Organization Science
Vol. 18, No. 4, July–August 2007, pp. 648–666
ISSN 1047-7039 | EISSN 1526-5455 | 07 | 1804 | 0648
DOI 10.1287/orsc.1060.0245

Innovators, Imitators, and the Evolving Architecture of Problem-Solving Networks

Myong-Hun Chang
Department of Economics, Cleveland State University, Cleveland, Ohio 44115,
m.chang@csuohio.edu

Joseph E. Harrington, Jr.
Department of Economics, Johns Hopkins University, Baltimore, Maryland 21218,
joe.harrington@jhu.edu

Scientific progress is driven by innovation—which serves to produce a diversity of ideas—and imitation through a social network—which serves to diffuse these ideas. In this paper, we develop an agent-based computational model of this process, in which the agents in the population are heterogeneous in their abilities to innovate and imitate. The model incorporates three primary forces: the discovery of new ideas, the observation and adoption of these ideas, and the endogenous development of networks. The objective is to explore the evolving architecture of problem-solving networks and the critical roles that different agents play in the process. A central finding is that the emergent network takes a chain structure with innovators (those most skilled at generating new ideas) being the main source of ideas and those most skilled at imitating acting as connectors between the innovators and the masses. The impact of agent heterogeneity and environmental volatility on the network architecture is also characterized.

Key words: innovator; imitator; connector; problem-solving networks; network architecture; agent-based model

1. Introduction

The scientific revolution of the seventeenth century is often attributed to the genius of a few solitary innovators. The archival records from this period, however, reveal that the *social* dimensions of the revolution—for example, the social networks that connected these scientists through time and space—were just as critical in bringing the revolution to its ultimate victory (Hunter 1998). Hatch (1998) describes the extensive correspondence networks that were established and operated by a few human connectors during this period. These social connectors included N.-C. Fabri de Peiresc (1580–1637), Marin Mersenne (1588–1648), Samuel Hartlib (c. 1600–1662), Ismaël Boulliau (1605–1694), and Henry Oldenburg (1618–1677). Although they did not originate the paradigm-shifting ideas themselves, as *connectors* they facilitated the wide dissemination of ideas through communication networks of influential acquaintances and contacts. Hatch's description of Boulliau's network hints at the expanse over which the networks operated as well as the extent to which they were used: "Embracing the humanist ideal of community and communication, [Ismaël] Boulliau established a decidedly scientific and European network. …Boulliau's correspondence network included some 4,200 letters for the years 1632–93;…[it] marks a critical transition in geographical distribution, which now extended beyond France, Holland, and Italy, to Poland, Scandinavia, and the Levant" (Hatch 1998, p. 55). In fact, luminaries such as Galileo, Huygens, Dupuy, Mersenne, Oldenburg, and Fermat were all connected to Boulliau's network.

Hull (1988), in proposing an evolutionary model of the dynamic process by which scientific progress is made, positions these social connectors at center stage with the innovators:

> According to the model that I am proposing, both discovery and dissemination are necessary…. As Lamarck (1809, p. 404) ruefully concluded his *Philosophie zoologique*, "Men who strive in their works to push back the limits of human knowledge know well that it is not enough to discover and prove a useful truth previously unknown, but that it is necessary also to be able to propagate it and get it recognized." …If science is a selection process, transmission is necessary. Disseminators are operative in this process. Perhaps they do not get the ceremonial citations that patron saints do, but they are liable to get much more in the way of substantive citations…. To the extent that disseminators substitute their own views for the patron saints whom they cite ceremoniously, they are functioning as germ-line parasites—the cowbirds of science. (Hull 1988, pp. 376–377)

Knowledge sharing through informal networks is not only relevant to grand questions in science, but also to everyday problem solving in organizations. Many organizational studies have documented the significant role that networks play in both sharing current knowledge and contributing to the creation of new knowledge (see Cross and Parker 2004, and the collection of papers in Cross et al. 2003). In one study, managers reported that information from other people is far more instrumental in a project's success than that gained from impersonal sources such as knowledge databases (Cross et al. 2001).

Furthermore, within this network of information sources, middle-level managers are critical for facilitating information flow even though the formal hierarchy would not have suggested such a significant role. Given the high centrality of some agents in problem-solving networks, research has begun to explore the characteristics of these connectors (see, for example, Borgatti and Cross 2003,[1] Klein et al. 2004, Uzzi and Dunlap 2005).

The empirical literature then finds that networks are instrumental in problem solving and that some agents have high centrality and act as *connectors* within the network. These observations generate a number of questions. Although there are agents with high centrality, how critical are they to the group's performance? If agents are acting in a decentralized manner, can we generally expect connectors to emerge? In other words, is their emergence typical or rare? What factors determine whether connectors emerge? What are the characteristics of connectors? Are connectors just people who, through happenstance, find themselves in a position of centrality, which then perpetuates itself—people link with connectors because connectors have a lot of knowledge, but the reason they have a lot of knowledge is that people link with them—or do connectors have special skills attuned to playing this role? Related to this point, Sparrowe et al. (2001) find that agents with higher centrality have higher individual performance. But what is the causal relationship between these variables? Is this due to the agent's centrality resulting in higher performance, or do these individuals have certain fundamental traits that may lead to both higher performance and higher centrality?

To address these questions, we develop a computational agent-based model—first introduced in Chang and Harrington (2005) for the homogeneous agents case—with several significant features.[2] First, we consider learning in a generic problem-solving environment. Rather than tailor the model to a particular setting, the intent is to derive some basic properties of networks that are broadly applicable. Second, the model allows the network to be endogenously formed through the decisions of individual agents regarding with whom they form links. In this way, we can assess the regularity with which connectors emerge.

Third, agents are modelled as allocating effort between innovation—discovering new ideas in isolation—and imitation—linking with other agents to learn what they know. Our model, then, allows for both knowledge transfer and creation; furthermore, these processes are intertwined, as agents can take new ideas they come up with and put them together with the ideas of others to derive a new solution to a problem. Fourth, agents are heterogeneous in terms of their innovative and imitative skills. Specifically, the population is comprised of three types: *Innovators* are highly productive in generating new ideas, *imitators* are highly productive in identifying the ideas of others, and *regular agents* are moderately productive at both activities. Such heterogeneity in skills, of course, has the first-order impact on the choices made by individuals in terms of the learning mechanism used for solving problems—innovation versus imitation. However, more importantly, the extent of the skill differentials and the distribution of the heterogeneous skills within the population tend to shape the architecture of the networks that evolve—i.e., who learns from whom and with how much intensity. This feature of our model, hence, allows us to explore both how the emergence of connectors depends on the distribution of the types as defined above and what the fundamental characteristics of connectors are when they do emerge in the networks. The relative contribution that *innovators* and *imitators* play in the performance of the overall population in solving problems can also be assessed in this framework.

Finally, we consider the interactions of the two learning mechanisms—innovation and imitation—in an environment within which the nature of the problem to be solved changes from one period to the next at some fixed rate. In a product market setting, a firm's problem changes with what its competitors are doing, as well as with consumer preferences. Thus, a more volatile environment means that competitors and consumers are less stable. Or, if the problem is to find a new vaccine, environmental volatility corresponds to the rate of mutation of the virus. Such volatility in the problem-solving environment determines the rate at which newly adopted ideas become obsolete over time, which, in turn, affects the value to an agent of learning through innovation relative to that of learning through imitation via network. The endogenous formation of networks and the emergence of connectors in the population are then directly influenced by the degree of stability in the problem-solving environment.

There are several significant findings. First, connectors do indeed naturally emerge and, in addition, it is *imitators* that assume this role and act as conduits between *innovators* and *regular agents*. That is, rather than directly connecting to *innovators*, *regular agents* evolve to connect to *imitators* to learn new information. That some agents have high centrality is then an emergent property and a decentralized process—i.e., one in which each individual agent adjusts his or her network connections autonomously—results in those agents who are best equipped to be connectors taking on this task. In exploring the socially optimal mix of types, group performance is maximized when there is a mix of *innovators* and *imitators*. This result is important, as it shows that certain agents—who are not themselves a source of innovation—can be critical in problem-solving networks.

We also derive results concerning the circumstances under which a network structure with connectors can be expected to emerge. Here, we focus on environmental stability as a driving factor, which is represented as the rate of change of the problem that agents are trying to solve. We find that connectors are less likely to emerge when the environment is less stable. When the environment changes rapidly, the returns to *regular agents* of connecting with *imitators* tend to be low, as the ideas of *imitators* available for copying come from a biased sample that is no longer suited for the new environment (that is, the new problem to solve). They are better off generating the ideas themselves, and this leads to an underdeveloped network with a diminished role for connectors. Hence, the issue of identifying who are critical connectors and maintaining them in the network is more important in more stable environments.

In concluding this introduction, let us relate this paper to two literatures. First, there is the work on learning in networks (some of which was mentioned above). As noted in the review article of Podolny and Page (1998), there are two ways in which networks can enhance learning. First, there is the diffusion of new ideas, for which there is a considerable body of work examining diffusion networks (e.g., the classic work of Rogers 2003). This work typically takes the network as fixed and explores how the structure of the network impacts the rate of diffusion of a given innovation. Second, new knowledge can be created by bringing together information at different nodes of a network. Work encompassing that feature is rare, although there has been some recent work related to the model of this paper (for a review, see Chang and Harrington 2006). Both of these learning forces are present in our model. Furthermore, our model is unique in that it endogenizes both the network and the innovation. In fact, as described above in the context of knowledge sharing and creation, innovations are themselves a product of the network that helps diffuse them.[3]

Within the literature on learning in networks, research in sociology is rich in empirics, although formal modelling is rare. There is, however, an extensive theoretical literature on networks in economics (for a review, see Jackson 2006). The subset of this work that is concerned with the dynamics of network formation (for a review, see Goyal 2005) is distinct from our approach in a substantive way. Economic models assume that the value of a particular network structure is fixed and explore how agents adjust their links to form a better network. In contrast, the value of a given network is endogenous in our model, because it depends on what agents know, and what is known evolves over time because of innovation and imitation through the network. The economic model is more appropriate for networks such as friendship, while ours is more designed to address problem solving. Alternatively stated, the learning that occurs in economic models is about what is the right network. In our model, there is that sort of learning as well, but there is also learning about how to solve a problem; it is noteworthy that the two learning processes are intertwined.

As our model encompasses the decision about whether to exert effort on innovation or imitation, it is also relevant to the literature on exploration and exploitation. The crux of that literature is understanding what determines whether agents engage in discovering new ideas or in exploiting existing ideas (such as through the diffusion of a new idea) and to what extent agents engage in a socially optimal mixture of these activities. A classic paper here is by March (1991), who assumes there is an organizational code that adapts over time and determines the relative rates of exploration and exploitation. The code influences what agents do, but the code itself is also influenced by the actions of better-performing agents. Thus, the code is a device that passes along the better ideas but is also a product of what is being done. Our model gets inside the "black box" of the organizational code by replacing it with a network—agents do not learn from the code but rather from other agents through their endogenously created network. Although the role of the code is exogenous in March's work, the role of the network is endogenous in our model, because agents can decide how many resources to put into developing a network and how many into developing their own ideas. The exploration/exploitation trade-off is also examined in Siggelkow and Levinthal (2003), although their focus is on the role of the formal hierarchy (in the allocation of authority) in problem solving. We replace the formal hierarchy with its fixed links with an informal endogenous network, and thus address a different set of questions related to the properties of emergent networks and individual agents' roles in the network.

