Computational Economics **18:** 9–24, 2001.

Learning to Be Thoughtless: Social Norms and Individual Computation

JOSHUA M. EPSTEIN
Economic Studies Program, The Brookings Institution, 1775 Massachusetts Ave., NW, Washington, D.C. 20036, U.S.A.; e-mail: jepstein@brook.edu and The External Faculty, Santa Fe Institute, U.S.A.

Abstract. This paper extends the literature on the evolution of norms with an agent-based model capturing a phenomenon that has been essentially ignored, namely that individual thought – or computing – is often inversely related to the strength of a social norm. Once a norm is entrenched, we conform thoughtlessly. In this model, agents learn how to behave (what norm to adopt), but – under a strategy I term Best Reply to Adaptive Sample Evidence – they also learn how much to think about how to behave. How much they are thinking affects how they behave, which – given how others behave – affects how much they think. In short, there is feedback between the social (inter-agent) and internal (intra-agent) dynamics. In addition, we generate the stylized facts regarding the spatio-temporal evolution of norms: local conformity, global diversity, and punctuated equilibria.

Key words: agent-based computational economics, evolution of norms

1. Two Features of Norms

When I'd had my coffee this morning and went upstairs to get dressed for work, I never considered being a nudist for the day. When I got in my car to drive to work, it never crossed my mind to drive on the left. And when I joined my colleagues at lunch, I did not consider eating my salad barehanded; without a thought, I used a fork.

The point here is that many social conventions have *two* features of interest. First, they are self-enforcing behavioral regularities (Lewis, 1969; Axelrod, 1986; Young, 1993a, 1995). But second, once entrenched, we conform *without thinking about it*. Indeed, this is one reason why social norms are useful; they obviate the need for a lot of individual computing. After all, if we had to go out and sample people on the street to see if nudism or dress were the norm, and then had to sample other drivers to see if left or right were the norm, and so on, we would spent most of the day figuring out how to operate, and we would not get much accomplished. Thoughtless conformity, while useful in such contexts, is frightening in others – as when norms of discrimination become entrenched. It seems to me that the literature on the evolution of norms and conventions has focused almost exclusively on the first feature of norms – that they are self-enforcing behavioral regularities, often

represented elegantly as equilibria of n-person coordination games possessing multiple pure-strategy Nash equilibria (Young 1993a, 1995; Kandori, Mailith, and Rob, 1991).

Goals

My aim here is to extend this literature with a simple agent-based model capturing the second feature noted above, that *individual thought – or computing – is inversely related to the strength of a social norm.* In this model, then, agents learn how to behave (what norm to adopt), but they also learn *how much to think about* how to behave. How much they are thinking affects how they behave, which – given how others behave – affects how much they think. In short, there is *feedback* between the social (inter-agent) and internal (intra-agent) dynamics. In addition, we are looking for the stylized facts regarding the spatio-temporal evolution of norms: local conformity, global diversity, and punctuated equilibria (Young, 1998).

2. An Agent-Based Computational Model

This model posits a ring of interacting agents. Each agent occupies a fixed position on the ring and is an object characterized by two attributes. One attribute is the agent's 'norm', which in this model is binary. We may think of these as 'drive on the right (R) vs. drive on the left (L)'. Initially, agents are assigned norms. Then, of course, agents update their norms based on observation of agents within some sampling radius. This radius is the second attribute and is typically heterogeneous across agents. An agent with a sampling radius of 5 takes data on the five agents to his left and the five agents to his right. Agents do not sample outside their current radius. Agents update, or 'adapt', their sampling radii incrementally according to the following simple rule:

Radius Update Rule

Imagine being an agent with current sampling radius of r. First, survey all r agents to the left and all r agents to the right. Some have L (drive on the left) as their norm and some have R (drive on the right). Compute the relative frequency of Rs at radius r; call the result $F(r)$. Now, make the same computation for radius $r+1$. If $F(r+1)$ *does not* equal $F(r)$, then increase your search radius to $r+1$.[1] Otherwise, compute $F(r-1)$. If $F(r-1)$ *does* equal $F(r)$, then reduce your search radius to $r-1$. If neither condition obtains (i.e., if $F(r+1) = F(r) \neq F(r-1)$), leave your search radius unchanged at r.

Agents are 'lazy statisticians', if you will. If they are getting a different result at a higher radius ($F(r+1) \neq F(r)$), they increase the radius – since, as statisticians, they know larger samples to be more reliable than smaller ones. But they are also lazy. Hence, if there's no difference at the higher radius, they check a lower one. If there is no difference between that and their current radius ($F(r-1) = F(r)$),