# The Emergence of Local Norms in Networks

MARY A. BURKE,[1] GARY M. FOURNIER,[2] AND KISLAYA PRASAD[3]
[1] Federal Reserve Bank of Boston, Boston, Massachusetts 02205; [2] Florida State University, Tallahassee, Florida 32306; and [3] University of Maryland at College Park, College Park, Maryland 20742

*Received October 15, 2004; revised January 26, 2006; accepted March 10, 2006*

*We develop an explanation of the emergence of local norms and the associated phenomenon of geographical variation in behavior. Individuals are assumed to interact locally with neighbors in an environment with a network externality. Although many patterns of behavior are possible, the dispersed interactive choices of agents are shown to select behavior that is locally uniform but globally diverse. The range of applications of the theory includes regional variation in the practice of medicine, technology choice, and corruption. The framework is also useful for further developing our understanding of important phenomena like lock-in, critical thresholds, and contagion.* © *2006 Wiley Periodicals, Inc. Complexity 11: 65–83, 2006*

**Key Words:** local norms; network externality; critical thresholds; geographical variations

## 1. INTRODUCTION

Geographical variations in behavior are a persistent puzzle. People in seemingly similar situations often choose to do different things. Their choices depend on where they live and the people they associate with. These circumstances give rise to uniformity of behavior within groups, together with global diversity across groups. Such differences are attributed, almost axiomatically, to differences in "culture" or "norms," as in matters of dress, speech, or driving, for example. For our purposes the term "norm" refers to an established behavior that is widespread, if not universal, within a particular community. It is also self-reinforcing: once the norm is in place it is in each individual's interest to conform to it, even though ex ante an alternative norm might have emerged.[1] In a given region, or within a given social group, there is the appearance of a socially agreed on way of responding to situations. Across groups, however, differences in behavior persist even when the groups are not isolated. The challenge is to understand how such locally uniform, globally diverse, patterns emerge. To paraphrase Morris [2], we want to know when and why we might expect "coexistent conventions."

Coexistent conventions, or local behavior norms, are pervasive. For example, the medical treatment that a patient receives depends, to an inordinate extent, on where he lives. Geographical variations in medical procedure choice were first documented by Glover [3], and subsequent studies (e.g., [4, 5]) have confirmed the presence of "small area

*Correspondence to: Kislaya Prasad, Robert H. Smith School of Business, University of Maryland at College Park, College Park, MD 20742 (e-mail: kprasad@rhsmith.umd.edu)*

[1] *In this definition a norm is the same thing as a convention, as defined in Young* [1], *and we will use the terms interchangeably.*


DOI 10.1002/cplx.20129

variations." Furthermore, choices appear relatively uniform at the local level [6]. In sharecropping, contracts between tenants and landlords often take a simple form where the tenant keeps a fixed fraction of the produce. Studies have shown that the specific fraction (e.g., half or two-thirds) tends to be uniform within regions, but can vary considerably across regions [7, 8]. The competition between alternative standards in technology [9, 10] is another example. Locally uniform convergence to a standard often arises, together with differences in choices by different groups. Likewise, there is considerable evidence of differences in norms of corruption of government officials across regions with similar governance structures, of courtesy and helpfulness towards strangers, and of the industriousness and entrepreneurship of workers. Numerous studies of organizations have indicated how the success of work groups depends on "corporate culture" or "social capital." Otherwise similar groups are capable of very different levels of performance based on a shared expectation of individual contributions to the group. Although uniform behavior within groups has been the subject of numerous inquiries stressing social influences on behavior, relatively little theoretical research in economics (we cite some examples below) has addressed the puzzle of geographical variation: why and when do different groups adopt different norms?[2]

In this article we examine a variety of regional variation phenomena using a model that incorporates local interaction within networks and social influence on choices. We build on much previous work, but especially the recent literature on evolutionary game theory (e.g., [13]). Rather than deducing equilibrium behavior from game-theoretic solution concepts predicated on strong rationality assumptions, these models describe the aggregate behavioral patterns that emerge when individuals adopt relatively simple, *boundedly rational* decision rules. Predictions focus on the stable, long-run behavior of the dynamics, which depend not only on the behavioral rules but also on the topology of the social interactions. Within this literature we follow most closely those articles that address the prospects for behavioral uniformity (and diversity) under various social maps and payoff structures, such as Morris [2], Goyal [14], and Sugden [15]. The most direct influence is the Young and Burke [7] model of the evolution of sharecropping norms. Unlike the latter article, however, the current framework does not rely on a specific parametrization of payoffs and admits a broader range of interpretations. In this more general setting analytical results are not forthcoming, and we rely on a computational approach. A computational approach to problems such as these is taken by Axtell et al. [16], Bowles et al. [17], Epstein and Axtell [18], and Epstein [19, 20] and Tesfatsion [21]. As in their work, our computational results are remarkably sharp and robust.

Our model has a network of agents with a defined neighborhood relation.[3] A central assumption is that choices of neighbors exert a direct social influence on an agent's decision. A particular choice yields a greater payoff, and so becomes more likely, if more neighbors have recently made the same choice. Decisions are myopic (and possibly error-prone). Agents take the current choices of neighbors as given and, in each period, try to choose an optimal response. However, with small probability they make the wrong choice. As this stochastic dynamic process evolves in time, we are able to witness the emergence of the characteristic pattern of locally uniform and globally diverse choices. As parameters of the model are varied, we also observe the presence of critical points that, when crossed, lead to a sudden qualitative change in the behavior of the system. As a result, norms within a region tend not to change gradually and, instead, respond suddenly as important thresholds are crossed. In the absence of errors the network could get locked into a number of possible norms. Errors, even when they are small, allow us to refine our predictions considerably.

The local uniformity is clearly a consequence of the local social influence (or network externality) assumption. Although network externality models can generate behavioral uniformity, they tend to do so on a global scale. We want to get diversity without resorting to the untenable assumption that groups are isolated. And so, our model admits the presence of a boundary, where people are under pressure from competing modes of behavior. Global diversity arises from the assumed heterogeneity in the environment, where heterogeneity occurs within as well as across regions. Global diversity can arise also in models with homogeneous agents, as in Morris [2], Sugden [15], and Goyal [14], among others. The extension to the heterogeneous case is nontrivial, and we are motivated by the richness of the applications that are afforded in such settings. For example, in the Young and Burke [7] model, regions differ in soil quality. One region may have more fertile soil on average, although all regions have both high- and low-quality plots. Some contracts are more ideally suited to high-quality plots, others to low quality plots. In the presence of local social influence, the contract chosen for a low-quality plot is likely to depend on the average soil quality in the region. In any region the landlord and tenant of the low-quality plot will be drawn toward the contract others choose, and in fertile regions this is likely to be the contract

[2] *Although several articles (e.g.,* [11, 12] *have pointed to the presence of multiple equilibria to explain geographic variation, the latent existence of multiple equilibria is not sufficient to explain the simultaneous selection of different equilibria at different, possibly contiguous, locations.*

[3] *The nature of social interactions in this model is admittedly simple, but nonetheless enables sharp and robust results. For example we assume a fixed exogenous neighborhood structure. We discuss possible extensions for future research, such as endogenous network formation, in the conclusion.*

appropriate for high-quality plots. A similar idea appears in the Burke et al. model of medical procedure choice. The ideal procedure for a patient depends on individual characteristics, such as age, which vary within the population. Physicians are also influenced by the choices of neighboring physicians. Locally uniform treatment will result, where the local norm depends on the typical patient characteristics in the region. For instance, in regions where older people are relatively more numerous the norm that emerges is for the use of procedures better suited for older patients, *even on younger patients*. In each case, an identical transaction (between identical landlords and tenants on identical plots, or identical physicians and patients) will result in very different outcomes at different locations. The predicted relationship between average regional characteristics and local norms holds up rather well in data sets on sharecropping in Illinois [7] and cardiac care in Florida [6].

The environment we present here extends Burke et al.'s. [6] finding of robust geographical variation and generalizes it to other applications. In [6], stable long-run variation requires an infinite geography in which physicians are assumed to be located on the integers $Z$, and their neighbors are the adjacent physicians. Here we obtain regional variation for a finite number of agents (a harder problem, surprisingly) and for a larger set of spatial arrangements. Moreover, errors in decisionmaking (not considered in that article) are shown to be crucial for producing regional variation in the finite case. When regional variation is a stable outcome of the noiseless process the noisy dynamic process approximates exactly this equilibrium (leading us to believe that regional variation is stable, in the sense of Kandori et al. [22] and Young [1]). Remarkably, even when regional variation is not a stable long-run outcome, we find that it arises as a robust phenomenon for considerable lengths of time.

The rest of the article is organized as follows. We describe the general model in section 2 and present the results from our simulations in section 3. Finally, we draw together our conclusions in section 4, emphasizing what we learn about the emergence of norms and discussing the key phenomena observed—regional variation, criticality, and lock-in.

## 2. MODEL

There are $K$ agents or decision makers. Each agent $x$ has a set of neighbors, $\mathcal{N}(x)$. We consider two arrangements of the agents (called geographies): (i) a circle and (ii) a torus. The circle has the virtue of simplicity, while still capturing the notion of local, overlapping interactions. Although any individual is influenced only by the actions of two neighboring agents, the behavior of others much farther away may exert an indirect influence. The torus model gives each agent 4 rather than 2 neighbors and affects the behavior at the boundaries between regions, as described below. We find that regional variation may arise and persist in both types of graphs, but the differences between the cases indicate that network size is important. In the circles, we index the $K$ agents by the numbers $1, 2, \ldots, K$ and define the neighborhoods by

$$\mathcal{N}(i) = \begin{cases} \{2, K\} & \text{if } i = 1 \\ \{1, K-1\} & \text{if } i = K \\ \{i-1, i+1\} & \text{if } K > i > 1 \end{cases}$$

In effect, agents are located along a single dimension, with the neighbors being those at adjacent locations. However, we define 1 and $K$ to be neighbors. In the torus geography, we do the same with a two-dimensional arrangement. Suppose that each location is identified by two coordinates $(i, j)$. We assume that $1 \le i \le M$ and $1 \le j \le M$, together with $K = M \times M$. Now the neighbors of $(i, j)$ are $\{(i+1, j), (i-1, j), (i, j+1), (i, j-1)\}$, with the obvious modification at the edges.