Another related work is Haas (2006). In the context of transnational teams working on knowledge-intensive projects, she looks at the roles of "cosmopolitans" and "locals" in the process of acquiring and applying both internal and external knowledge to their teams. The exploration/exploitation trade-off is also behind one of the results, which is that the ideal team composition entails a *mix* of both types.

The model is described in §2, and how we conduct our computational experiments is reviewed in §3. Results pertaining to the emergent properties of networks are presented in §§4 and 5. In §6, we address the issue of the optimal mix of agent types and relate these findings to the role of various agents in the emergent structure. Section 7 concludes by suggesting some managerial implications of our work and describing how our model can be modified to address other questions that have been raised in the literature on networks and organizations.

2. The Model

2.1. Agents, Tasks, Goal, and Performance

The social system consists of L individuals. Each individual engages in an operation that can be broken down into H separate tasks. There are several different methods that can be used to perform each task. The method an agent chooses for a given task is represented by a sequence of d bits (zero or one) such that there are 2^d possible methods available for each task. Let $\underline{z}_i^h(t)$ denote the method used by individual i in task h in period t. In any period t, an individual i is then fully characterized by a binary vector of $H \cdot d$ dimensions, which we denote by $\underline{z}_i(t)$, where $\underline{z}_i(t)$ is a connected sequence of methods, $\underline{z}_i^1(t), \underline{z}_i^2(t), \ldots,$ and $\underline{z}_i^H(t)$—one method (a string of d bits) for each task. To be more concrete, consider an operation having five separate tasks with four dimensions to each task so that $H = 5$ and $d = 4$:

Task (h):	#1	#2	#3	#4	#5
Methods ($\underline{z}_i^h(t)$):	1001	1101	0001	1010	0101

There are 16 (2^4) different methods for each task. Because the operation is completely described by a vector of 20 (5×4) bits, there are 2^{20} possible bit configurations (i.e., methods vectors) for the overall operation.

The degree of heterogeneity between two methods vectors, $\underline{z}_i$ and $\underline{z}_j$, is measured using *Hamming distance*, which is defined as the number of positions for which the corresponding bits differ. We shall denote it by $D(\underline{z}_i, \underline{z}_j)$.

In period t, the population faces a common goal vector, $\hat{\underline{z}}(t)$, which is also a binary vector of $H \cdot d$ dimensions. The degree of turbulence in task environments is captured by intertemporal variability in $\hat{\underline{z}}(t)$, the details of which are explained in §2.4.

The individuals are uninformed about the goal vector $\hat{\underline{z}}(t)$ ex ante but engage in "search" to get as close to it as possible. Given H tasks with d bits in each task and the goal vector $\hat{\underline{z}}(t)$, the period t performance of individual i is then measured by $\pi_i(t)$, where

$$\pi_i(t) = H \cdot d - D(\underline{z}_i(t), \hat{\underline{z}}(t)). \tag{1}$$

Hence, the performance of agent i is greater as the Hamming distance to the goal vector is shorter.[4] The performance of a social system is measured by how close the individuals are to the common goal. We let $\widehat{\pi}(t)$ denote the aggregate social performance in period t such that it is the simple sum of L agents' performance levels in t: $\widehat{\pi}(t) = \sum_{i=1}^{L} \pi_i(t)$.

2.2. Modeling Innovation and Imitation

In a given period, an individual's search for the current optimum is carried out through two distinct mechanisms: innovation and imitation. Innovation occurs when an individual independently discovers and considers for implementation a random method for a randomly chosen task. Imitation is when an individual selects someone (probabilistically) and then observes and considers implementing the method currently deployed by that agent for one randomly chosen task.

Although each act of innovation or imitation is assumed to be a single task, this is without loss of generality: If we choose to define a task as including d' dimensions, the case of a single act of innovation or imitation involving two tasks can be handled by setting $d = 2d'$.[5] In essence, what we are calling a "task" is defined as the unit of discovery or observation. The actual substantive condition is instead the relationship between d and H, as an agent's innovation or imitation involves a smaller part of the possible solution when d/H is smaller.

Whether obtained through innovation or imitation, an experimental method is actually adopted if and only if its adoption brings the agent closer to the goal by decreasing the Hamming distance between the agent's new methods vector and the goal vector. For clarity, let us consider the following example with $H = 5$ and $d = 2$:

common goal vector: 01 | 10 | 10 | 01 | 01
agent i's current methods vector: 01 | 01 | 11 | 00 | 11

The relevant operation has five tasks. In each task, there are four distinct methods that can be tried: (0, 0), (0, 1), (1, 0), and (1, 1). Agent i with the above current methods vector is then employing the method (0, 1) for task 1, (0, 1) for task 2, (1, 1) for task 3, (0, 0) for task 4, and (1, 1) for task 5. The Hamming distance between i's current methods vector and the goal vector is four. Suppose i chooses to innovate in task 1. For task 1, agent i randomly selects a method from the set of all available methods, $\{(0,0), (0,1), (1,0), (1,1)\}$. Let us assume that i comes up with the idea of (1, 1) for task 1. The experimental methods vector for agent i is then

agent i's experimental methods vector:
11 | 01 | 11 | 00 | 11,

where the method (0, 1) in task 1 is replaced with (1, 1). This raises the Hamming distance to the goal vector from four to five, and hence is rejected by the agent. Alternatively, suppose that agent i chooses to imitate and ends up observing the method used for task 4 by another agent j ($\neq i$) whose methods vector is

agent j's current methods vector: 10 | 10 | 11 | 01 | 01.

Because j's method in task 4 is (0, 1), when it is tried by agent i, his or her experimental methods vector becomes

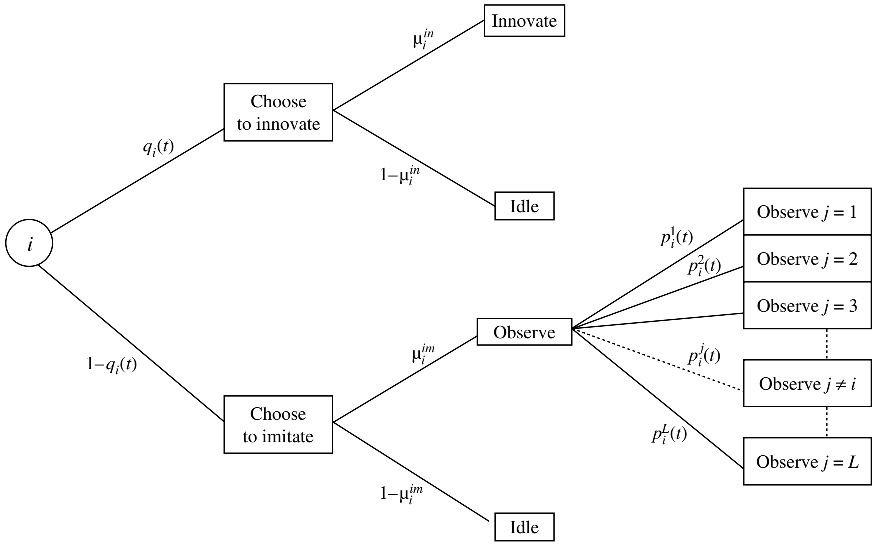
agent i's experimental methods vector:
01 | 01 | 11 | 01 | 11,

which then reduces the Hamming distance to the goal vector to three; hence, the experimental methods vector becomes i's new methods vector.

2.3. Endogenizing Choices for Innovation and Imitation

We assume that in each period an individual may engage in *either* innovation *or* imitation by using the network. How exactly does an individual choose between innovation and imitation and, if he or she chooses to imitate, how does he or she choose whom to imitate? We model this as a two-stage stochastic decision process with rein-

Figure 1 Decision Sequence of Individual i Period t



forcement learning. Figure 1 describes the timing of decisions in our model. In stage 1 of period t, individual i is in possession of the current methods vector, $\underline{z}_i(t)$, and chooses to innovate with probability $q_i(t)$ and imitate with probability $1-q_i(t)$. If the agent chooses to innovate, then, with probability μ_i^{in}, he or she generates an idea that is a randomly chosen task $h \in \{1, \ldots, H\}$ and a randomly chosen method for that task such that the experimental method vector, denoted $\underline{z}_i'(t)$, has the same methods as $\underline{z}_i(t)$ in all tasks except for the chosen task h. The method for the chosen task h will be replaced with the randomly chosen method, as explained in the example provided in the previous subsection. This experimental vector is adopted by i if and only if its adoption decreases the Hamming distance between the agent and the current goal vector, $\hat{\underline{z}}(t)$, in which case the methods vector in period $t+1$ is the experimental vector, $\underline{z}_i'(t)$. Otherwise, the experimental vector is discarded and the methods vector in $t+1$ is the same as $\underline{z}_i(t)$.[6] Alternatively, when the individual fails to generate an idea, which occurs with probability $1-\mu_i^{in}$, the methods vector in $t+1$ remains the same as $\underline{z}_i(t)$.

Now suppose individual i chooses to imitate in stage 1. Given that the agent decides to imitate someone else, he or she taps into the network to make an observation. Tapping into the network is also a probabilistic event, in which with probability μ_i^{im} the agent is connected to the network, while with probability $1-\mu_i^{im}$ the agent fails to connect. An agent who is connected then enters stage 2 of the decision process, in which he or she must select another agent to be studied for possible imitation. Let $p_i^j(t)$ be the probability with which i observes j in period t, so $\sum_{j\neq i} p_i^j(t) = 1$ for all i. If agent i observes another agent l, that observation involves a randomly chosen task h and the current method used by agent l in that task, $\underline{z}_l^h(t)$. Let $\underline{z}_i''(t)$ be the experimental vector such that it has the same methods as in $\underline{z}_i(t)$ for all tasks except for task h, and the method in h is replaced with $\underline{z}_l^h(t)$. Adoption or rejection of the observed method is based on the Hamming distance criterion, such that it is adopted if and only if it reduces the Hamming distance to the goal vector $\hat{\underline{z}}(t)$; the new methods vector in $t+1$ is, hence, the experimental vector, $\underline{z}_i''(t)$, in the case of adoption. Otherwise, it remains the same as $\underline{z}_i(t)$. Again, if the agent fails to connect to the network, which occurs with probability $1-\mu_i^{im}$, the new methods vector remains the same as $\underline{z}_i(t)$.

