in a given location, k, given that the state in the previous period was ω. Regarding choice of action, the stochastic best reply dynamic a la

[3] However, see Section III.C for an example with overlapping populations.

[4] Introducing noise in the choice of location will not fundamentally alter our results.

[5] Our results hold for a more general class of monotone dynamics.

K-M-R (1993) is assumed. Denoting the fraction of agents choosing action s_1 at location k in response to state ω by $b^k(\omega)$, we have:

$$b^k(\omega)\begin{cases} = 1, & if \quad u(s_2,k,\omega) < u(s_1,k,\omega); \\ \in \Delta(\{s_1,s_2\}), & if \quad u(s_2,k,\omega) = u(s_1,k,\omega); \\ = 0, & if \quad u(s_2,k,\omega) > u(s_1,k,\omega). \end{cases}$$

The function $z_{t+1} = b(\omega_t)$ gives the profile of strategies that will be used in each location at $t+1$, given that the time t state is ω_t. Since the stage game is a coordination one, the dynamic described above leads to a set of states where everyone plays an identical action, G or B, in each location. Actions could vary across locations. The long run behavior of the system depends on the initial distribution of agents across states. This is resolved if noise is introduced into the system. Assume that with probability ϵ, each player *trembles* and chooses an action from a full support distribution. These choices are assumed to be *iid* across players and time. This yields a stochastic dynamical system that defines a Markov chain on the finite state space - ^{P}. This Markov chain has a unique invariant distribution for a given rate $\epsilon > 0$. The invariant distribution is globally stable and ergodic and is interpreted as giving the proportion of time that the society spends in each state. We will concentrate on the support of the distribution μ, which results as the limit case when ϵ is driven to zero.

In this paper we are interested in studying the effects of exogenous policies that restrict mobility on the long run payoffs of agents in an environment where agents are heterogenous in their mobility levels. As mentioned earlier, mobile agents might create an externality that is enjoyed by themselves, as well as by a subset of the immobile agents. We shall, proceed by studying how equilibrium selection might change as a function of different restrictions on the level of the externality. For now, we do not consider interactions across countries. Therefore we will focus on a single country. The results extend trivially to the full state space. An element ω of - describes a distribution of agents across locations (within a country) and actions. We first fix the size of a country, by assuming that each country contains i locations where $i \in D_k$, where D_k is the set of divisors of k, and concentrate on the set of states $\tilde{\Omega} \subset \Omega$ that have the following three properties:

(1) all mobile agents within the country are piled in one location;

(2) action G is played in that pile;

(3) action B is played in each of the remaining locations.

We will make extensive use of techniques developed in Ellison (1995). For definitions and additional applications the reader is referred to Fudenberg and Levine (1998). The characterization of the long run predictions of the model (the stochastically stable outcomes) relies on the calculation of the *radius* and the *coradius* of the basin of attraction of a family of sets. While the formalization of these concepts requires the use of some mathematical notation, they are very intuitive to grasp. Suppose the system is in an absorbing set A. The *radius* of the basin of attraction of A corresponds to the minimum number of trembles necessary to leave the basin of attraction. Next we need to compute the minimum number of trembles needed to reach the basin of attraction of A, starting from an absorbing set outside A. Do the same for all other absorbing sets outside A, and determine the maximum of these numbers. This number is the *coradius* of the basin of attraction of A. Ellison derives a sufficient condition for an absorbing set to be uniquely selected by the learning process: If the radius of the basin of attraction of an absorbing set A is larger than its coradius, then all stochastically stable sets are contained in A.

We are interested in absorbing sets of the process, where play settles down to a stationary distribution of choices of locations and actions. Let $P_{\omega\omega'}$ denote the probability of transition from state ω to state ω'.

Definition 1 *A set of states Ω' is absorbing if (i) for all $\omega' \in \Omega'$, $\omega \notin \Omega'$, $P_{\omega'\omega} = 0$, and (ii) $\nexists \Omega'' \subset \Omega'$, $\Omega'' \neq \Omega'$ such that (i) holds for Ω''.*

The first condition requires that once the process enters the absorbing set, it will not leave it. The second condition requires that absorbing sets are minimal. Let A be a subset of the set of absorbing sets of the model without noise. The *basin of attraction* of A, denoted by $D(A)$, is the set of all states from which the unperturbed Markov process converges to a state in A with probability one,

$$D(A) = \{\omega \in - \mid \Pr(\exists \tau' \text{ such that } \omega^\tau \in A \; \forall \tau > \tau' | \omega^0 = \omega) = 1\}.$$