At each date, every agent observes a private signal $\sigma_i \in \{\alpha, \beta\}$ and chooses a decision $d_i \in \{A, B\}$. Payoffs depend upon $(d_i, \sigma_i)$, as well as on the decisions of neighbors.[4] Let $n$ be the number of neighbors of agent $i$ who chose action $A$ *in the previous period:* we denote payoffs by $\pi(d_i, \sigma_i, n)$. Define the payoff difference $\Delta^\sigma(n) = \pi(A, \sigma, n) - \pi(B, \sigma, n)$. Our key *social influence* or *network externality* assumption is as follows:

**Assumption 1.** *$\Delta^\sigma(n)$ is an increasing function of $n$ for all $\sigma$.*

In other words, $A$ becomes relatively more attractive if more neighbors decided to play $A$ in the previous period. The previous definitions lead to two possibilities: (i) either $\Delta^\sigma(n) \le 0$ for all $n$ or $\Delta^\sigma(n) \ge 0$ for all $n$, or (ii) $\Delta^\sigma(n)$ changes sign for some value of $n$. In the former case the choices of neighbors do not affect the ranking of actions, whereas in the latter case they do. We focus on case (ii)—for both signals the relative ranking of the actions changes with $n$.[5] Of particular interest is how rankings change with the signal. Let $N$ be the total number of neighbors (which is an even number for both the circle and the torus). When $n = N/2$, the neighbors are equally split between playing the two actions, so the effect of the two is, in some sense, neutralized. We assume

**Assumption 2.** *At $n = N/2$, the sign of $\Delta^\sigma(n)$ changes with the signal; specifically, $\Delta^\alpha(N/2) > 0$ and $\Delta^\beta(N/2) < 0$.*

This assumption captures the idea that $A$ is the better choice when signal $\alpha$ is observed, whereas $B$ is the better choice if signal $\beta$ is observed (once social influences are neutralized). Together with the assumption 1, and the focus on case (ii), this implies that an agent should choose $A$ if all $N$ neighbors

---

[4] *Mathematically, the structure here is an interacting particle system. These are discussed by Liggett* [23] *and Schinazi* [24]. *Such systems were introduced fruitfully into game theory by Blume* [25].

[5] *For case (i), our arguments imply that the emergent behavior would be to play the action superior for all $n$.*

do so and $B$ if all neighbors choose $B$. In other words, "lock-in" to either action is a possibility.

We assume that decision making is prone to error, according to

**Assumption 3.** *Given $(\sigma, n)$, with probability $(1 - \varepsilon)$ an agent chooses the action that maximizes payoffs, and with probability $\varepsilon$ chooses the inferior action.*

Our final assumption relates to properties of signals. A region is a fixed, contiguous set of locations. Signals arrive according to a fixed probability distribution within a region, but the probability can differ across regions. To simplify, we allow only two possible values for the probability of receiving signal $\alpha$, $p$, and $q$. For example, suppose there are two regions, called East and West. In the circle, West is defined to be the set $\{i | i \leq K/2\}$; in the torus it is $\{(i, j) | j \leq M/2\}$. A possible distribution of signal probabilities (the *East-West* distribution) is the following: The probability that a location $x \in$ East receives signal $\alpha$ is $p$, and the probability location $y \in$ West receives signal $\alpha$ is $q$. There can be more than two regions, but in that case (given our assumption of only two values of the signal probability) some will share a common value of the signal probability. Formally, we have

**Assumption 4.** *There are two types of regions, distinguished by the probability of receiving signal $\alpha$. Some regions have probability $p$, others have probability $q$. In general, $p > \frac{1}{2}$ and $q < \frac{1}{2}$.*

An important feature of these regions is that they are not isolated—there is a boundary at which agents from one region come into contact with agents from the other region. There are other ways to model contact between regions. One alternative that is similar to our circle model consists of *two* circles, with the property that occasionally an agent in one circle imitates a randomly selected individual from the other circle.[6]

In the next section, regions other than East-West are also considered. We find it convenient to define regions with reference to some pattern of choices (choice distributions) or some pattern of signal probabilities (signal probability distributions). A choice distribution and a signal probability distribution will be said to *conform* if the distribution of signals assigns probability $p > \frac{1}{2}$ at locations where the choice is $A$, and probability $q < \frac{1}{2}$ at locations where the choice is $B$. Once an initial set of choices is specified, we have a fully specified stochastic dynamic system. The state consists of an assignment of choices (from $\{A, B\}$) to each location. In each iteration, a signal is generated for each location and choices are updated. This leads to a new state, and the process can be repeated.

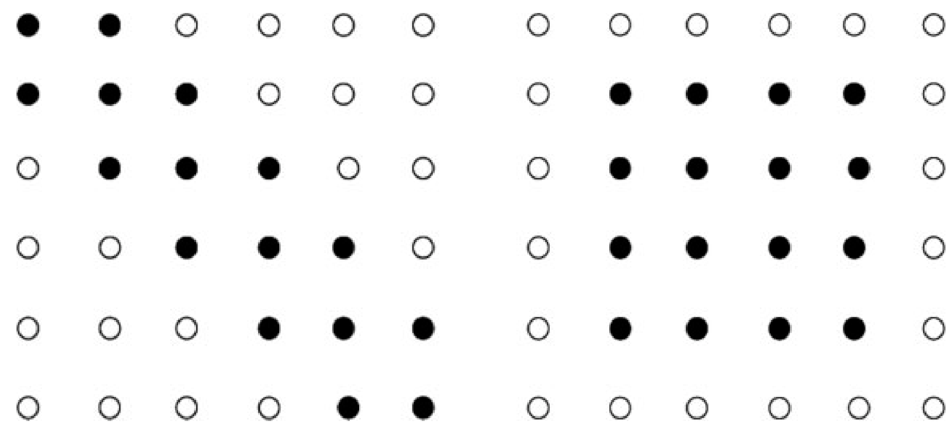
[6]*We thank Rinaldo Schinazi for this observation.*

We now illustrate the reach of this model with three applications, beginning with a finite version of the *medical procedure choice* problem discussed in Burke et al. [6]. The decisionmakers are physicians who, in each period, get a new patient with a specific condition (say, coronary atherosclerosis). Signals are now to be interpreted as patient characteristics, e.g., age ($\alpha$ is "old" and $\beta$ is "young"). After observing patient characteristics, a physician must choose between two procedures [drugs (A) or surgery (B)]. Physicians are influenced by the choices made recently by their neighbors—either because they talk to, and learn from, neighbors or from a desire to conform with local practice. In the manner of assumption 2, procedure $A$ is better for $\alpha$ (old) patients, whereas $B$ is better for $\beta$ (young) patients. Patient characteristics (age distributions) can differ across regions ($p \neq q$). We want to know whether stable patterns of procedure use evolve and whether patients in different regions are likely to receive different treatments.

Our second example is of technology choice with a network externality. The decisionmakers are problem solvers who belong to one of two professional groups. There are two available technologies ($A$ and $B$) from which an individual must choose. An agent's neighbors are people he interacts with, typically from the same profession. We index people in such a manner that a region comprises all the people in the same profession (the people with ties across professions are placed on the boundary). For concreteness, the East comprises of graphic designers and the West writers. At each date, each individual gets a task which may be intensive in the use of images ($\alpha$) or of text ($\beta$). Image-intensive tasks are best solved using technology $A$, whereas technology $B$ is best for text-intensive tasks. There is also a network externality present. You are more likely to use a technology if your neighbors use the same technology. People in both professions get both types of tasks, but graphic designers are more likely to get image-intensive tasks ($p > q$). The questions now concern whether stable patterns of technology adoption arise and whether technology use differs across professional groups.

Our final example concerns corruption of government employees, and we assume the geography of a circle. The agents are now officials, who choose to either solicit a bribe ($A$) or not ($B$). The payoffs imply that officials are more likely to solicit bribes if their neighbors do so (perhaps because this lowers the risk of getting caught or reduces the stigma associated with corruption). A signal is now associated with the arrival of a private individual with some business transaction. This person has some observable characteristic that can take one of two values, $\alpha$ or $\beta$, such as rich or poor, doctor or lawyer, member of one ethnic group or another. A given individual's corruptibility, that is, his willingness to pay the bribe rather than report the official, cannot be observed in advance, but it is correlated with the observable characteristic. For example, suppose that 70% of $\alpha$'s are corruptible (pay the bribes),

**FIGURE 1**



"Diagonal" and "Square" choice distribution—unstable in the long-run.

but only 30% of $\beta$ types are. The payoffs to the bribe solicitor, conditional on the corruptibility of the citizen, are denoted by $\hat{\pi}(d_i, b_k, n)$, where $b_k \in \{c, \neg c\}$ indicates whether the citizen is corruptible or not. Let $\hat{\pi}(A, c, n) = n, \hat{\pi}(A, \neg c, n) = n - 2$, and let $\hat{\pi}(B, b_k, n) = 0$ for either value of $b_k$. Given these payoffs and the conditional probabilities of corruptibility, the expected payoff for an official that solicits a bribe from an $\alpha$ type is thus

$$\pi(A, \alpha, n) = 0.7n + 0.3(n - 2) = n - 0.6.$$

Similarly, the expected payoff from soliciting a bribe from a $\beta$ type is $\pi(A, \beta, n) = n - 1.4$. Assuming $N = 2$, it can readily be confirmed that Assumptions 1 and 2 hold. We assume regions differ in their proportions of observable types, and so regions also differ in their rates of corruptibility. Again we want to know whether the emergent patterns of corruption among officials display regional variation, as well as which circumstances lead to a noncorrupt governance norm.

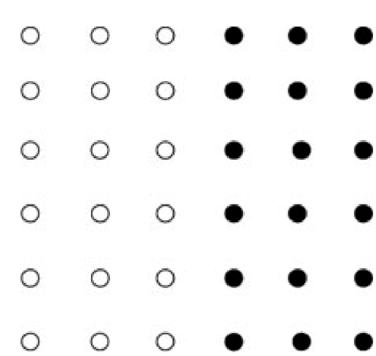
## 3. RESULTS

We focus our results discussion on the case of the torus model. Because the stochastic dynamical systems described above are Markov chains, long-run behavior in the "zero noise" (no decision error) case is fairly easy to establish analytically. We record these results below for reference and thereafter describe the results of our computational simulations of the zero noise model and of its corresponding variants involving noise. The simulations not only lend insight into behavior when analytical results are unavailable, but they also serve as a selection device when the theory indicates multiple long-run outcomes. Further, the simulations reveal interesting phenomena that would be missed in an analysis of only long-run behavior.