The probabilities, $q_i(t)$ and $\{p_i^1(t), \ldots, p_i^{i-1}(t), p_i^{i+1}(t), \ldots, p_i^L(t)\}$, are adjusted over time by individual agents according to a reinforcement learning rule. We adopt a version of the *experience-weighted attraction* (*EWA*) learning rule as described in Camerer and Ho (1999). Under this rule, an agent has a numerical attraction for each possible action. The learning rule specifies how attractions are updated by the agent's experience and how the probabilities of choosing different actions depend on attractions. The main feature of the rule is that *a positive outcome realized from a course of action reinforces the likelihood of that same action being chosen again.*

Using the EWA rule, $q_i(t)$ is adjusted each period on the basis of evolving attraction measures, $A_i^{in}(t)$ for innovation and $A_i^{im}(t)$ for imitation. The following process drives the evolution of $A_i^{in}(t)$ and $A_i^{im}(t)$. If the agent chose to pursue *innovation* and discovered and then adopted the new idea, the attraction measure for *innovation* increases by one after allowing for the decay factor of ϕ $(0<\phi\leq 1)$ on the previous attraction level—that is, $A_i^{in}(t+1) = \phi A_i^{in}(t)+1$. If the agent chose to innovate but was unsuccessful (either because he or she failed to generate an idea or because the idea generated was not useful) or if the agent instead chose to imitate, then the attraction measure for innovation is simply the attraction level from the previous period decayed by the factor ϕ—that is, $A_i^{in}(t+1) = \phi A_i^{in}(t)$. Similarly, a success or failure in imitation at t has the identical influence on $A_i^{im}(t+1)$, such that $A_i^{im}(t+1) = \phi A_i^{im}(t) + 1$ if i adopted a method through imitation in t, while $A_i^{im}(t+1) = \phi A_i^{im}(t)$ otherwise. Given $A_i^{in}(t)$ and $A_i^{im}(t)$, one derives the choice probability of innovation in period t as follows:

$$q_i(t) = \frac{(A_i^{in}(t))^\lambda}{(A_i^{in}(t))^\lambda + (A_i^{im}(t))^\lambda}, \qquad (2)$$

where $\lambda > 0$. The parameter λ measures sensitivity of players to attractions. A high value of λ means that a single success has more of an impact on the likelihood of repeating that activity (innovation or imitation).[7] The probability of imitation is, of course, $1 - q_i(t)$. The expression in (2) says that a favorable experience through innovation (imitation) raises the probability that an agent will choose to innovate (imitate) again in the future.

The stage 2 attractions and the probabilities are derived in a similar manner. Let $B_i^j(t)$ be agent i's attraction to another agent j in period t. Its evolution follows the same rule as that of $A_i^{in}(t)$ and $A_i^{im}(t)$, in that $B_i^j(t+1) = \phi B_i^j(t) + 1$ if agent i successfully imitated another agent j in t, while $B_i^j(t+1) = \phi B_i^j(t)$ otherwise. The probability that agent i observes agent j in period t is adjusted each period on the basis of the attraction measures, $\{B_i^j(t)\}_{j \neq i}$:

$$p_i^j(t) = \frac{(B_i^j(t))^\lambda}{\sum_{h \neq i}(B_i^h(t))^\lambda} \qquad (3)$$

for all i and for all $j \neq i$, where $\lambda > 0$.[8] Agent i's success in imitating another agent j then further raises the probability that the same agent will be observed again relative to others.[9]

There are two distinct sets of probabilities in our model. One set of probabilities, $q_i(t)$ and $\{p_i^j(t)\}_{j \neq i}$, is endogenously derived and evolves over time in response to the personal experiences of agent i. Another set of probabilities, μ_i^{in} and μ_i^{im}, is exogenously specified and is imposed on the model as parameters. They control the capabilities of individual agents to independently innovate or to imitate someone else in the population via social learning. It is particularly interesting to understand how these parameters influence the structure and performance of the network.

2.4. Modeling Turbulence in Task Environment

If agents only faced one fixed problem, then eventually all would end up at the global optimum. In that case, the measure of performance is the speed with which the optimum is achieved. Although there are some problems like that—for example, the race for a drug—most organizations and societies face ongoing challenges. This is surely the case with business organizations, that face a series of problems, and the current problem that they are working on may change because of, for example, the actions of competing companies or technological advances in another industry. Rather than model agents as facing a fixed problem, we choose to model them as facing a series of related problems. For analytical tractability, this is done by allowing the problem itself to evolve stochastically over time. Performance is then measured by the average quality of the network's solution rather than simply by the speed with which a problem is solved. In addition, an important feature of a network ought to be how well it adapts to change and not simply how fast it solves a problem.

Change or turbulence is specified in our model by first assigning an initial goal vector, $\hat{z}(0)$, to the population and then specifying a dynamic process by which it shifts over time. In period t (including $t = 0$), all agents in the population have the common goal vector of $\underline{\hat{z}}(t)$. In period $t + 1$, the goal stays the same with probability σ—i.e., $\underline{\hat{z}}(t+1) = \underline{\hat{z}}(t)$—and changes with probability $(1 - \sigma)$. The goal in $t+1$, if different from $\underline{\hat{z}}(t)$, is then an *iid* selection from the set of points that lie within the Hamming distance ρ of $\underline{\hat{z}}(t)$. The goal vector for the population then stochastically shifts while remaining within Hamming distance ρ of the *current* goal. This allows us to control the possible size of the intertemporal change. The lower σ is and the greater ρ is, the more frequent and variable is the change, respectively, in the population's goal vector.

This section provided a detailed description of the model. The definitions of the parameters introduced in this section and referred to throughout the paper are provided in Table 1. The values of these parameters used in the computational experiments are also given in Table 1. The notation for the relevant endogenous variables is provided in Table 2.

3. Design of Computational Experiments

The underlying simulation model specifies that $H = 24$ and $d = 4$, so there are 96 total bits in a methods vector and more than 7.9×10^{28} ($\cong 2^{96}$) possibilities in the search space.

We assume a population of 50 individuals: $L = 50$.[10] The population is divided into three separate groups: *innovators*, *imitators*, and *regular agents*. Let N represent (and denote) the group of *innovators* and M the group of *imitators*. The group of *regular agents* is denoted as R. There are 10 super types, such that $|N| + |M| = 10$ and $|R| = 40$. The baseline case we consider initially assumes the following configuration of capabilities for the agents in these three groups: $(\mu_i^{in}, \mu_i^{im}) = (1, 0)$ for all i in N; $(\mu_i^{in}, \mu_i^{im}) = (0, 1)$ for all i in M; and $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.25)$ for all i in R. Later we will

Table 1 List of Parameters

Notations	Definitions	Parameter values considered
L	No. of agents in the population	50
H	No. of separate tasks	24
d	No. of dimensions per task	4
σ	Probability that the goal vector stays the same from t to $t+1$	{0.5, 0.7, 0.8, 0.9}
ρ	Maximum no. of dimensions in the goal vector that can change from t to $t+1$	{1, 4, 9}
μ_i^{in}	Probability that agent i generates an idea in any given period	{0, 0.25, 0.5, 0.75, 1}
μ_i^{im}	Probability that agent i taps into its network to imitate another agent	{0, 0.25, 0.5, 0.75, 1}
$A_i^{in}(0)$	Agent i's attraction for innovation in $t=0$	1
$A_i^{im}(0)$	Agent i's attraction for imitation in $t=0$	1
$B_i^j(0)$	Agent i's attraction to agent j in $t=0$	1
ϕ	Decay factor for attractions	1
λ	Sensitivity of agents to attractions	1
$\|N\|$	Size of the set, N, of *innovators*	{0, 1, 2, ..., 10}
$\|M\|$	Size of the set, M, of *imitators*	{0, 1, 2, ..., 10}
$\|R\|$	Size of the set, R, of *regular* agents	40

consider two extensions: (1) $(\mu_i^{in}, \mu_i^{im}) = (0.75, 0.25)$ for all i in N, $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.75)$ for all i in M, and $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.25)$ for all i in R; and (2) $(\mu_i^{in}, \mu_i^{im}) = (0.75, 0.25)$ for all i in N, $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.75)$ for all i in M, and $(\mu_i^{in}, \mu_i^{im}) = (0.5, 0.5)$ for all i in R. These extensions will allow us to check the robustness of the properties we identify in the baseline case.

We assume that the *initial* practices of the agents are completely homogeneous, so that $\underline{z}_i(0) = \underline{z}_j(0)\ \forall\, i \neq j$. This is to ensure that any social learning (imitation) occurring over the horizon under study entails only newly generated knowledge. Otherwise, the initial variation in the information levels of the agents will induce some imitation activities, introducing unnecessary random noise into the system.[11] The common initial methods vector is assumed to be an independent draw from $\{0, 1\}^{Hd}$.

The parameters affecting the endogenous variables are $|N|{:}|M|$—the composition of the super-type individuals in the population—as well as σ and ρ—the frequency and magnitude of the environmental changes for the population. Keeping the total size of the super types at 10,

Table 2 List of Endogenous Probabilities

Notations	Definitions
$q_i(t)$	Probability that agent i chooses to innovate in t
$1-q_i(t)$	Probability that agent i chooses to imitate in t
$p_i^j(t)$	Probability that agent i observes agent j in t
f_{rs}	Steady-state probability with which an average agent in group r observes an average agent in group s

we consider the ratio of $|N|{:}|M|$ from {10:0, 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, 2:8, 1:9, 0:10}. We consider values of σ from $\{0.5, 0.7, 0.8, 0.9\}$ and ρ from $\{1, 4, 9\}$.

Additional parameters are ϕ (decay factor) and λ (sensitivity to attractions), which control the evolution of the attraction measures—that is, $A_i^{in}(t)$, $A_i^{im}(t)$, and $B_i^j(t)$. We assume that $\phi = 1$ and $\lambda = 1$ so that attractions do not decay and the agents are moderately sensitive to attractions. These values remain fixed over the relevant horizon. Finally, the initial attraction stocks are set at $B_i^j(0) = 1$ for all i and for all $j \neq i$, and $A_i^{in}(0) = A_i^{im}(0) = 1$ for all i. Hence, an individual in $t = 0$ is equally likely to engage in innovation and imitation—that is, $q_i(0) = 0.5$—and has no inclination to observe one individual over another ex ante—that is, $p_i^j(0) = 1/(L-1)$ ($=1/49 \cong 0.0204$ in our experiments) for all i and for all $j \neq i$.

All computational experiments carried out here assume a horizon of 15,000 periods. The time series of the performance measures is observed to reach a steady state by the 2,000th period.[12] We measure the steady-state performance of individual i, denoted $\overline{\pi}_i$, to be the average over the last 5,000 periods of this horizon so that $\overline{\pi}_i = (1/5{,}000) \sum_{t=10{,}001}^{15{,}000} \pi_i(t)$. The *aggregate* steady-state performance of the entire population is then denoted $\overline{\pi} \equiv \sum_{i=1}^{L} \overline{\pi}_i$. Likewise, the endogenous steady-state innovation probability, denoted $\overline{q}_i$, is computed for each agent as the average over the last 5,000 periods: $\overline{q}_i = (1/5{,}000) \sum_{t=10{,}001}^{15{,}000} q_i(t)$. Finally, the endogenous steady-state imitation probabilities, denoted $\overline{p}_i^j$, are computed to be the average over the last 5,000 periods: $\overline{p}_i^j = (1/5{,}000) \sum_{t=10{,}001}^{15{,}000} p_i^j(t)$.