The *radius* of the set $D(A)$ is the number of trembles necessary to leave the set, starting from a state in A. Let $c(\omega, \omega')$ be the number of trembles needed for the system to transit from state ω to state ω'. That is, $c(\cdot)$ measures the *transition cost* between these states. Define a *path* by a finite sequence $(\omega^1, \omega^2, \ldots, \omega^k)$ of distinct states. The *cost* of such a path is defined by

$$c(\omega^1, \omega^2, \ldots, \omega^k) = \Sigma_{\tau=1}^{k-1} c(\omega^\tau, \omega^{\tau+1}).$$

The *radius* of A is the least costly (in terms of trembles) path leading from any state in A to some state outside the basin of attraction of A.

Definition 2 *The radius of the basin of attraction of a set of absorbing sets A is:*

$$R(A) = \min_{(\omega^1, \ldots, \omega^k)} c(\omega^1, \ldots, \omega^k) \text{ such that } \omega^1 \in A, \omega^k \notin D(A).$$

The path $(\omega^1, \ldots, \omega^k)$, defining the radius of $D(A)$, thus describes the cheapest way out of that set. The *coradius* of the basin of attraction of a set of absorbing sets is defined by the number of trembles necessary to reach this set from the state where the minimum number of trembles required to reach $D(A)$ is maximized.

Definition 3 *The coradius of the basin of attraction of a set of absorbing sets A is:*

$$CR(A) = \max_{\omega^1 \notin A} \min_{(\omega^1, \ldots, \omega^k)} c(\omega^1, \ldots, \omega^k) \; s.t. \; \omega^k \in D(A).$$

The smaller the coradius, the likelier is the event that simultaneous trembles shift the system from any absorbing state to some state in $D(A)$. We will make use of the following concept that is related to the coradius.

Definition 4 *The modified coradius of the basin of attraction of a set of absorbing sets A is:*

$$CR^*(A) = \max_{\omega^1 \notin A} \min_{(\omega^1,\ldots,\omega^k)} \{c(\omega^1,\ldots,\omega^k) - \sum_{l=2}^{L-1} R(-_l)\} s.t.\ \omega^k \in D(A)$$

and $\{-_l\}$ *is the sequence of absorbing sets through which* $(\omega^1,\ldots,\omega^k)$ *passes.*

A sufficient condition for a set of states to be stochastically stable is that the radius of its basin of attraction exceeds the modified coradius. Intuitively, while, in the presence of noise, all states in the system are transient, the stochastically stable states are the ones where the system will spend the largest fraction of its time in the "long run." We will, thus, concentrate our attention to such states. Formally, let $\mu(\epsilon)$ denote the unique ergodic distribution associated with the Markov chain, and define the limit distribution by $\mu^* = \lim_{\epsilon\to 0} \mu(\epsilon)$. The *limit set* is defined by

$$-^* = \{\omega \in - : \mu^*(\omega) > 0\}.$$

Define $\mu^*(A) = \Sigma_{\omega\in A}\mu^*(\omega)$, with $\mu^*(-^*) = 1$. We have the following:

Theorem 1 *(Ellison, 1995) For any absorbing set* A*, if* $R(A) > CR^*(A)$*, then* $\mu^*(A) = 1$.

The result obviously holds if we substitute $CR(A)$ for $CR^*(A)$. Roughly, the result says that a set A contains the stochastically stable states if it is more difficult to leave the set than it is to enter it from the worst possible starting point.

Let $R_i(\Theta)$, $CR_i(\Theta)$, and $CR_i^*(\Theta)$ stand for the radius, coradius, and the modified coradius of a set of states $\Theta \subseteq -$ when mobile agents are free to locate across i locations. The following gives a sufficient condition for $\tilde{-}$ to be selected. The proof proceeds by first demonstrating that $\tilde{-}$ is the unique set belonging in the intersection of two appropriately defined sets of states, $\tilde{-} \cup -_1$ and $\tilde{-} \cup -_2$. It is then shown that $R_i(\tilde{-} \cup -_1) > CR_i^*(\tilde{-} \cup -_1)$ and that $R_i(\tilde{-} \cup -_2) > CR_i(\tilde{-} \cup -_2)$. This, in turn, implies that $\tilde{-}$ contains the unique stochastically stable set.