### 3.1. Long-run Outcomes in the "Zero Noise" Case

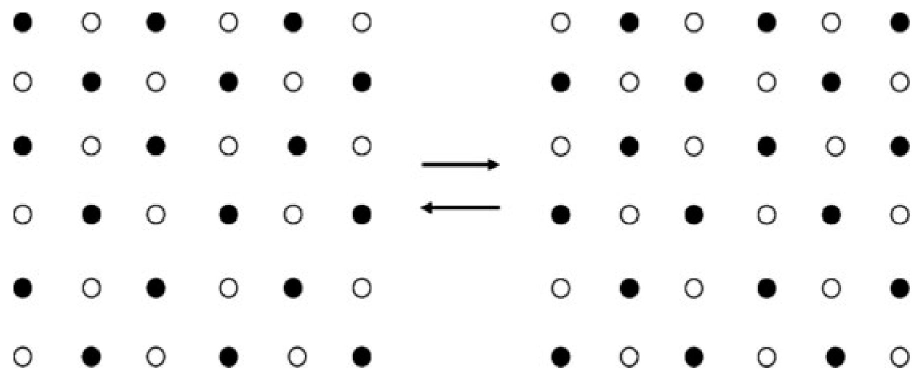
We use a somewhat informal (simulation-based) definition of *long-run outcome*. Let $T$ denote the length of the simulation of a dynamical system. A state is a possible long-run outcome if, starting from *some* set of initial conditions, it arises for a strictly positive fraction of time as $T \to \infty$. Such a state is also said to be *stable* in the *long-run*. In our model, the perfectly uniform states **A** and **B** are clearly long-run outcomes. Once the system is in one of these states, it never leaves. In the next proposition we examine other distributions of choices—some that are long-run outcomes and others that are not. These configurations will play an important role in our simulations and are depicted in Figures 1–4 for $M = 6$. Each circle specifies a location, with the neighborhood relation as defined in section 2. A black dot denotes choice of $A$, whereas a white dot denotes a choice of $B$.

**FIGURE 2**



Choices conforming with the "East-West" signal distribution—stable in the long-run.

FIGURE 3



The "blinking" cycle.

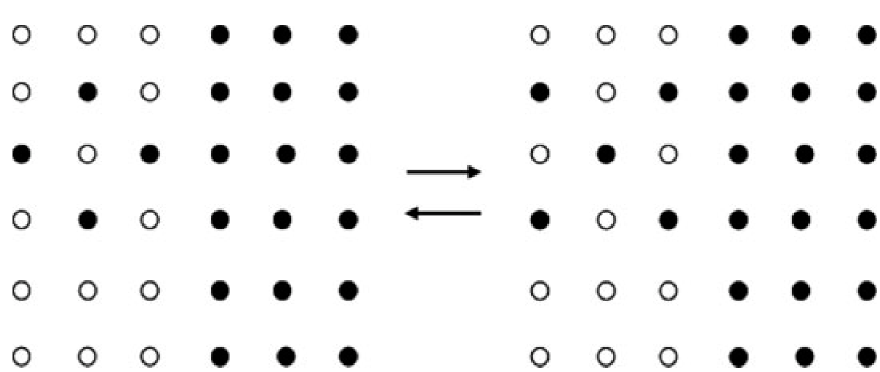
**Proposition 1.** *Suppose Assumptions 1–4 are satisfied for the torus model and there is no error in decisionmaking* ($\varepsilon = 0$) *Then,*

1. **A**, **B**, *"East-West," the "blinking" cycle, and mixed cycles as in Figure 4 are possible long-run outcomes.*
2. *"Diagonal" and "Square" cannot be long-run outcomes.*

The proof is immediate. For **A**, **B** and choices *conforming* (refer to the prior definition following Assumption 4) to the "East-West" distribution of signals we note that each agent has at least three neighbors matching his own choice. So, regardless of the signal received, each individual will continue with this choice. In fact, any horizontal (vertical) band of length (height) six and width at least two will be stable in the long-run. For the "blinking" and mixed cycles, each arrow in the figure denotes a transition from one state to the other. Inspection of each location in Figure 3 shows that in either state, all four neighbors of an agent are people making the opposite choice. Thus each individual alternates between his two choices, and each transition between the depicted states will occur with probability one. A similar argument shows that each transition in Figure 4 also occurs with probability one.

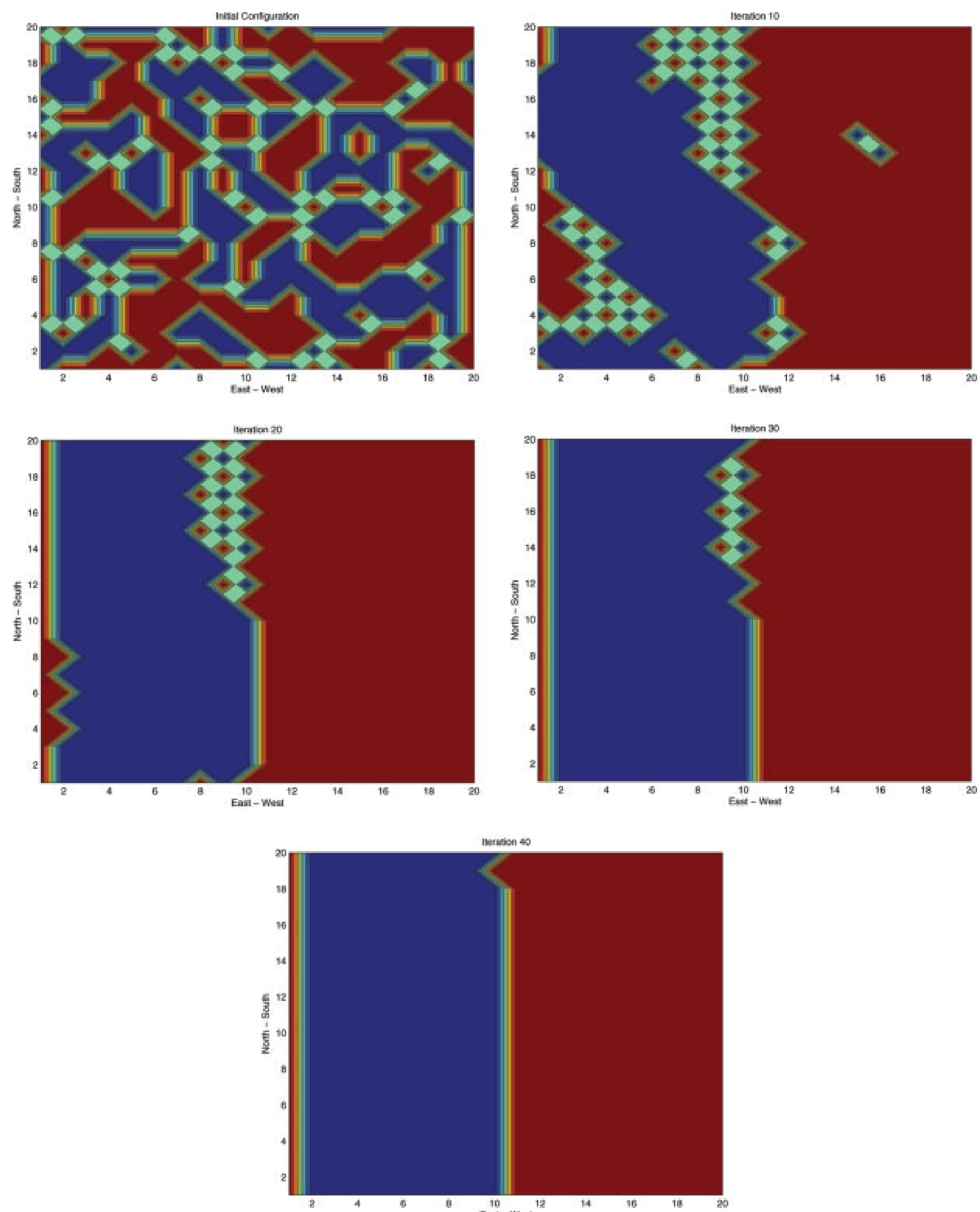
Now consider the "diagonal" distribution of choices. We show that this must be a transient state. Each location at the boundary of the diagonal band has two neighbors whose last choice was *A*, and two neighbors whose last choice was *B*. So the choice at these locations depends on the value of the signal. There is positive probability that signals will be such that agents on the boundary change their choice from *A* to *B*. The boundary now recedes, and again there is positive probability of its receding yet further. Because the diagonal band is finite, the probability of its disappearing (leaving the system

FIGURE 4



A stable mixed cycle.

**FIGURE 5**



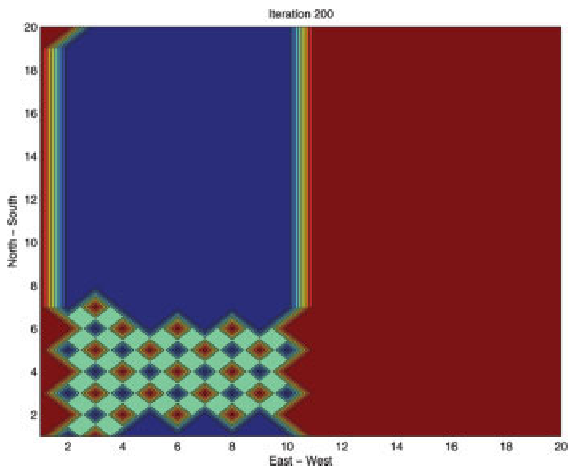

Convergence in the zero noise case from a random initial configuration.

in state **B**) is strictly positive. Once the system is in state **B**, it never leaves **B**. The time to reach **B** (unless it reaches **A**, or some other absorbing set of states first) may be large, but is finite. As the length of the simulation $T \rightarrow \infty$ the fraction of time in the diagonal state must approach zero. A similar argument applies for the "square" choice distribution. This completes the proof.

### 3.2. Computational Results

The general format for each experiment is to begin with a particular distribution of signals (such as *East-West*). Then, starting from a random initial configuration of choices, we examine the evolution of the system. To illustrate, consider the following example. The signal distribution is *East-West*, the error probability is zero, and $M = 40$. In addition,

**FIGURE 6**




An alternative simulation in the zero noise case leads to a mixed state.

$p = 0.7$, $q = 0.3$. In Figure 5, we illustrate convergence, starting from a random initial assignment of choices.[7] *A* is depicted in red, and *B* in blue. In this instance, we happen to get convergence to regional variation of choices (a distribution of choices that conforms to the signal distribution). The choice is *A* at locations where the signal $\alpha$ has probability 0.7; it is *B* in locations where the probability of $\alpha$ is 0.3. People make different (signal-independent) choices in the two regions.

As one would expect from Proposition 1, there are other possibilities. Another simulation, starting once again from random initial conditions, arrived in iteration 200 to the state in Figure 6.

Broad inferences about the properties of the system require a more thorough investigation, involving repetitions and variations in parameters. Our overall design involves a very large number of simulations, and Table 1 provides a summary and guide to the experiments and reported results. An "experiment" is a fixed number of repetitions (always taken to be 50) with the same set of parameter values. Each repetition involves 1000 iterations (in every iteration a decision is made by each individual, after observing their signal). So, for instance, we conduct an experiment where the signal probability distribution is *East-West*, where $M = 20$, $\varepsilon = 0.01$, initial choices are randomly generated, and $(p, q) = (0.7, 0.3)$. The parameters we choose to vary are not arbitrary. The qualitative nature of results depend in significant ways upon the values of $M$, $\varepsilon$, and $(p, q)$, and our experiments illustrate this. In addition to *East-West*, we consider three other signal probability distributions: *Diagonal, Square*, and *Random*. In *Diagonal* (*Square*) the probability of $\alpha$ is $p$ along a central diagonal band (square region) and $q$ elsewhere. We want to determine if a conforming choice distribution (as in Figure 1) will arise. In the *Random* distribution, signal probability $p$ or $q$ is assigned to each location according to some probability distribution.