All the experiments were based on 100 replications, each using a fresh set of random numbers.[13] Hence, the steady-state performance and probability measures reported are the averages over those 100 replications.

4. Results on the Evolving Architecture of Problem-Solving Networks: A Baseline Model

We start our analysis with the baseline case where $(\mu_i^{in}, \mu_i^{im}) = (1, 0)$ for all i in N, $(\mu_i^{in}, \mu_i^{im}) = (0, 1)$ for all i in M, and $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.25)$ for all i in R. Hence, *innovators* are true solitary geniuses who do not observe or copy other agents. The *imitators* are pure copycats with no ability to make independent discoveries. They rely exclusively on imitating someone else in the population through their networks. Finally, the *regular agents* have modest innate ability in both innovation and imitation.

In our model, the problem-solving network is defined in terms of the observation probabilities that the agents possess. Thus, we must examine the steady-state probabilities, $\overline{p}_i^j$s, to analyze the evolving architecture of the network. Recall that $\overline{p}_i^j$ is the probability with which

Figure 2 $\bar{P}_i^j$ **for** $\sigma = 0.8$ **and** $|N|:|M| = 4:6$

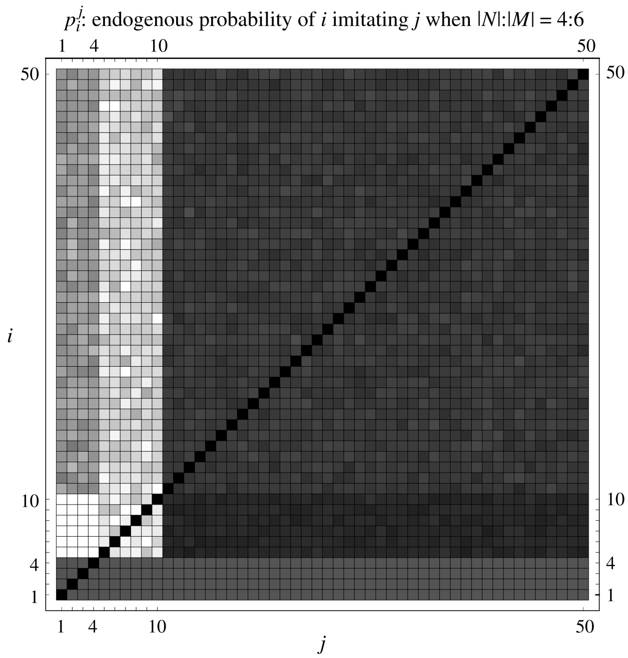
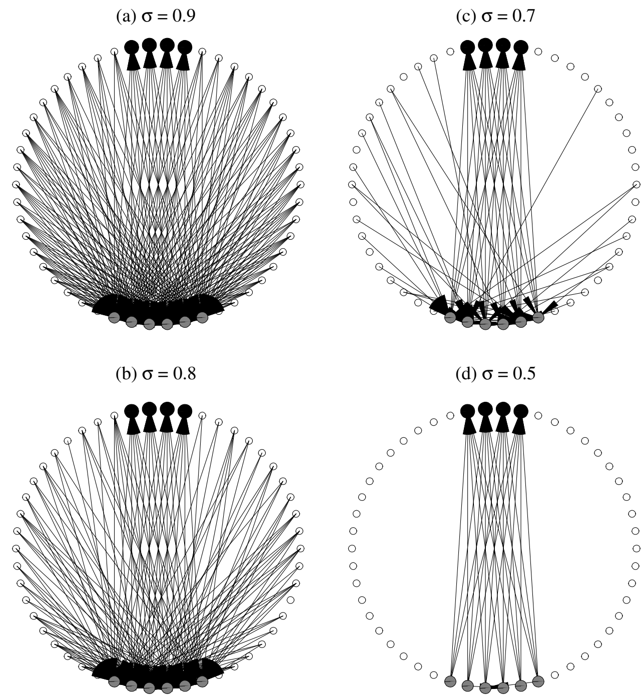


Figure 3 **Networks with Significant Links** ($\bar{p}_i^j > 0.05$)



agent i observes another agent j along the steady state. Given a population of 50 agents, each agent has these probabilities for 49 other agents. Figure 2 captures these probabilities for when $\sigma = 0.8$, $\rho = 1$, and $|N|:|M| = 4:6$. The vertical frame indicates the identity of the observer (agent i) and the horizontal frame the identity of the target (agent j). The figure visualizes the complete sets of probabilities for all 50 agents by representing the size of a probability with the brightness of a cell. The brighter (darker) the given cell, the higher (lower) the corresponding probability. The diagonal cells are completely black, as an agent observes itself with zero probability.

The simulation that generated the output for Figure 2 specifies that agents 1–4 are *innovators* (group N), agents 5–10 are *imitators* (group M), and agents 11–50 are *regular agents* (group R). One can immediately see that there is a unique structure to this network. The four *innovators* observe others (and themselves) with equal probabilities, which implies that their networks remain undeveloped throughout the entire horizon.[14] The six *imitators* (agents 5–10) observe the first four *innovators* with high probabilities, other *imitators* with somewhat lower probabilities, and the *regular agents* with the lowest probabilities. *Regular agents* (11–50) observe the *imitators* with high probabilities, the *innovators* with lower probabilities, and other *regular agents* with the lowest probabilities. This clearly suggests a *chain structure* to this network: *Innovators* engage in individual learning without any reliance on networks; *imitators* learn mainly from *innovators*; and *regular agents* learn mainly from *imitators*.[15] In this structure, *imitators* then play the role of *connectors* (between *innovators* and *regular agents*) by acting as the transmitters of ideas from *innovators* to the rest of the population.

An alternative way of visualizing the chain structure is presented in Figure 3, in which the networks consist of a set of nodes (agents) and links. More specifically, 50 agents (dots) are positioned along a circle. The four black dots on the upper part of the circle represent the *innovators* (agents 1–4), while the six grey-shaded dots on the lower part of the circle are the *imitators* (agents 5–10). The remaining hollow dots around the circle represent the 40 *regular agents*. A directed link (an arrow) is drawn from agent i to agent j if and only if the steady-state probability of observation, $\bar{p}_i^j$, is greater than some prespecified threshold level of $\hat{p}$. In Figure 3, $\hat{p} = 0.05$. Recall that a purely random network at $t = 0$ entails $p_i^j(0) = 1/49 \approx 0.0204$ for all i and j ($i \neq j$). By assuming $\hat{p} = 0.05$, we are defining a link from i to j as being significant if the network is sufficiently developed so that the steady-state probability of i observing j is greater than 0.05. Figures 3(a)–3(d) then capture the steady-state networks for when $\sigma \in \{0.9, 0.8, 0.7, 0.5\}$. In fact, Figure 3(b) uses the same probability data that generated Figure 2. The existence of the chain structure is quite clear. All the links coming out of the hollow dots (*regular agents*) are directed toward the grey-shaded dots (*imitators*). All the links out of the grey-shaded dots are directed towards the black dots (*innovators*) or to other grey-shaded dots. Furthermore, as σ decreases, the networks appear to be less well developed and the chain structure less pronounced.

How robust is this structural property, and how is it affected by the relevant parameters such as σ, ρ, and the mix of the super types, $|N|{:}|M|$? Given the enormous size of the probability sets among which we must make systematic comparisons, we simplify our analysis by eliminating redundant information. Because the observation probabilities among agents belonging to the same group are similar, we compute the probability with which an *average* agent in a given group observes an *average* agent in another group. Let f_{rs} denote the probability with which an average agent in group r observes an average agent in group s. Given three groups, $\{N, M, R\}$, we look for the probability with which an average agent in group g learns from an average agent in group g', where $g \in \{N, M, R\}$, $g' \in \{N, M, R\}$, and $g \neq g'$. Because an agent may also learn from other agents in his or her own group, we define two mean probabilities, f_{gg} and $f_{gg'}$, as follows:

$$f_{gg} = \frac{1}{|g|} \sum_{\forall i \in g} \left(\frac{1}{|g|-1} \right) \sum_{\substack{\forall j \in g \\ j \neq i}} \bar{p}_i^j \tag{4}$$

$$f_{gg'} = \frac{1}{|g|} \sum_{\forall i \in g} \left(\frac{1}{|g'|} \right) \sum_{\forall j \in g'} \bar{p}_i^j, \tag{5}$$

where $|g|$ is the size of group g. There are then nine different mean probabilities to be computed. Define a matrix F as the probability matrix showing all nine of them:

$$F = \begin{vmatrix} f_{NN} & f_{NM} & f_{NR} \\ f_{MN} & f_{MM} & f_{MR} \\ f_{RN} & f_{RM} & f_{RR} \end{vmatrix}. \tag{6}$$

For the baseline case, where *innovators* do not communicate at all, it is clear that $f_{NN} = f_{NM} = f_{NR}$. This is because *innovators* start out with networks that are completely undeveloped—i.e., they observe others with equal probabilities—and they never get to develop the networks over the horizon. However, the *imitators* in group M and the *regular agents* in group R do develop their networks, and the steady-state probabilities that ultimately emerge for these agents depend on all three parameters considered in this paper, σ, ρ, and $|N|{:}|M|$.

Shown in Figure 4 are the probability matrices, F, for $\sigma \in \{0.9, 0.8, 0.7, 0.5\}$ when $\rho = 1$ and the $|N|{:}|M| =$ 4:6. As expected, we get $f_{NN} = f_{NM} = f_{NR}$ $(=1/49 \approx 0.0204)$ for all cases. We also observe that $f_{MN} > f_{MM} > f_{MR}$ and $f_{RM} > f_{RN} > f_{RR}$ in all cases. The tendency for the network to take the chain structure is very clear: An agent in group M focuses mainly on observing an agent in group N, and an agent in group R focuses mainly on observing an agent in group M. It also appears that both f_{MN} and f_{RM} increase in σ, hinting at the possibility that the chainlike network structure is more pronounced in a more stable learning environment.[16] We have also examined these endogenous probabilities from various individual runs and verified that their properties are fully consistent with the ones from those averaged over 100 replications.