Our long run predictions do not depend on initial conditions. Without loss of generality, we shall assume that mobile and immobile agents are assigned symmetrically across locations in period 0 (see Figure 1, top). The next proposition describes a sufficient condition for the stochastically stable set to be a subset of $\tilde{-}$. Figure 1 (bottom) illustrates the outcome predicted by the proposition. In each of the four countries, the mobile players cluster in one location. At this location, the efficient equilibrium, G, is played. Everywhere else, the risk-dominant equilibrium B prevails. It is worth mentioning that, absent mobility across locations, each individual location would correspond to a K-M-R model, resulting in the risk dominant equilibrium being played everywhere. In our case, the mobile agents are able to create a sufficient cluster in which the efficient outcome is played. At the same time, the immobile agents in the location where the G pile is formed benefit from the externality..

Proposition 2 *Fix the number of accessible locations, $i > 1$, for all mobile agents. Then the stochastically stable set is a subset of $\tilde{-}$ if $x \in \left(\frac{1}{2}, \frac{n+im}{2n+im}\right)$.*

Proof. Let $\tilde{-}$ be the set of states where all mobile agents are piled in one location and G is played at that location and all immobile agents play B everywhere else. Let $-_1$ be the set of states where in every country all agents, mobile and immobile, play (B) and where mobile agents are located arbitrarily across locations. Finally, let $-_2$ be the set of states where mobile agents are located arbitrarily across locations and only one action is played in any given location (actions and the size of piles could vary across locations but we exclude $\tilde{-}$ and $-_1$ from this set). These three sets include all absorbing states. The cheapest way out of $\tilde{-} \cup -_1$ is to move from $\tilde{-}$ to $-_2$. This involves (any) one pile of immobile agents switching from B to G. Therefore, we have that $R_i(\tilde{-} \cup -_1) = xn$. To calculate the modified coradius of $\tilde{-} \cup -_1$, we only need to consider paths starting in $\tilde{-} \cup -_2$. The starting point giving rise to the most expensive such transition is the one where all agents play G at all locations. Note that from there we can reach a state where all mobile agents are piled at a single location at zero cost. From there, the minimal cost path into $\tilde{-} \cup -_1$ involves $(1-x)n(i-1)$ trembles; this is the minimal number of trembles needed to switch play at the $i-1$ locations that are only inhabited by immobile players from G to B. The modified coradius is this cost minus the radii of the intermediate steps. Each of these equals $(1-x)n$, and there are $i-2$ intermediate steps. Hence, $CR_i^*(\tilde{-} \cup -_1) = (1-x)n(i-1) - (1-x)n(i-2) = (1-x)n$.

Therefore, $R_i(\tilde{-} \cup -_1) > CR_i^*(\tilde{-} \cup -_1)$ iff $xn > (1-x)n$, or iff $x > \frac{n}{2n}$. The cheapest way out of $\tilde{-} \cup -_2$ is to move from $\tilde{-}$ to $-_1$. This involves the big pile of mobile agents playing G to switch to B. We thus have that $R_i(\tilde{-} \cup -_2) = (1-x)(n+im)$. On the other hand, if every agent in every location plays B, the state where all mobile agents are piled in one location is reached with positive probability. Then, xn trembles are needed to lead to immobile agents in one location to change to G and, subsequently, have all mobile agents moving to that location leading to $\tilde{-}$. Hence, we have that $CR_i(\tilde{-} \cup -_2) = xn$. Therefore, $R_i(\tilde{-} \cup -_2) > CR_i(\tilde{-} \cup -_2)$ iff $(1-x)(n+im) > xn$, or iff $x < \frac{n+im}{2n+im}$. ■

It can be seen that the condition in Proposition 2 will tend to be satisfied if the fraction of mobile agents is sufficiently high. Otherwise, the risk dominant outcome will prevail everywhere. On the other hand, if mobility becomes sufficiently restricted, it might become impossible to support the G outcome in any location. The intuition behind this result is simple. The emergence of a G formation relies on the externality between agents that exhibit good behavior being sufficiently strong. Thus, it requires a sufficient mass of mobile agents. Barriers to mobility might prevent the critical mass from being formed, leading to the B outcome everywhere being the only outcome in the stochastically stable set. The following Proposition asserts that, for certain parameters, dividing a continent into too many small countries, i.e., restricting mobility to be very limited, will lead to $-_1$ being the only element of the stochastically stable set. This is done by looking at the modified coradius of $-_1$ in relation to the number of locations, i, that each mobile agent has access to.