[7] *All figures were generated using Matlab. Each location is represented as a diamond, which will be either red or blue depending on the choice. The background is light blue. However, when adjacent cells are of the same color, Matlab fills in the background using that same color. The Matlab programs are available from the authors. The iteration in which we take a snapshot appears at the top of the frame.*

**TABLE 1**

Guide to Results for the Torus Geography

| | Signal Probability Distribution | | | |
|---|---|---|---|---|
| | *East-West* | *Diagonal* | *Square* | *Random* |
| Number of agents is $M \times M$ | $M = 20, 40$ | $M = 20$ | $M = 20$ | $M = 20, 40, 100$ |
| Error probability | $\varepsilon = 0, 0.01, 0.05$ | $\varepsilon = 0, 0.01$ | $\varepsilon = 0, 0.01$ | $\varepsilon = 0, 0.01, 0.05$* |
| Starting choice configuration | Random **A**, **B**, "Blinking" | Random | Random | Random |
| No. of repetitions (in each experiment) | 50 | 50 | 50 | 50 |
| No. of iterations (in each simulation) | 1000 | 1000<br>5000* | 1000<br>5000* | 1000 |
| Probability values ($p$, $q$) | (0.7, 0.3), (0.7, 0.4)<br>(0.7, 0.3), (0.7, 0.4)<br>(0.7, 0.5), (0.7, 0.7)<br>(0.5, 0.7), (0.4, 0.7)<br>(0.3, 0.7) | (0.7, 0.3) | (0.7, 0.3) | (0.7, 0.3) |

Simulations with an asterisk are discussed, but not reported in the article.

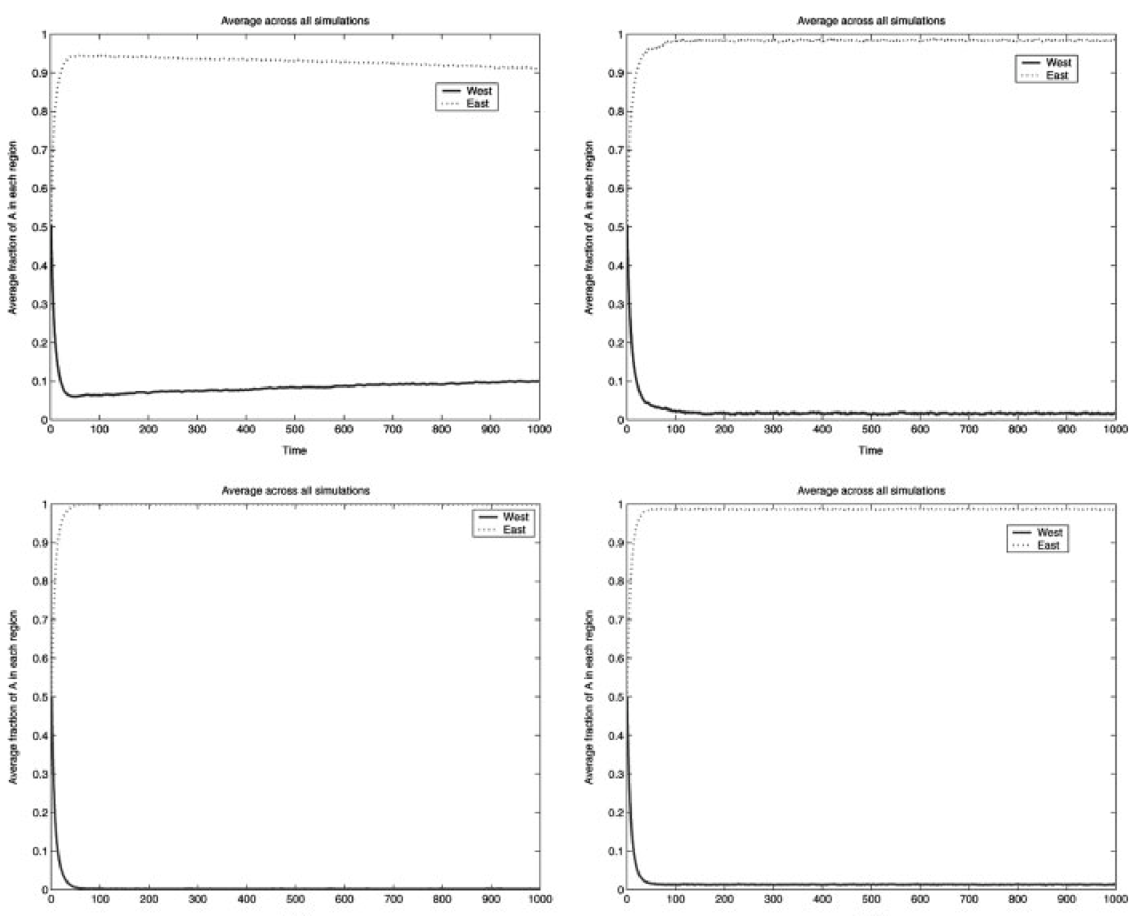
### 3.2.1. East-West

The results for the *East-West* distribution of signals are presented in Figure 7 for $M = 20$ and $M = 40$ when $(p, q) = (0.7, 0.3)$ for random initial choices. The top panel is for $M = 20$, with the results for noisy decision-making ($\varepsilon = 0.01$) on the right. Each graph shows the average fraction of $A$'s in the two regions (where the average is across 50 repetitions of the simulation). There is clear evidence that a conforming segregation of choices by region occurs. $A$ is used at a very high proportion of the 200 locations in the East, and a low proportion of locations in the West. Despite the fact that **A** and **B** are theoretically possible when $\varepsilon = 0$, they never occur. In the zero noise case mixed states, as in Figure 4 and 6, often occur; this is what explains the departure from full segregation in the top left graph. There also appears to be a tendency for the blinking regions to increase with time. The effect of the introduction of small noise is striking. Even a one in hundred chance of error (top right graph) leads to a much sharper segregation of choices. The noise tends to make blinking cycles unstable, as they are invaded by uniform regions. The lower set of graphs display results from a repetition of the experiment for $M = 40$. There is now complete (conforming) segregation, even in the zero noise case (use of $A$ in the East is close to 100%, in the West close to 0). This persists for the higher values of $M$ that we have tested (the opposite is true when $M$ is smaller than 20). The reason, pieced together from an examination of initial conditions, appears to be as follows. The stability of a blinking segment requires that it span across a region. When regions are large, it is unlikely for such a pattern (or one leading to it) to arise purely by chance when the random assignment of initial choices is made.

### 3.2.2. Diagonal and Square

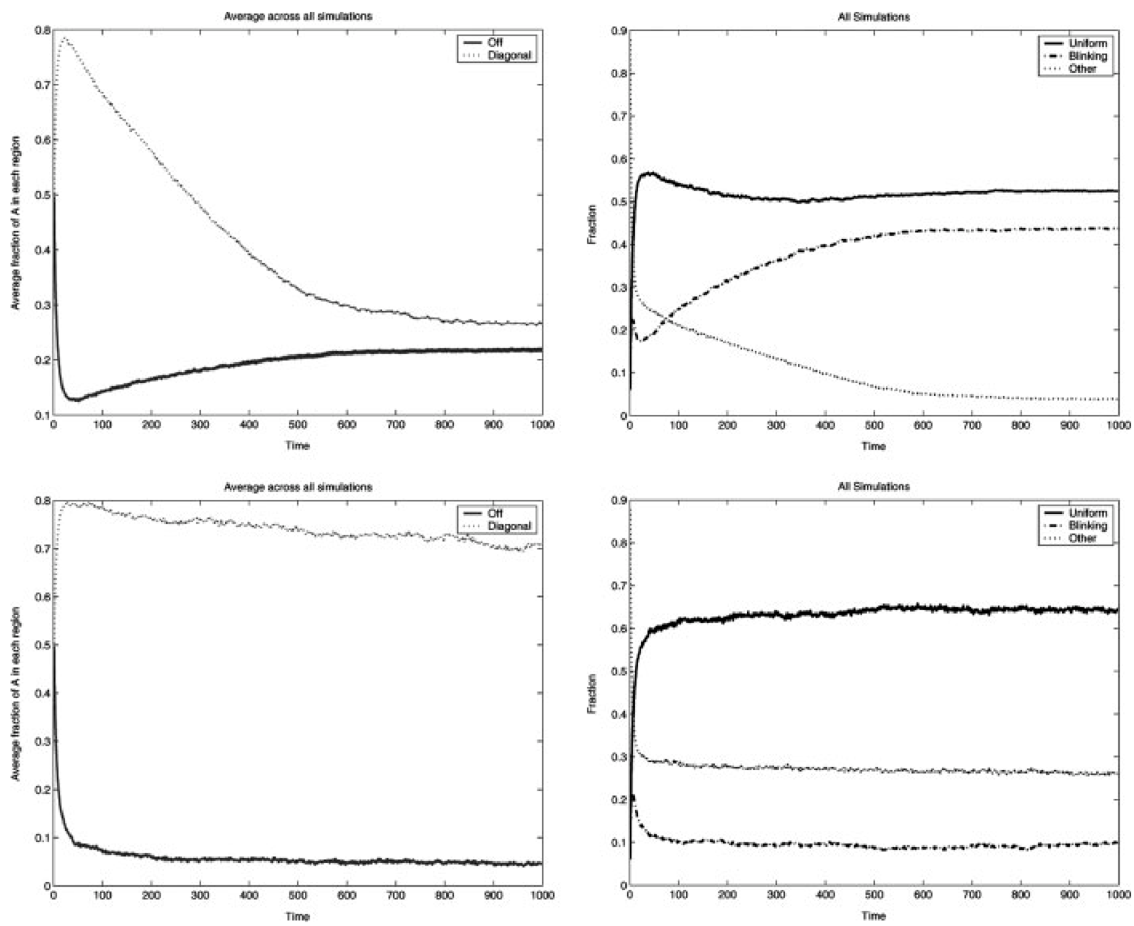
Suppose that the signal probability distribution is Diagonal (the case of Square is very similar and so will not be

**FIGURE 7**




The "East-West" configuration for $M = 20$ and 40 (with noise on the right).