Figure 4 F **Matrix for** $\rho = 1$ **and** $|N|{:}|M| = 4{:}6$

$$F = \begin{pmatrix} f_{NN} & f_{NM} & f_{NR} \\ f_{MN} & f_{MM} & f_{MR} \\ f_{RN} & f_{RM} & f_{RR} \end{pmatrix}$$

(a) $\sigma = 0.9$: $\begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0873 & 0.0582 & 0.0090 \\ 0.0356 & 0.0554 & 0.0135 \end{pmatrix}$

(b) $\sigma = 0.8$: $\begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0843 & 0.0549 & 0.0097 \\ 0.0362 & 0.0512 & 0.0140 \end{pmatrix}$

(c) $\sigma = 0.7$: $\begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0815 & 0.0523 & 0.0103 \\ 0.0355 & 0.0457 & 0.0150 \end{pmatrix}$

(d) $\sigma = 0.5$: $\begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0776 & 0.0462 & 0.0115 \\ 0.0338 & 0.0383 & 0.0163 \end{pmatrix}$

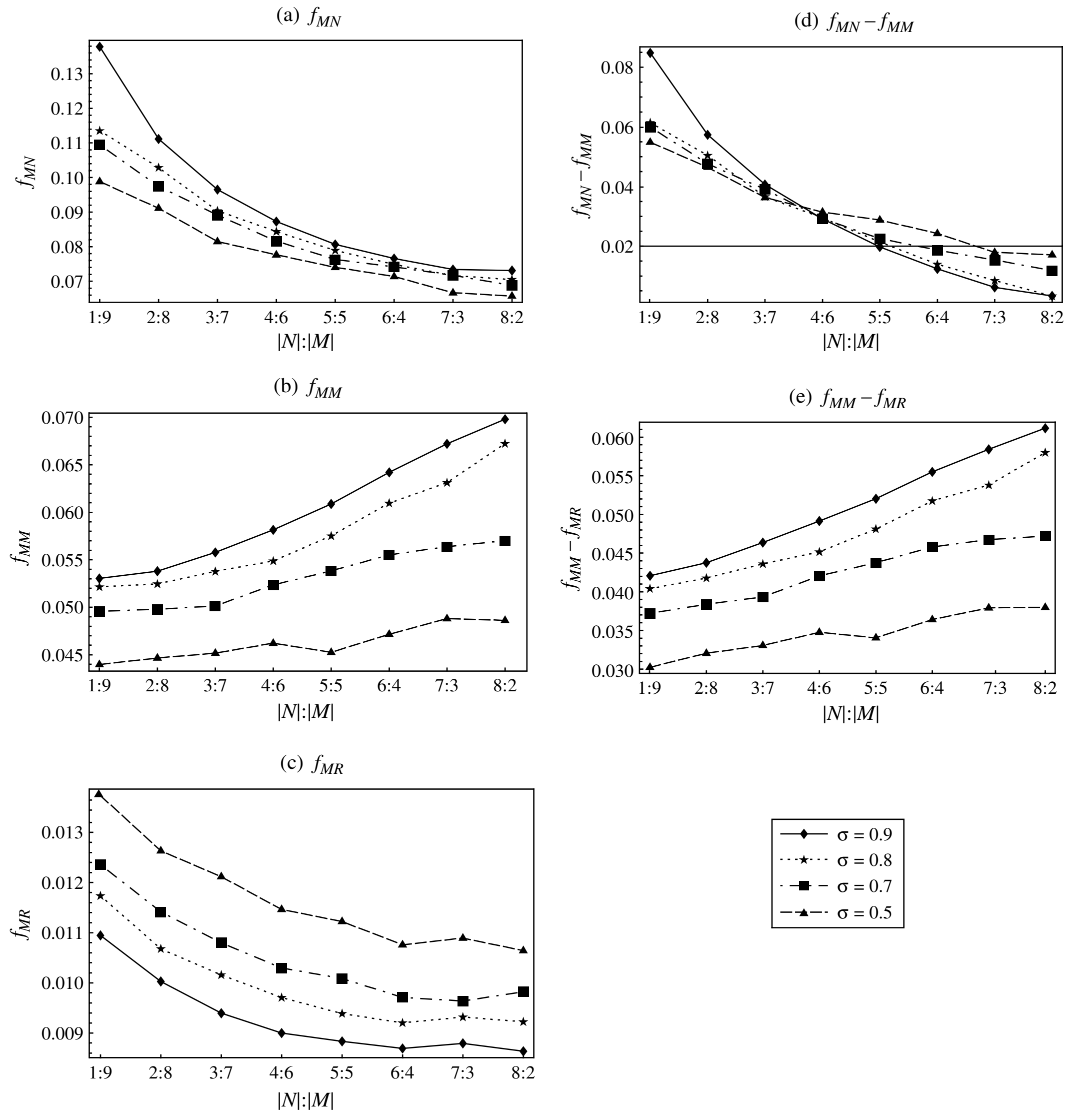
To confirm the generality of the properties observed in Figure 4 and to further explore the ways in which these probabilities respond to the changes in the parameter values, we resort to visualization of these probabilities in the next section. We focus on the observation probabilities of an *imitator*, $f_{M\cdot}$, and those of a *regular agent*, $f_{R\cdot}$ but ignore the *innovators* because they do not develop their networks in the baseline case.

4.1. Steady-State Architecture and the Impact of the Environmental Volatility

The collection of figures in Figure 5 displays the average individual probabilities, f_{MN}, f_{MM}, and f_{MR}, as well as their differentials, $(f_{MN} - f_{MM})$ and $(f_{MM} - f_{MR})$, for various σ and the $|N|{:}|M|$ mixes, given $\rho = 1$. Each figure plots the probability or the probability differential as a function of $|N|{:}|M|$ for all $\sigma \in \{0.9, 0.8, 0.7, 0.5\}$. The first observation to make is from Figures 5(d) and 5(e). They show that both $(f_{MN} - f_{MM})$ and $(f_{MM} - f_{MR})$ are strictly positive for all values of $|N|{:}|M|$ and σ, which implies that $f_{MN} > f_{MM} > f_{MR}$. Hence, an *imitator* observes an *innovator* with the highest probability, another *imitator* with a moderate probability, and a *regular agent* with the lowest probability.

Figures 6(a)–6(e) convey similar information on (f_{RN}, f_{RM}, f_{RR}) and $(f_{RM} - f_{RN}, f_{RN} - f_{RR})$. Both of the probability differentials plotted in 6(d) and 6(e) are again strictly positive for all values of $|N|{:}|M|$ and σ, which

Figure 5 Dependence of (f_{MN}, f_{MM}, f_{MR}) **on** $|N|{:}|M|$ **and** σ



implies that $f_{RM} > f_{RN} > f_{RR}$. A *regular agent* hence observes an *imitator* with a higher probability than he or she observes an *innovator*.

Figures 5 and 6 show that these properties hold for all σ and the $|N|{:}|M|$ mixtures considered in our simulations, given $\rho = 1$. In a similar fashion, Figure 7 (for *imitators*) and Figure 8 (for *regular agents*) plot the probabilities and their differentials as functions of the $|N|{:}|M|$ ratio and ρ, given $\sigma = 0.9$. Again, the previous properties hold for all $|N|{:}|M|$ ratios and all $\rho \in \{1, 4, 9\}$. Although not reported here, these properties also hold for various population sizes.[17]

Property 1. *When the* innovators *are solitary geniuses and the* imitators *are pure copycats, the social network evolves into a chain structure, where an* imitator *learns mainly from an* innovator *and a* regular agent *learns mainly from an* imitator: (a) $f_{MN} > f_{MM} > f_{MR}$ *and* (b) $f_{RM} > f_{RN} > f_{RR}$.

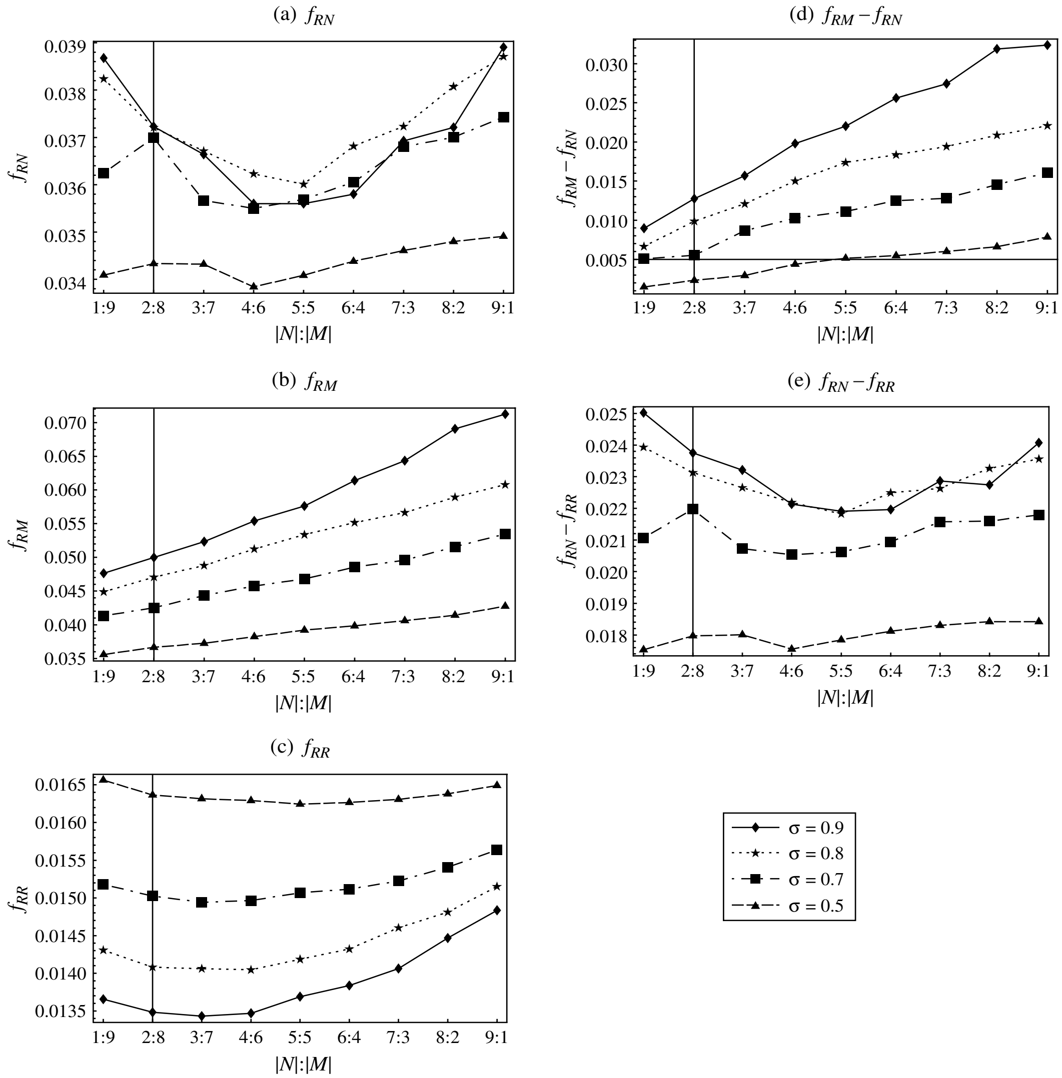
How does the degree of environmental volatility, as captured by σ and ρ, affect the chain-like network structure identified above? From Figures 5(a)–(c) and 7(a)–(c), we find that an increase in volatility (i.e., a decrease in σ or an increase in ρ) lowers f_{MN} and f_{MM}, while it raises f_{MR}. Similarly, Figures 6(a)–(c) and 8(a)–(c) show that an increase in volatility lowers f_{RM} and raises f_{RR}.[18]

Property 2. *When the environment is more volatile,* (1) imitators *shift the probability of observation from* innovators *and* imitators *to* regular agents, *and* (2) regular agents *shift the probability of observation from* imitators *to* regular agents.