Proposition 3 *Suppose that* $i < \frac{2x-1}{1-x}\frac{n}{m}$. *Then* $R_i(-_1) > CR_i^*(-_1)$.

Proof. We have that $R_i(-_1) = xn$, since the state where only immobile agents occupy a given location is reached with positive probability and we need only one location to switch to G in order to get out of $-_1$. We also have that $CR_i^*(-_1) = (n+im)(1-x)$. This is based on the following reasoning. From the state where everyone plays G, the state where every mobile agent is piled in one location is reached with positive probability. Then we need $(1-x)n(i-1)$ trembles to switch all locations except the one where the big pile is located to B. Finally, we need $(1-x)(n+im)$ trembles in order to switch the final location to B. The sum of the coradii of the intermediate steps is given by $(1-x)n(i-1)$ and $CR_i^*(-_1) = (1-x)n(i-1)+(1-x)(n+im)-(1-x)n(i-1) = (n+im)(1-x)$. Therefore, $R_i(-_1) > CR_i^*(-_1)$ iff $xn > (n+im)(1-x)$ iff $i < \frac{2x-1}{1-x}\frac{n}{m}$. ■

The above two Propositions taken together imply some intuitive comparative statics. When the number of locations in a country, i, is sufficiently small, then the risk dominant outcome is likely to prevail everywhere (Proposition 3). On the other hand, for sufficiently large i, mobile agents will pile themselves in locations in which the Pareto dominant action is played exclusively. In addition, if the number of immobile agents, n, is low, then the mobile agents will pile themselves in locations in which the Pareto dominant action is played exclusively (Proposition 2) while if the number of immobile players is large, the risk dominant outcome prevails (Proposition 2). The reverse implications are true if we vary the number of mobile players.

Note that the conditions in the statements of Propositions 2 and 3 are tight in the sense that for any combination of the parameters i, x, m, and n we can characterize the stochastically stable set. In other words if $x \in \left(\frac{1}{2}, \frac{n+im}{2n+im}\right)$ the stochastically stable set is a subset of $\tilde{-}$ and if $x \in \left(\frac{n+im}{2n+im}, 1\right)$ the stochastically stable set is a subset of $-_1$.

The next Proposition provides a rationale for why the optimal country size might be smaller than the size of the entire continent. Intuitively, the probability that the pile of mobile players forms at a given location in a country is inversely proportional to the country's size, provided the pile forms at all. Our earlier results show that at the location where the pile forms all players play G and the immobile players at that location benefit from an externality generated by the mobile players. In a smaller country, a given immobile player is more likely to be visited by the pile and therefore to benefit from this externality. Thus, an optimal level of "unification" emerges due to two opposing effects. The first effect requires that there is enough mobility for the externality to be sufficiently strong in order for the G pile to form in some location. The second effect requires that mobility is sufficiently restricted so that each individual location has the G pile being form in that location relatively frequently. The proof is clear and will be omitted.

Proposition 4 *Provided, each country is sufficiently large to render* $\tilde{-}$ *stochastically stable, we have: (a) the fraction of time that the G pile of each country spends in any given location increases in the number of countries* k *(decreases in the number of locations,* i*). (b) the number of immobile agents in the G state is* $n \times \rho$*.*

III. Some Extensions

In this section we present some examples of extensions of the basic model presented in the previous section. We also discuss the robustness of the basic model to different assumptions regarding the meeting technology. One example we study in that direction is the case where the matching technology allows for meetings between populations that are overlapping.

A. Income Distribution

Consider an example where there are two countries, each containing two symmetric locations.[6] Suppose there are three types of agents. Agents of the first type are completely immobile. Agents of the second type are mobile across locations within a country but immobile across countries. Finally, agents of the third type are completely mobile across all locations in all countries. As we mentioned earlier, in the context of our stylized model mobility across locations could also be interpreted as mobility across different skills for workers. Like before, we will assume without loss of generality that agents of all mobility levels are initially assigned symmetrically across locations (see Figure 2, top). In this environment fully mobile agents enjoy a higher frequency of the high payoff since they can relocate to any location in which a G pile is formed. Partially mobile agents enjoy a lower frequency of the high payoff since they can relocate to any location *in their own country* in which a G pile is formed. Finally, immobile agents enjoy the lowest frequency of the high payoff since they can only take advantage of a G pile if it is formed in their own location. Thus,

[6] The intuition of this example easily generalizes to many countries and many locations.

a distribution of payoffs across agents is generated with higher expected payoffs corresponding to higher levels of mobility. In addition, there are two possibilities regarding outcomes in the two countries. If one country alone does not have the critical mass of mobile agents in order to support the efficient outcome in one of its locations, the risk dominant outcome will prevail everywhere in that country (see Figure 2, middle). This is the result of the highly mobile agents immigrating and help building a G pile in one of the locations of the other country. If, on the other hand, each single country has a critical mass of mobile agents, the G outcome will prevail in one location of each country (see Figure 2, bottom).