**FIGURE 8**

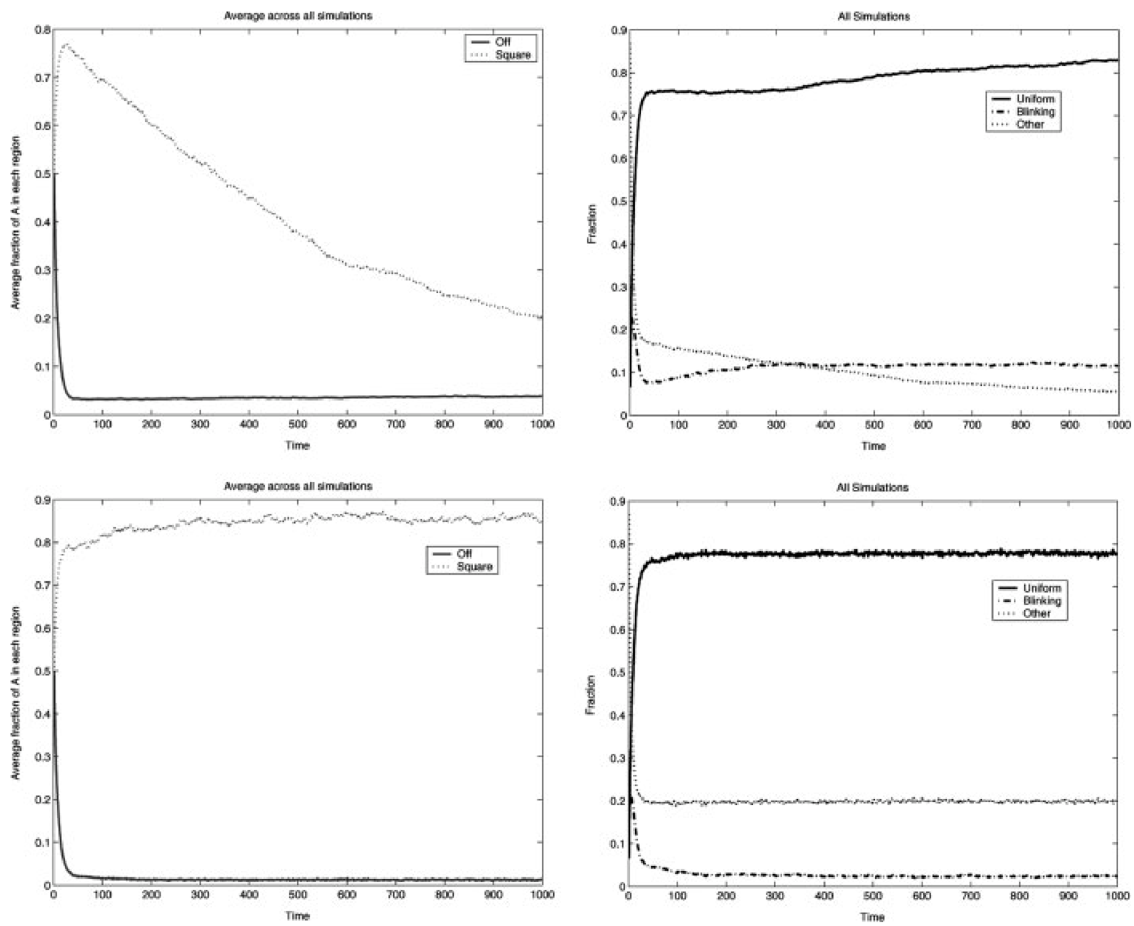



The "Diagonal" with noise (bottom panel) and without (top panel).

discussed separately). In other words, all locations along a diagonal region have signal probability $p > 0.5$, whereas all off-diagonal locations have probability $q < 0.5$. In a conforming set of choices, all agents on the diagonal would choose $A$, and all others would choose $B$. From Proposition 1, we know that such a pattern of choices cannot be long-run stable when $\varepsilon = 0$. But we run an experiment, for $M = 20$, $(p, q) = (0.7, 0.3)$ and starting with random initial choices. There are 190 locations in the diagonal, and 210 locations are off-diagonal (in the case of the square, $M = 20$ and the square is $10 \times 10$). The top panel of Figure 8 depicts the results without noise, and results for $\varepsilon = 0.01$ are in the bottom panel. The left graph depicts the average fraction of $A$'s in each region. In the graph on the right we classify behavior at each location as being uniform, blinking, or something else using the following procedure. For every location, we determine whether the individual makes the same choice as *all* of his neighbors. If so, he is classified in the uniform category.[8] If the individual makes a choice that is the opposite of *all* neighbors' choices he is classified as blinking. Everyone else is classified in a residual "other" category. For an individual simulation, such a graph gives us a fairly accurate picture of the extent to which there are regional norms. A difficulty arises when averaging these numbers across simulations. We can get the same averages either through a sequence of all uniform and all blinking configurations or through a sequence of mixed distributions. The variance of these percentages is computed and can be useful in distinguishing the two types of situations.

[8]*This measure is quite conservative, especially when noise is present. A single discordant choice in the middle of a uniform region causes four neighbors to be classified in the "other" category.*

**FIGURE 9**




The "Square" with noise (bottom panel) and without (top panel).

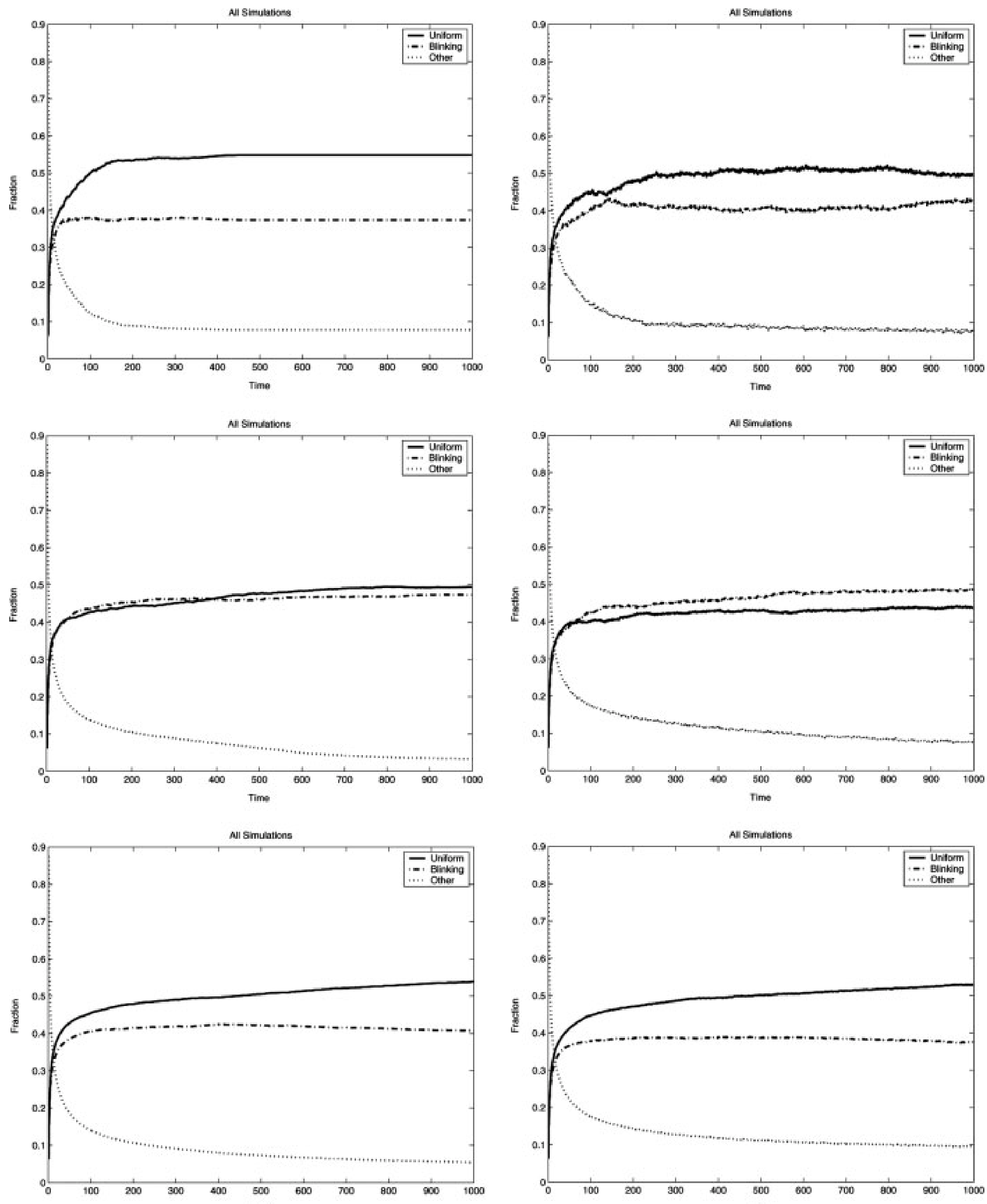
From Figure 8 we see that, in very few iterations, the system organizes into a conforming regional variation pattern. In the noiseless case the pattern then decays slowly, but it is quite a while before the two regions become indistinguishable. We observe either an invasion of the diagonal by the surrounding uniform region or by a blinking pattern. In the simulation of the diagonal the latter is more often the case. But for the square (Figure 9) we are more likely to get invasion by surrounding uniform regions. The introduction of small noise has a very marked effect. Choices conforming to the regional distribution of signal probabilities are much more persistent (this is also seen when we separate the percentage of uniform states in the right-hand graph by choice). Another difference between the diagonal and square distributions is that in the former the average fraction of $A$'s in the diagonal region shows a very slight decline, which is not true for the square. In both cases we extended the simulation to 5000 iterations—and find that our results hold.

### *3.2.3. Random Regions*

In this section we allow for heterogeneity of signal probability within a region. We continue with two possible values, $p$ and $q$, but no longer require all locations within an exogenously defined region to have the same probability. The signal probability distribution can be specified as follows: let $\theta$ be the probability that the $\alpha$-probability at a location within a region is $q(=0.3)$. In the first instance, there is a single value of $\theta$ for the whole network.