The implication of Property 2 is that the chainlike network structure is more pronounced when the environment is more stable.

Having described our results concerning the network architecture for the baseline case, let us now dig deeper

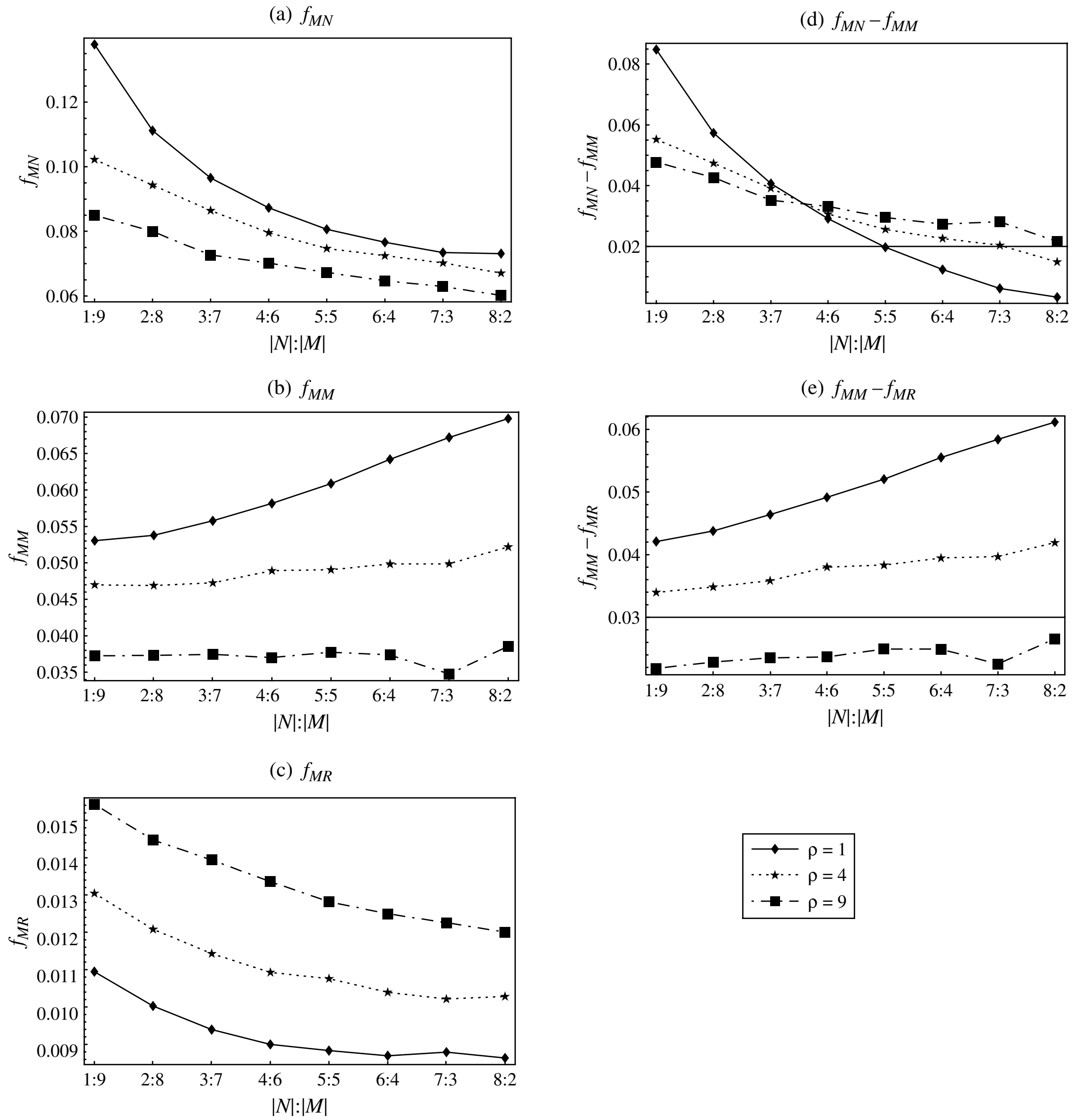
Figure 6 Dependence of (f_{RN}, f_{RM}, f_{RR}) **on** $|N|{:}|M|$ **and** σ



into the underlying mechanisms. What causes the network structure to evolve into a chain? Why do *imitators* choose to imitate *innovators* (rather than other *imitators*), while *regular agents* choose to imitate *imitators* (rather than *innovators*)? Why is this network structure more pronounced when the environment is more stable?

In contrasting the strategy of observing *imitators* with that of observing *innovators*, we must remember that these two agent types differ in terms of where their ideas originate.[19] *Innovators* make random draws from the entire space of ideas. *Imitators* make random draws from the ideas that are being used by others; those ideas are a mix of ones that have been selected and ones that are random (ideas that agents were endowed with and have not had a chance to change or those that are effectively random because they were adopted long ago when the environment was very different). If the environment is stable, the ideas of *imitators* should be better than the ideas of *innovators*, because the former type's implemented ideas come from a (favorably) biased sample. In contrast, when the environment is highly volatile, the ideas of *innovators* may be better because *imitators*' ideas come from an unfavorably biased sample, that is, ideas adopted for a different environment. One way to verify this claim is to examine the performance levels of a typical *innovator* and a typical *imitator* and see how they are affected by σ, the degree of environmental stability. The underlying motivation is that the level of an individual's performance is a direct indication of how well suited the implemented ideas are to the current environment. To this end, let us denote by $\overline{\overline{\pi}}^k$ the average performance of a *single* agent in group $k \in \{N, M, R\}$ such that $\overline{\overline{\pi}}^k = (1/|k|)\sum_{\forall i \in k} \overline{\pi}_i$. We have examined $\overline{\overline{\pi}}^M$ and $\overline{\overline{\pi}}^N$ for $\sigma \in \{0.5, 0.7, 0.8, 0.9\}$ and for three typical super-type mixes, $|N|{:}|M| \in \{8{:}2, 6{:}4, 4{:}6\}$. We found (as expected) that both $\overline{\overline{\pi}}^M$ and $\overline{\overline{\pi}}^N$ monotonically increase in σ so that

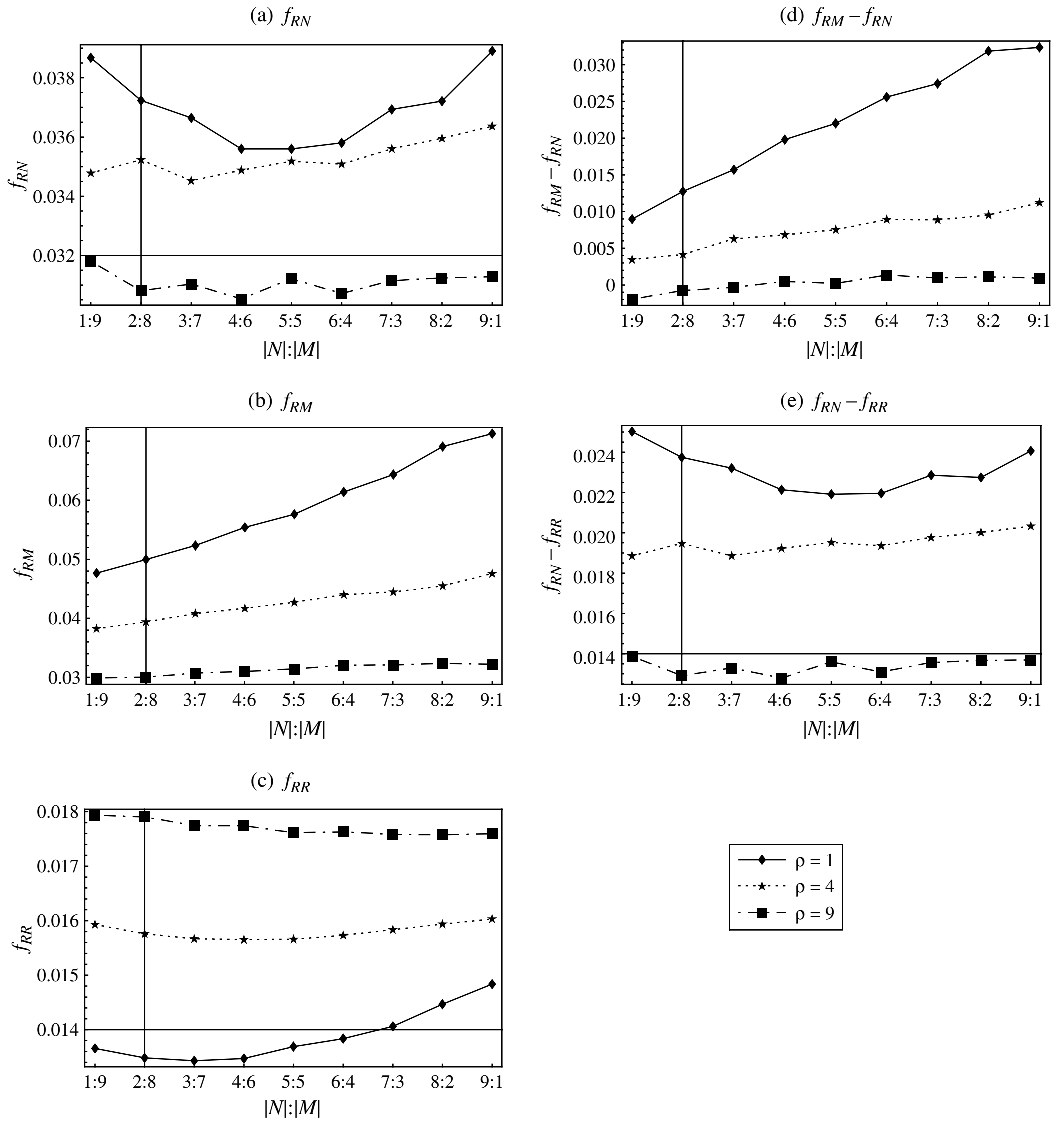
Figure 7 Dependence of (f_{MN}, f_{MM}, f_{MR}) **on** $|N|{:}|M|$ **and** ρ



a more stable environment raises the performance levels of all individuals. More interesting, however, is the performance differential between an *imitator* and an *innovator*. In Figure 9, we plot $(\overline{\pi}^M - \overline{\pi}^N)$—the performance advantage an *imitator* has over an *innovator*—as a function of σ for the three different mixtures of the super types. Note that $(\overline{\pi}^M - \overline{\pi}^N)$ is positive for all values of σ and $|N|{:}|M|$, implying that a typical *imitator* tends to outperform a typical *innovator* in all cases. Most important for our analysis, this differential is greater for a higher value of σ for all three super-type mixes. Although a greater stability in the problem-solving environment benefits everyone, it benefits an *imitator* to a greater extent than it does an *innovator*. This result then is consistent with the "biased-sample" argument presented earlier.