B. Rich and Poor Countries

Consider again a variant of the basic model of the previous section where there are two countries, "rich" and "poor", and $\frac{i}{2}$ symmetric locations in each country (see Figure 3, top). In order to explore the implications of having one country that is "richer" than the other, we assume that the coordination game played in that country has a payoff matrix that has an arbitrarily small positive constant, α, added to each payoff. Of course, if we restrict mobility across the two countries, provided that the risk dominant equilibrium is not "too risk dominant," a G pile will be observed in one location of each country in the long run (see Figure 3, middle). Now assume that the border opens. Suppose that a G pile is formed in one location of the poor country. In that case, xn trembles in one of the locations in the rich country will switch that location to the G outcome. Since the payoff of the G outcome is higher in the rich country than in the poor country, the pile of mobile players will immediately relocate to the rich country. On the other hand, to destroy such a pile if it is formed in one of the locations of the rich country a total of $(1-x)(n+im)$ trembles is needed. This example suggests that the G pile will visit the rich country infinitely more often than the poor one. In other words, absent any restrictions on immigration of the mobile agents, the mobile agents will pile in the rich country (see Figure 3, bottom). We identify this with a "brain drain" effect. An arbitrarily small initial difference in payoffs between the two countries will lead to dramatic difference in the long run odds for coordination to the Pareto superior outcome in the poor country. Absent any restrictions on mobility, the mobile agents of the poor country will relocate leading to the positive externality from mobility being present only in the rich country.

C. Overlapping Populations

Consider a continent with three locations, ordered from left to right, and two types of agents, immobile ones and partially mobile ones. One set of partially mobile agents has access to the two locations on the left, another set to the two locations on the right (see Figure 4). It can be shown that our results have analogs in this environment. Consider the following classes of states, which are distinguished by the actions played in the three locations:

$$
\begin{array}{c}
-_1 \begin{bmatrix} BBB \end{bmatrix} \\
-_2 \begin{bmatrix} BGG \\ GGB \\ GGG \end{bmatrix} \\
-_3 \begin{bmatrix} GBG \\ BBG \\ GBB \\ BGB \end{bmatrix}
\end{array}
$$

We can establish that the stochastically stable set lies in $-_3$. This set includes only polymorphic states, in which different conventions coexist across locations that are partially overlapping. The argument is an analog to the proof of Proposition 2. We need xn trembles to move from BBG to BGG, and note that there is no less costly transition out of the basin of attraction of $-_1 \cup -_3$. Similarly, we need $(1-x)n$ trembles in order to move from any state in $-_2$ to $-_1 \cup -_3$. So we have that $R(-_1 \cup -_3) = xn$ and $CR(-_1 \cup -_3) = (1-x)n$. Clearly, since $x > \frac{1}{2}$, $R(-_1 \cup -_3) > CR(-_1 \cup -_3)$. On the other hand, we need $(1-x)(n+2m)$ trembles to move from BBG to

BBB,and note that there is no less costly transition out of the basin of attraction of $\Gamma_2 \cup \Gamma_3$. Similarly, we need xn trembles in order to move from any state in Γ_1 to GBB. So we have that $R(\Gamma_2 \cup \Gamma_3) = (1-x)(n+2m)$ and $CR(\Gamma_2 \cup \Gamma_3) = xn$. Therefore, for $R(\Gamma_2 \cup \Gamma_3) > CR(\Gamma_2 \cup \Gamma_3)$ we need that $(1-x)(n+2m) > xn$, or, $x < \frac{n+2m}{2n+2m}$. Notice, that this condition will not be satisfied unless the fraction of mobile players, m, is sufficiently high. Since Γ_3 lies in the above two unions, we conclude that, provided that the above condition holds, the stochastically stable set lies in Γ_3. This example demonstrates that our main analysis may be carried out under different specifications of the matching technology.