Let $\theta = 0.5$. At each location the probability of arrival of signal $\alpha$ is chosen from $\{p, q\}$, with each choice being given equal probability. We then run a simulation, starting from a random initial configuration of choices. Three sets of simulations are run, for $M = 20$, 40, 100 and $\varepsilon \in \{0, 0.01\}$. The results are presented in Figures 10 and 11. Quite surprisingly, large regions with uniform behavior patterns form, coexisting with large regions where the blinking cycle is present. Both these regions persist even when there is noise in the

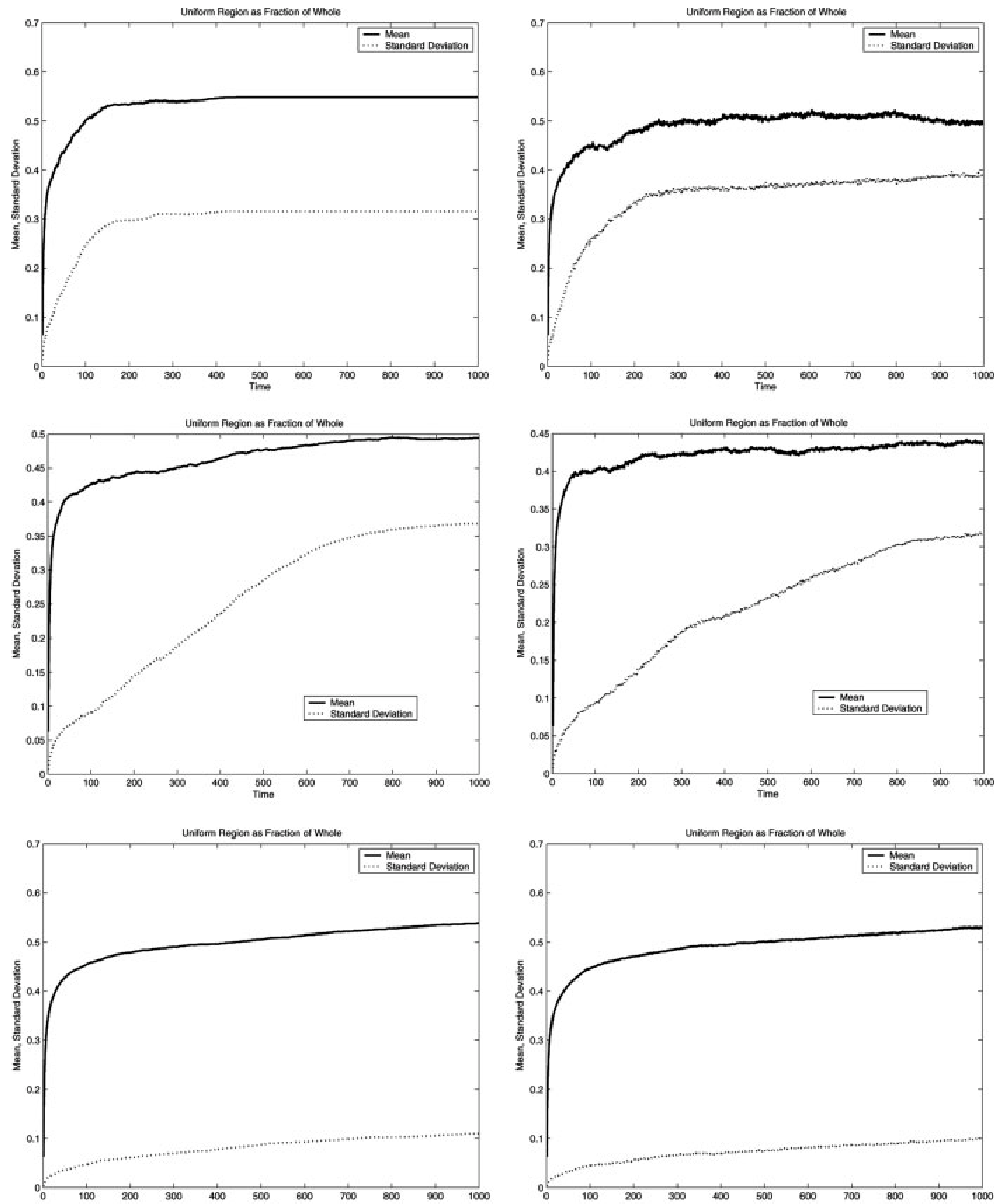
## FIGURE 10




The “random” configuration for, from top to bottom, $M = 20, 40, 100$. Simulations with noise ($\varepsilon = 0.01$) shown on the right.

environment. A typical outcome (for $M = 100$ and $\varepsilon = 0.01$) is displayed in Figure 12. In general, the average of the percentage of uniform cells gets close to 50%. Every individual simulation is more likely to end up in a mixed pattern with a similar percentage of uniform cells when $M$ is larger (simulations with small $M$ are more likely to be of the “all or nothing” variety). This is apparent by looking at Figure 11, where the standard deviation of the fraction of space that is uniform is

**FIGURE 11**

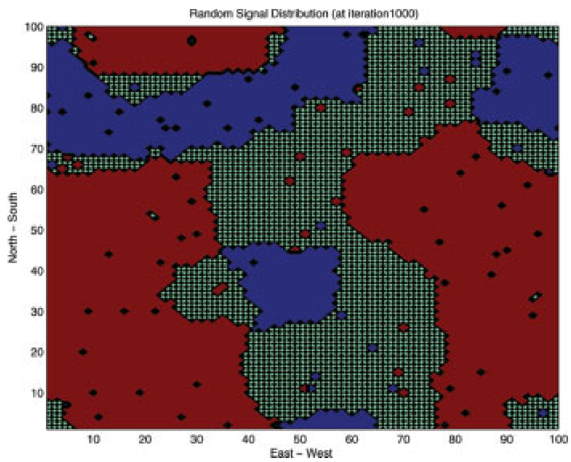



Mean and standard deviation of the fraction of space that was uniform for top to bottom; $M = 20$, 40, and 100. Simulation with noise ($\varepsilon = 0.01$) shown on the right.

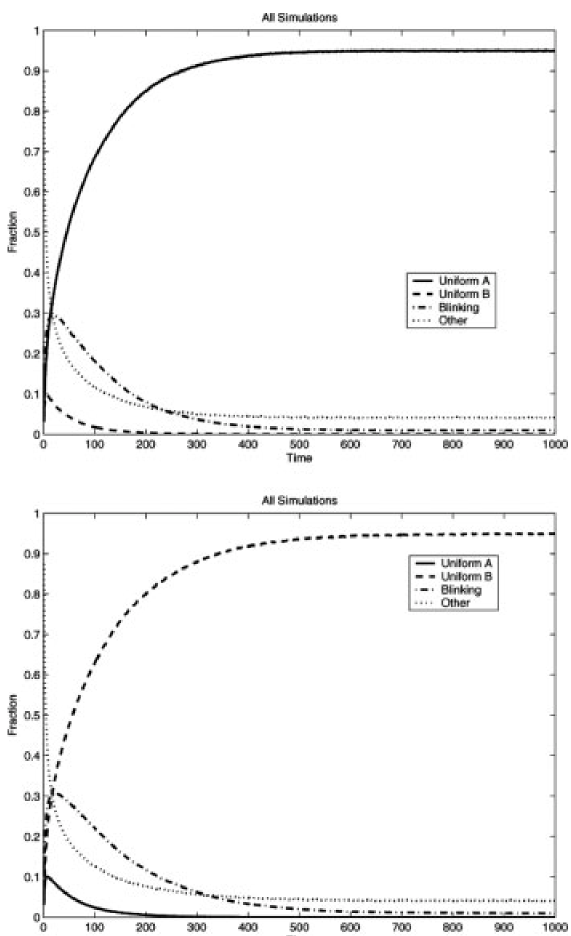
also plotted. This is seen to be a much smaller number when $M = 100$.

Since noise has been seen to be inimical to the existence of blinking regions, we repeated our experiment (for $N = 100$) with the noise level raised to $\varepsilon = 0.05$. The results are largely unchanged (and so the corresponding graphs are not presented here). We did find an increase in the "Other" category (as expected), and a slight decline in the "Blinking" category.

## FIGURES 12 AND 13

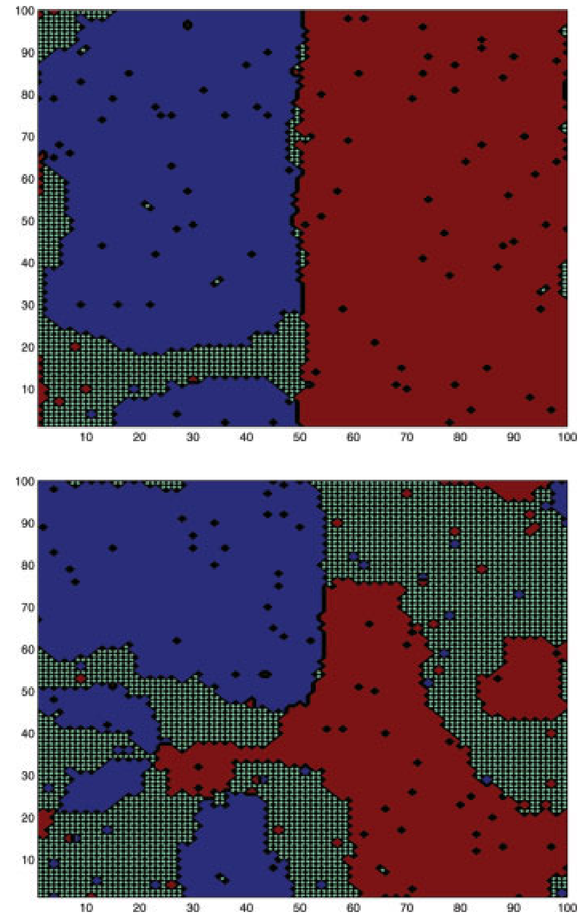



Final state in a simulation with a random signal probability distribution.




Signal probabilities are randomly chosen from {0.3, 0.7} at each location. In the top frame the event ($q = 0.3$) has probability $\theta = 0.4$ and in the bottom frame $\theta = 0.6$.

## FIGURE 14



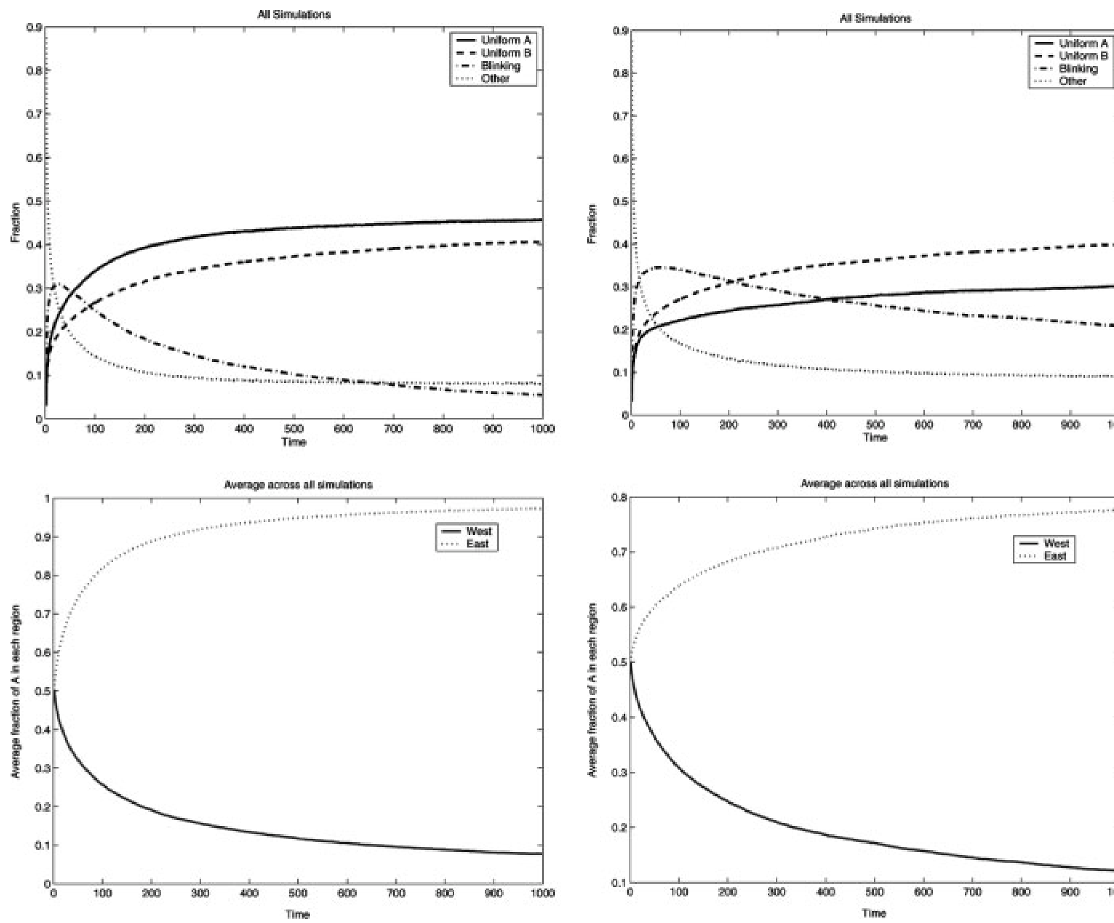
East and West differ in $\theta$. Top, $\theta = 0.6$ in the West and $\theta = 0.4$ in the East; bottom, these probabilities are 0.55 and 0.45, respectively.