The above logic explains why *regular agents* prefer to imitate *imitators* when the environment is relatively stable—that is, σ close to one. When the environment is more volatile (σ low), note from Figure 6(a)–6(c) that *regular agents* imitate *imitators* less and other *regular agents* more, but they do not necessarily imitate *innovators* more. The biased-sample argument explains why *regular agents* imitate *imitators* less when the environment is more volatile. As to why they observe other *regular agents* more, it is useful to note that the network of a *regular agent* is less well formed when the environment is more volatile. Because they have a modest ability to innovate, they tend to shift probability from imitation to innovation when σ is lower. Figure 10 verifies this by plotting the steady-state probability of innovation, $\overline{q}_i$, for a *regular agent* as a function of $|N|{:}|M|$ ratio for all $\sigma \in \{0.5, 0.7, 0.8, 0.9\}$: $\overline{q}_i$, monotonically increases for all values of $|N|{:}|M|$, as σ is lower. This implies that *regular agents* spend less time developing

Figure 8 **Dependence of** (f_{RN}, f_{RM}, f_{RR}) **on** $|N|{:}|M|$ **and** ρ



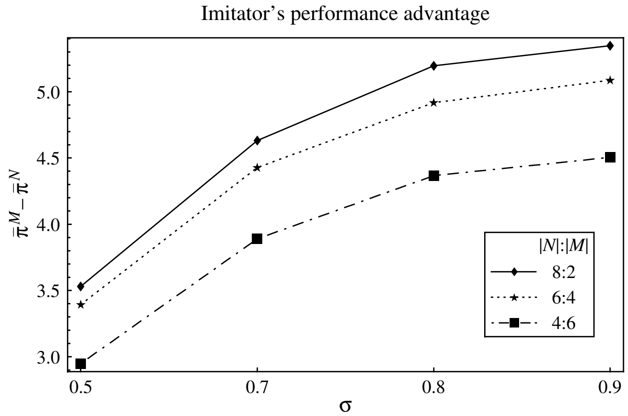
their social networks when the environment is more volatile.

The underdevelopment of the networks for *regular agents* explains why the links to other *regular agents* are stronger in a more volatile environment. But then why do *regular agents* not imitate *innovators* more (rather than *imitators*)?[20] There are two counteracting forces in this case. The sample bias argument for imitating *imitators* suggests that *regular agents* ought to move from imitating *imitators* to imitating *innovators* when σ is lower. The countervailing force is the underdevelopment of the *regular agents*' networks, which tends to shift the probability mass toward being uniform across all links.

The final property to be explained is why *imitators* prefer to imitate *innovators* rather than other *imitators*. Here it is important to recall that there are two conditions required for a particular round of imitation to lead to the adoption of an idea. First, the identified idea of another agent has to be worthwhile. Second, it cannot be an idea that an agent has already seen. As *regular agents* have a low rate of imitation—both because they are less productive when they do choose to imitate and because they also engage in innovation—it is relatively unlikely that they will see the same idea again. However, that is not necessarily the case with *imitators*. They imitate at a high rate, and if one were to imitate other *imitators* a lot, they may see the same ideas over and over again. In contrast, this is much less likely to occur when imitating *innovators*.

To summarize, *regular agents* imitate *imitators* because their ideas are better as a result of those ideas coming from a biased sample—ideas that have already met the approval of another agent; *innovators* draw their ideas randomly (although with both agent types, adoption is based on the idea's quality). *Imitators* imitate *innovators* because they are more likely to find new ideas, while

Figure 9 Performance Advantage of an *Imitator* over an *Innovator*

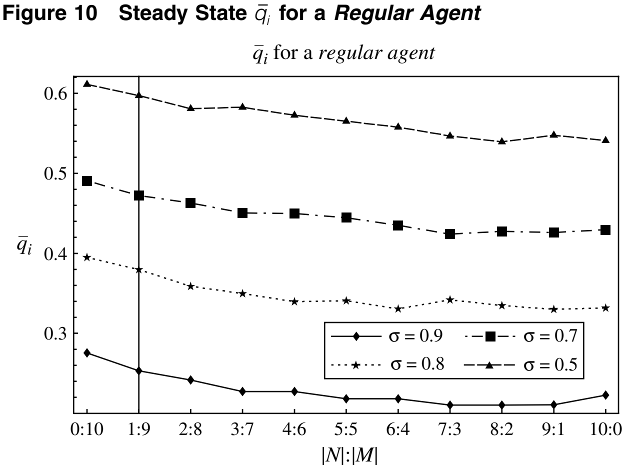


imitating other *imitators* is likely to generate repetitive ideas. Thus, *imitators* look to *innovators* to get fresh ideas, and *regular agents* look to *imitators* to get good ideas. These tendencies are stronger when the environment is more stable.

4.2. Impact of the Innovator/Imitator Mix

From Figure 5 we note that, for all values of σ, f_{MM} monotonically increases, while f_{MN} and f_{MR} monotonically decrease in the $|N|{:}|M|$ ratio. The intensity with which an *imitator* focuses his or her attention on an *innovator*, f_{MN}, tends to diminish as the number of *innovators* relative to *imitators* increases. The freed-up attention now goes to observing other *imitators* who have the favorably biased sample of ideas from an increasing number of *innovators*: Imitating *imitators* now becomes increasingly attractive because the risk of picking up redundant ideas goes down, as there are more *innovators* and less *imitators*. Consequently, $f_{MN} - f_{MM}$, the extent to which an *imitator* imitates an *innovator* rather than another *imitator*, monotonically declines when there exists a greater proportion of *innovators* in the population.

Figure 10 Steady State $\bar{q}_i$ for a *Regular Agent*



Similar observations can be made from Figure 6, which captures the probabilities of a *regular agent*: f_{RM} rises in the $|N|{:}|M|$ ratio, while f_{RN} and f_{RR} are nonmonotonic (U shaped) in the ratio. We observe in Figure 6(d) that $f_{RM} - f_{RN}$, the extent to which a *regular agent* observes an *imitator* rather than an *innovator*, monotonically increases in the $|N|{:}|M|$ ratio. The intuition is same as before.

Figures 7 and 8 provide further support for these results for various values of ρ. In the context of the chain network structure in which *imitators* learn from *innovators* and *regular agents* learn from *imitators*, it is then clear that the importance of a super-type agent (an *innovator* or an *imitator*) is positively related to its relative scarcity in the system. When there is a decline in the relative availability of *innovators* in the social system, an average *imitator* observes an average *innovator* with a higher probability. This is because the ideas held by the *imitators* tend to be redundant and the return to sampling an idea of an *innovator* is comparatively higher. Conversely, when there is a decline in the relative availability of *imitators* in the social system, an average *regular agent* observes an *imitator* with a higher probability, because the risk of getting redundant ideas is less.

4.3. Robustness

All of the results presented here for the baseline case have also been replicated for alternative cases: (1) $(\mu_i^{in}, \mu_i^{im}) = (0.75, 0.25)$ for all i in group N, $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.75)$ for all i in group M, and $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.25)$ for all i in group R, and (2) $(\mu_i^{in}, \mu_i^{im}) = (0.75, 0.25)$ for all i in group N, $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.75)$ for all i in group M, and $(\mu_i^{in}, \mu_i^{im}) = (0.5, 0.5)$ for all i in group R.

5. Results on the Network Architecture When *Innovators* Can Imitate and *Imitators* Can Innovate

We will now diverge from our baseline model and consider agent types that are more balanced. Let us endow *innovators* in group N with some ability to imitate and *imitators* in group M with some ability to innovate: $(\mu_i^{in}, \mu_i^{im}) = (0.75, 0.25)$ for all i in group N, $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.75)$ for all i in group M, and $(\mu_i^{in}, \mu_i^{im}) = (0.25, 0.25)$ for all i in group R. Galileo undoubtedly had some ability to network with his contemporaries and Mersenne surely had some ability to innovate and make discoveries on his own.[21] How is the architecture of problem-solving networks affected by the availability of alternative learning mechanisms for the super-type individuals? As mentioned in §4.3, all the results presented for the baseline case hold with this extension. However, there is an additional result with implications for the architecture of the network. Unlike our earlier case, *innovators* and *imitators* now tend to communicate directly with each other. While *imitators* continue to connect to *innovators* to imitate their ideas, *innovators* prefer to connect to *imitators* rather than other *innovators.* Not only do *imitators* serve their usual purpose, but they are now also sought after by *innovators*, who find that connecting with them is more productive than connecting with fellow *innovators*.

PROPERTY 3. *When all agents can both innovate and imitate* (*though to varying degrees*), *we observe the following*:

(1) *When the environment is relatively stable* (σ *high*), innovators *imitate* imitators *more than they imitate* innovators*. There is no clear difference when the environment is volatile* (σ *low*).

(2) Imitators *imitate* innovators *more than they imitate other* imitators.

(3) Regular agents *imitate* imitators *more than they imitate* innovators.