IV. Conclusions

We studied a model where boundedly rational agents are randomly matched repeatedly in order to play a coordination game with tension between risk dominance and Pareto dominance. This game can be thought of as a metaphor for many social situations in which multiple Pareto ranked equilibria are possible. We concentrated on the role of introducing different levels of mobility to equilibrium selection and, thus, to the long run distribution of payoffs across players and locations. The model had a number of predictions. Mobile agents weakly benefit from increased mobility. If mobility of the mobile agents is restricted, the risk-dominant equilibrium obtains at every location. In contrast, if country size is sufficiently large, there are enough mobile players to ensure efficient play at one of the locations and all mobile agents will be at that location. Immobile agents benefit from increased mobility at low levels of mobility, and country size has to be sufficiently large to ensure efficient play at one location. The immobile agents at that location benefit from the externality generated by the presence of the mobile players. Immobile agents lose from increased mobility at high levels of mobility. Income inequality is weakly increasing in mobility. Thus, there is an optimal country size. Finally, if there are arbitrarily small payoff differences between two countries that favor one country, opening borders causes a "brain drain" effect; in the long run, all mobile agents reside in the favored (former) country and efficiency is attained only in that country.

Our predictions are restricted to the long run outcomes of an evolutionary process. Such predictions have been criticized on the basis of the slow rate of convergence. In other words, given the small probability of the *iid* random trembles, the status quo state (whatever that might be) is likely to be observed for a large number of periods before a transition is observed. There are several lines of defence against this criticism. First, Ellison (1993) has shown that "local interactions," that is local matching rules, under which agents are matched with high probability with a small group of neighbors, dramatically improve the rates of convergence without changing equilibrium selection predictions. We expect our results to be true under local matching. A second avenue, would be to assume that the tremble probability, ϵ, is not arbitrarily close to zero but, rather, $2-20\%$. In future work, we are planning to simulate the model for such tremble rates. Finally we could drop the independence assumption on the trembles. If the trembles are positively correlated across agents, the rates of convergence will improve. At the same time, provided that trembles are not "almost perfectly correlated" (for that case no equilibrium selection occurs) the selection results will not change.[7]

In our model, an individual's payoff options depend on which community, or set of accessible locations, he belongs to. This links our model to recent work on inequality that has emphasized the role of group effects for socioeconomic outcomes and has been referred as the *memberships theory of inequality* by Durlauf (1997).[8] As noted by Durlauf, strong community identification and membership can foster both pernicious and benign social norms. Our highly stylized model delivers predictions consistent with this assessment, and adds a role for community size. In our model, high-risk, high-payoff actions are less prevalent in small communities. Larger communities make such actions more likely. However, community ties may play a beneficial role when they lead to the retention of agents who are mobile within a community.

[7] See Young (1998a,b) for a discussion of this issue.

[8] One of the central tenets of this theory is: "Individual preferences, beliefs, and opportunities are strongly influenced by one's memberships in various groups. Such groups may be fixed, such as race, or may be determined by the economy or society, such as neighborhoods, schools, or firms," see Durlauf, 1997.

V. References

Akerlof, G.A. (1997), "Social Distance and Social Decision," *Econometrica,* **65**, 1005-1027.

Anwar A.W. (1996), "When Does Immigration Facilitate Efficiency," UCL working paper.

Bhaskar, V. and F. Vega-Redondo (1996), "Migration and the Evolution of Conventions," University of Alicante Working Paper.

Dieckmann, T. (1999), "The Evolution of Conventions with Mobile Players," *Journal of Economic Behavior and Organization,* **38** 93-111.

Durlauf, S.N. (1997), "The Membership Theory of Inequality: Ideas and Implications," University of Wisconsin at Madison Working Paper.

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

Ellison, G. (1995), "Basins of Attraction, Long Run Equilibria, and the Speed of Step–by–Step Evolution," MIT working paper.

Ely, J. (1996), "Local Conventions," Manuscript.

Freidlin, M. I., and A. D. Wentzell (1984), *Random Perturbations of Dynamical Systems,* New York: Springer Verlag.

Fudenberg, D. and D.K. Levine (1998), *The Theory of Learning in Games,* Cambridge, MA: MIT Press.

Greenwood, M.J. (1975), "Research on Internal Migration in the United States: A Survey," *Journal of Economic Literature* **13** 397-433.

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

Oechsler, J. (1997), "Decentralization and the Coordination Problem," *Journal of Economic Behavior and Organization,* **32** 119-139.

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

Young, P. H. (1998), "Conventional Contracts," *Review of Economic Studies* **65,** 773-792.

Young, P. H. (1998), *Individual Strategy and Social Structure*, Princeton University Press; Princeton, NJ.