However, regions with the blinking cycle do appear to be quite stable.[9]

We now change the probability with which a location gets assigned a signal probability (from $\{p,q\}$). First, let the probability, $\theta$, of choosing "$q = 0.3$" be 0.4. In other words, at roughly 40% of the locations there is a 0.3 probability that the signal will be $\alpha$; at about 60% of the locations the $\alpha$-probability is 0.7. Clearly, circumstances favor the use of $A$. This is what we find in the top panel of Figure 13: about 95% of the locations can be classified as using $A$. In the lower panel, we repeat the experiment with $\theta = 0.6$. Now choices are uniformly $B$.

[9]*There is some danger that our randomly generated signal probability distribution is exceptional. Consequently, as a robustness check, we did 50 simulations generating new signal probability distributions (results not reported). We find our results to be quite representative.*

**FIGURE 15**




East and West differ in $\theta$. Left column, $\theta = 0.6$ in the West and $\theta = 0.4$ in the East; right column, these probabilities are 0.55 and 0.45, respectively.

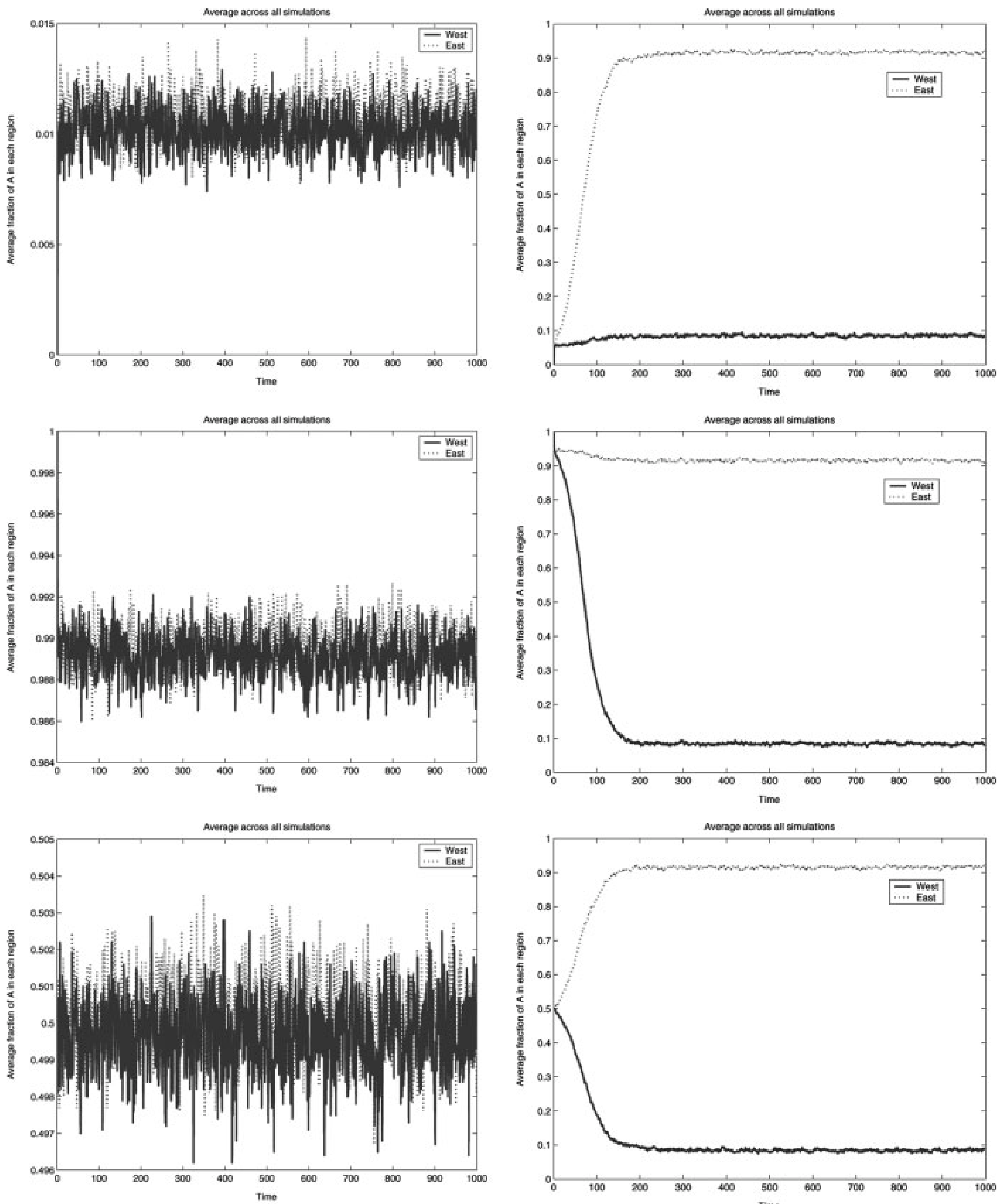
In Figures 14 and 15, we again consider East and West. Only now, in the West the probability $\theta$ is greater than 0.5, whereas in the East it is less than 0.5. We find that regional variation arises again as a robust phenomenon, with $A$ chosen in regions where $\theta < 0.5$, and $B$ chosen in regions with $\theta > 0.5$. Even a small majority of locations with $\alpha$-probability $p > 0.5 > q$ leads to the emergence of a regional norm where $A$ is used exclusively. Figure 14 shows a typical snapshot, after starting from random initial conditions. Figure 15 presents statistics. We report, in the right-hand column, the effect of taking the regional $\theta$ values closer to 0.5. Even when the value only mildly favors an action (e.g. $\theta = 0.45$, favoring $A$ in the East), this action arises as the local norm. Of course, the results are much sharper when $\theta$ is further away from 0.5. We appear to have a "critical threshold" result, with $\theta = 0.5$ being the critical value. Coexistence of different norms occurs at $\theta = 0.5$, but otherwise requires regions that differ in $\theta$ (with values on either side of the critical threshold).

### *3.2.4. Robustness and Criticality*

In Figure 16 we return to the *East-West* distribution, and show that possible long-run outcomes of the noiseless process other than regional variation (i.e., uniform and blinking) are unstable when small error is present. The simulations begin, respectively, in (1) the uniform state with everyone playing $B$, (2) everyone playing $A$, and (3) in the blinking cycle. Results for $\varepsilon = 0.01$ are displayed on the left; those for $\varepsilon = 0.05$ are on the right. In the $\varepsilon = 0.05$ case we get fairly quick convergence to the regional variation state. However, smaller noise ($\varepsilon = 0.01$) was not enough to upset the uniform and blinking states (at least during the length of the simulation).

Figure 17 is a robustness check on the value of probabilities (in the context of the *East-West* distribution). It also illustrates the "criticality" phenomenon: the existence of critical parameter thresholds that, when crossed, effect precipitous change in the behavior of the system. However, on

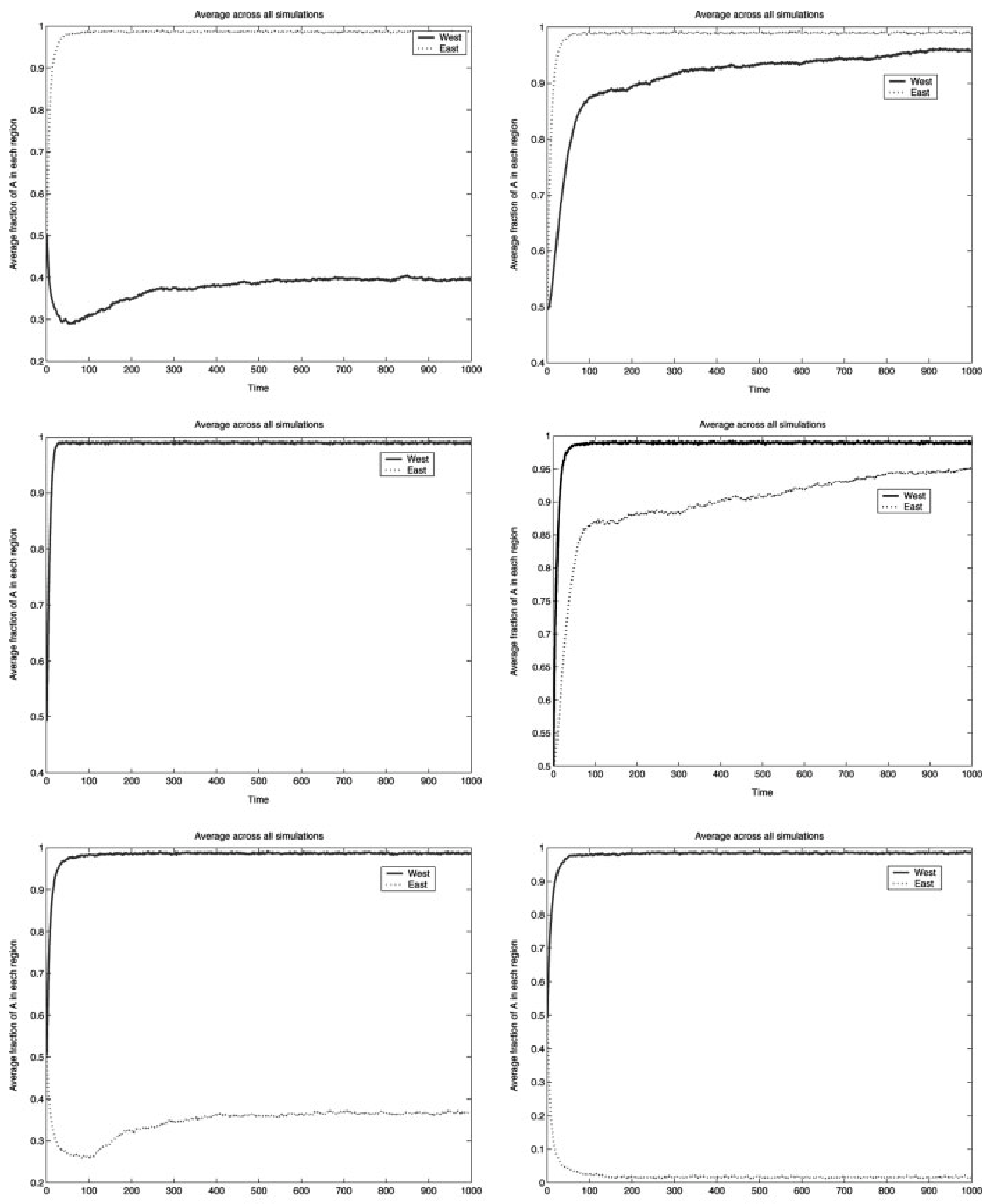
## FIGURE 16




Starting from stable states **B** (top), **A** (middle), and "blinking" (bottom). Simulations for $\varepsilon = 0.01$ are on the left and for $\varepsilon = 0.05$ are on the right.

either side of the critical point variation in fundamentals is largely irrelevant. The parameters that exhibit threshold effects in this model are $p$ and $q$. In the West, the prevailing norm depends on whether $q$ is greater than or less than 1/2. We start, in the top left-hand frame of Figure 17, where the value of $q$ is changed from 0.3 (used in all previous simulations) to 0.4 (with $p = 0.7$). The move closer to 1/2 has some effect as there are now more *A*'s in the West. This number increases as we change $q$ to 0.5 (top right). At $q = 0.7$ the graphs for East and West are virtually

**FIGURE 17**




Changing the probability of signals.

identical (middle left). Thereafter, we reduce $p$ to 0.5, and then to 0.4, and 0.3 (for symmetric effects). Changes further away from the 0.5 boundary do not have perceptible effects.

### 3.2.5. The Circle Model

In the circle model, unlike the torus, regional variation (*East-West*) is not a long-run steady-state of the noiseless process. In contrast to the *East-West* choice distribution for the torus, a boundary point of a choice region must have an equal number of neighbors making either choice. So choices at the boundary are always signal-dependent, and a judiciously chosen sequence of signals will lead to complete uniformity. A uniform state can be reached with nonzero probability, and once reached, will persist forever.

Although this is a bit like the diagonal case above, the results for the noiseless case point to an interesting difference. Despite the fact that regional variation is *not* a long-run outcome, we find that it arises and persist for multiple time periods (results not reported here). For large $M$ (e.g., $M = 100$) we have found it almost impossible to get to total uniformity once we are in the regional variation state. Convergence to one of the long-run steady states appears to be *very slow*. In the circle model as well, globally uniform choices are easily upset by the introduction of small noise and regional variation arises as the stable phenomenon.

## 4. CONCLUSION

The model tells us that alternative norms may coexist in close proximity to each other for indefinite periods of time. Despite a tendency toward local conformity, globally uniform states almost never occur and are unstable in the presence of small noise. There are several interesting details to this coexistence result. First, even when regional variation is not a long-run outcome of the dynamic process the system quickly organizes itself into this state, and any decay of the pattern is very slow. Second, we find that errors in decisionmaking, even when they are very small, can have important effects. For instance, the square distribution of choice appears to be stable when there is just a 1 in 100 chance of error. By contrast, there is fairly steady decay in the zero-noise case. Noise also decreases the likelihood of finding a blinking cycle. An exception is when the signal probability distribution is random—here mixed patterns that include the blinking cycle are typical. Third, the size of the network matters in a number of ways. Larger networks make regions with blinking patterns less likely for the East-West distribution, although not for the random distribution. They can also facilitate the survival of (long-run unstable) regional variation by making a transition to a uniform state a more remote possibility.

Our results show that the distribution of signals determines what the local norm will be. Specifically, the majority signal within a region will dictate the content of the norm. When the dominant signal differs across regions, local norms respond accordingly, and we witness global diversity as a stable phenomenon. The concept of "lock-in" has been very influential in the study of institutions and organizations, and our results in the noiseless case illustrate how it could arise. But small noise, as in the theory of Kandori et al. [22], Young [1], and Ellison [26], allows us to refine predictions considerably. New patterns of behavior, although they arise by error, can spread contagiously until they become locally prevalent. Lock-in does not have quite as tenacious a hold as in the noiseless case.

The locality of interactions is crucial for our results. In such settings, global majorities need not dictate global norms. Alternative norms will survive as long as there exist regions (or subgroups) within which the globally dominant signal or type forms a minority, even though interactions straddle regional boundaries. As informal evidence, we observe that minority languages are sustained by the presence of ethnic residential enclaves. We also observe ethnically and regionally specific slang and dress codes, as well as pockets of dedicated Mac users in a Windows-dominated world. Furthermore, as our criticality results indicate, stable coexistence does not require extreme differences in the composition of the population across regions or groups. The criticality of the 50% threshold means that norms can shift rapidly within a region with even a small change in demographics around the threshold. The results caution against making inferences about preferences, both within and across regions, based on observed behavior. A single norm can accommodate a diversity of types, just as relatively small changes in group characteristics may cause a discrete shift in the dominant behavior.

The condition of regional variation, involving local uniformity and global diversity, embeds certain social tensions. We have shown that within any region or group there may be a large number of agents (i.e., the minority type) who would be better off (a) living in a different region, in which their preferred norm prevails or (b) living in the same region but being a member of the majority type. In the signal interpretation, any given individual will face inferior payoffs whenever she receives the locally less common signal type. We have treated location and characteristics as exogenous. However, if geographic location or social network were made endogenous, we would expect self-selection into locations or networks by type, i.e., spontaneous physical or social segregation. Alternatively if locations remain fixed but type could change, preferences might adapt to surroundings.

Depending on the application under consideration, the degree to which location or characteristics are in fact endogenous will vary, as will the welfare implications. Although the prospect of spontaneous segregation echoes Schelling [27], in the context of our model, segregation by characteristics (i.e., the signal) may be socially preferred to an outcome in which regions or groups have a mix of types. Luckily, the implications need not be politically unsavory. For example, ignoring transport costs, it is desirable that medical patients be transferred to the treatment location that specializes in the treatment that best suits her type. In the case of corruption or corporate malfeasance, however, endogeneity may have negative consequences. We might expect, for example, that initially honest types, witnessing rampant corruption or receiving frequent invitations to embezzle, might eventually suffer moral decay. If so, replacing dishonest workers with honest workers on a piecemeal basis would be futile. Segregation and assimilation are not inevitable in every application, however. For example, graphics and text processing may be complements within a firm. If so the firm faces a tradeoff between compatibility across workers or tasks and supplying the best tool for each task. In our model the benefits of compatibility produce local uniformity, but an alternative outcome might involve innovations which render the opposing technologies more compatible.


## ACKNOWLEDGMENTS

We are grateful to an anonymous referee for very helpful suggestions on the design of the study; specifically, the investigation of section 3.2.3 was outlined by the referee.


## REFERENCES


1. Young, H.P. The evolution of conventions. Econometrica 1993, 61, 57–84.
2. Morris, S. Contagion. Rev Econ Studies 2000, 67(1), 57–78.
3. Glover, J.A. The incidence of tonsillectomy in school children. Proc Royal Soc Med 1938, 31, 95–112.
4. Wennberg, J.; Gittelsohn, A. Small area variations in health care delivery. Science 1973, 182, 1102–1108.
5. Wennberg, J.; Cooper, M., Eds. The Dartmouth Atlas of Health Care 1999; Dartmouth Medical School and American Hospital Association, Chicago, 1999.
6. Burke, M.A.; Fournier, G.M.; Prasad, K. Physician Social Networks and Geographical Variation in Medical Care, Florida State University working paper, 2004.
7. Young, H.P.; Burke, M.A. Competition and custom in economic contracts: A case study of Illinois agriculture. Am Econ Rev 2001, 91(3), 559–573.
8. Burke, M.A. On the Distributional Effects of Contractual Norms: The Case of Cropshare Contracts, unpublished manuscript, 2004, available on request from the author: mary.burke@bos.frb.org.
9. Arthur, W.B. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. Econ J 1989, 99(4), 116–131.
10. David, P. Clio and the economics of QWERTY. Am Econ Rev 1985, 75(2), 332–337.
11. Kremer, M. The O-ring theory of development. Quart J Econ 1993, 108(3), 551–575.
12. Selod, H.; Zenou, Y. Social interactions, ethnic minorities and urban unemployment. Annal Econ Statist 2001, 63–64, 183–214.
13. Young, H.P. Individual Strategy and Social Structure: An Evolutionary Theory of Institutions; Princeton University Press: Princeton, NJ, 1998.
14. Goyal, S. Interaction structure and social Change. J Institutional Theor Econ 1996, 152(3), 472–494.
15. Sugden, R. The Coexistence of Conventions. J Econ Behav Organization 1996, 28(2), 241–256.
16. Axtell, R.; Epstein, J.; Peyton Young, J. The emergence of economic classes in an agent-based bargaining model, In: Social Dynamics; Durlaul, S., Peyton Young, H.; Eds.; Oxford University Press, 1999.
17. Bowles, S.; J-K Choi, J.-K.; Hopfensitz, A. The coevolution of individual behaviors and social institutions. J Theor Biol 2003, 223, 135–147.
18. Epstein, J.M.; Axtell, R. Growing Artificial Societies: Social Science from the Bottom Up; MIT Press, 1996.
19. Epstein, J.M. Zones of cooperation in demographic prisoners' dilemma. Complexity 1998, 4(2), 36–48.
20. Epstein, J.M. Learning to be thoughtless: Social norms and individual computation, Comput Econ 2001, 18, 9–24.
21. Tesfatsion, L. Agent-based computational economics: Growing economies from the bottom up. Artif Life, 2002, 8(1), 55–82.
22. Kandori, M.; Mailath, G.; Rob, R. Learning, mutation, and long run equilibria in games. Econometrica 1993, 61, 29–56.
23. Liggett, T.M. Stochastic Interacting Systems: Contact, Voter and Exclusion Processes; Heidelberg: Springer-Verlag, 1999.
24. Schinazi, R. Classical and Spatial Stochastic Processes; Birkhauser, 1999.
25. Blume, L.E. The statistical mechanics of strategic interaction. Games Econ Behav 1993, 4, 387–424.
26. Ellison, G. Learning, local interaction and coordination. Econometrica 1993, 61(5), 1047–1073.
27. Schelling, T.C. Dynamic Models of Segregation. J Math Sociol 1971, 1, 143–186.