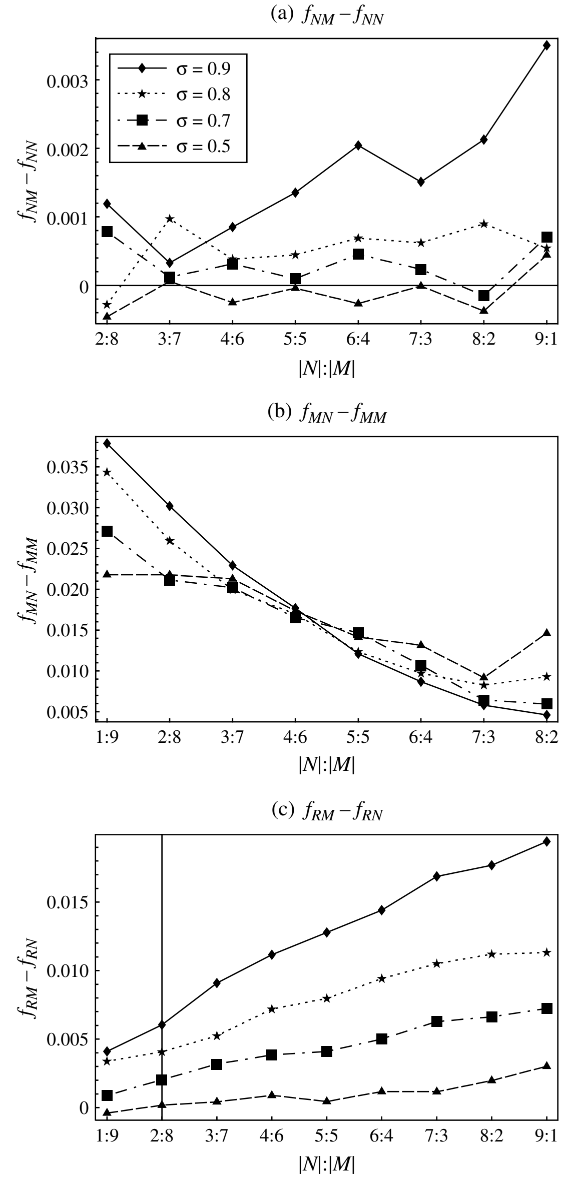
The above property is demonstrated in Figure 11, which plots the differential probabilities of $f_{NM} - f_{NN}$, $f_{MN} - f_{MM}$, and $f_{RM} - f_{RN}$ as functions of $|N|{:}|M|$ ratio for $\sigma \in \{0.9, 0.8, 0.7, 0.5\}$, given $\rho = 1$. Both $f_{NM} - f_{NN}$ and $f_{RM} - f_{RN}$ are positive for most values of $|N|{:}|M|$ and σ, but tend to decline with σ.[22] Focusing on when σ is high, we need to explain why *regular agents* and *innovators* choose to imitate *imitators*, while *imitators* choose to imitate *innovators*. The argument stated previously in the baseline case also explains why *regular agents* imitate *imitators* and *imitators* imitate *innovators*. The reason that *innovators* imitate *imitators* is the same reason that *regular agents* do—*imitators* have better practices (due to a favorably biased sample) and, because *innovators* (like *regular agents*) do not engage in as much imitation, there is less of a concern of rediscovering the same ideas.

When σ is low so that the environment is volatile, we need to explain the difference between *innovators* and *regular agents*—*regular agents* continue to imitate *imitators* more than *innovators*, while *innovators* do not seem to distinguish between *imitators* and *innovators*; note that Figure 11(a) shows $f_{NM} - f_{NN} \cong 0$ for σ low. One possible explanation is that the networks of *innovators* are less well formed than those of *regular agents*. Because *innovators* engage in more innovation than *regular agents*, this necessarily means less imitation, and it is the act of imitation that leads to a well-formed network.

6. Socially Optimal Mix of *Innovators* and *Imitators*

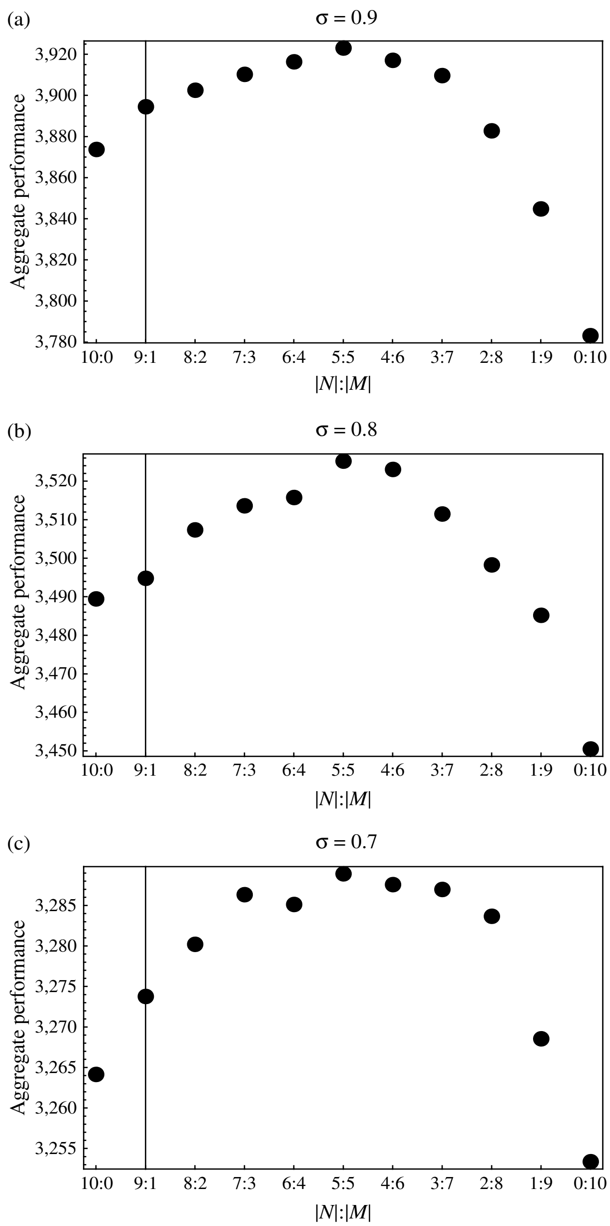
Having established the central role that *imitators* play in the evolving networks, we now explore the *socially optimal* mix of super-type agents.[23] Given that the innovators are the ones generating new ideas and thus providing raw materials for progress, is the social system best off with the super types consisting solely of *innovators*, or is it better off with some heterogeneous mixture of *innovators* and *imitators*? Furthermore, how is this optimal mix affected by the relevant environmental parameters, if at all?

Figure 11 $f_{NM} - f_{NN}$; $f_{MN} - f_{MM}$; $f_{RM} - f_{RN}$ [$\rho = 1$]



For the baseline parameter configurations, Figure 12 captures the steady-state aggregate performance, $\overline{\overline{\pi}}$, as a function of the $|N|{:}|M|$ ratio for $\sigma \in \{0.9, 0.8, 0.7\}$ and $\rho = 1$. It is clear from the figure that the aggregate performance is nonmonotonic in the mix, with $|N|{:}|M| =$ 5:5 emerging as the social optimum. The nonmonotonicity of $\overline{\overline{\pi}}$ is robust in that we observe the same property for all $\sigma \in \{0.5, 0.7, 0.8, 0.9\}$ and $\rho \in \{1, 4, 9\}$. Although the optimal mix of 5:5 is invariant to modest changes in parameters, we have found that large changes in parameter values can cause it to range from 7:3 to 4:6. What is always the case, however, is that the optimum has a mix of *innovators* and *imitators*. The configurations of super types consisting solely of *innovators* (10:0) or *imitators* (0:10) produces strictly inferior aggregate performance.

Figure 12 Aggregate Performance for $\rho = 1$

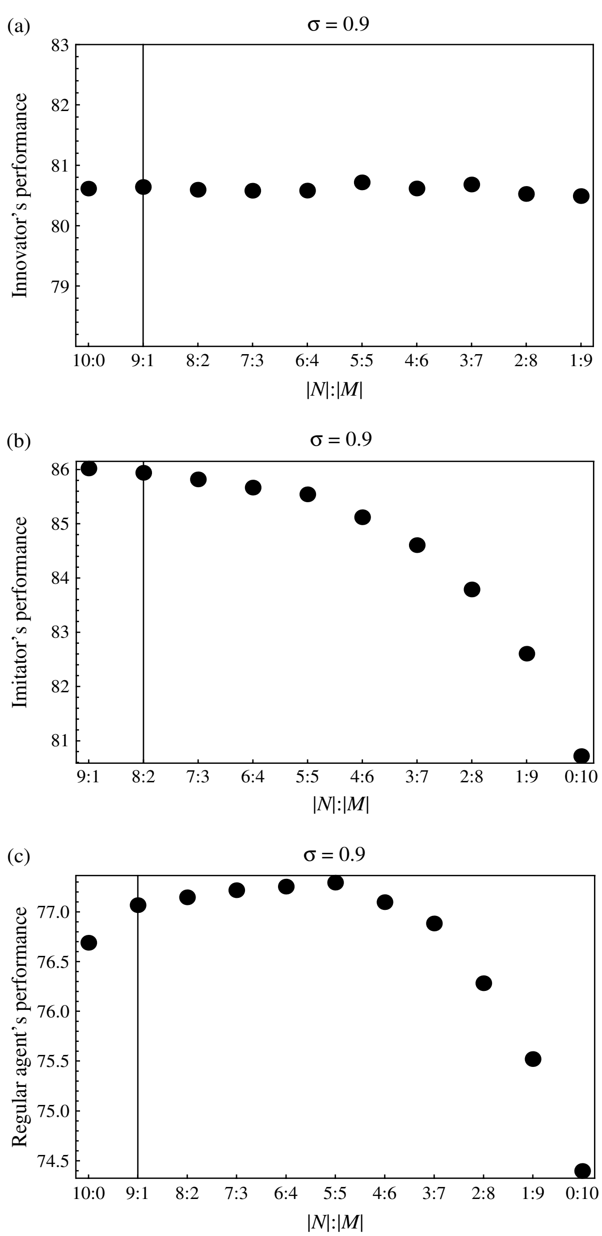


Property 4. *The aggregate performance is maximized when there is a mix of* innovators *and* imitators *in the population.*

Why is it that a mix of *innovators* and *imitators* is socially beneficial? What causes the marginal social gain from an additional *imitator* to outweigh (be outweighed by) the marginal social loss from one less *innovator* when the ratio of *innovators* to *imitators* is relatively high (low)? We conjecture that it is due to the fact that the *imitators*, in the course of imitating others, unintentionally play the role of integrating the distributed knowledge in the social system. Note that the baseline case involves *innovators*, who are only capable of generating new ideas. Imitation is done by *imitators* and *regular agents*. While the new ideas are generated by *innovators*, these original ideas are scattered among them. An average agent must observe a relatively large number of *innovators* to find a set of valuable ideas. However, given their superior abilities to imitate others, *imitators* easily observe and copy many of the valuable ideas that are distributed amongst the *innovators*. When $|N| \gg |M|$, an average *imitator* comes to be in possession of a wide variety of ideas originating from a relatively large number of *innovators*, thereby making it more productive for an average agent to observe a single *imitator* than to observe the original sources. Hence, the *imitator* facilitates efficient dissemination of valuable ideas within the social system. The determining factor for the social optimum is then the balancing of two forces: generation of new ideas by *innovators* and dissemination of existing ideas by *imitators*. When $|N| \gg |M|$, the relative value of dissemination is important, because the marginal social gain from an additional *imitator* outweighs the marginal social cost of one less *innovator*, and social performance improves. Once the proportion of *imitators* exceeds a threshold level, there is now insufficient generation of ideas within the social system; the marginal social value of an *innovator* outweighs that of an *imitator*, leading to a decline in the aggregate performance. The socially optimal mix is then strictly interior.

We take a step toward confirming the conjectured role of *imitators* by investigating how the performance at the *individual* level for each type is affected by the mixture of the super types. To that end, we examine the average performance of an individual agent belonging to each type. Figure 13 captures the individual-level performance for each type as a function of the $|N|:|M|$ ratio, given $\sigma = 0.9$ and $\rho = 1$. Figure 13(a) shows that $\overline{\pi}^N$ tends to be independent of the mix of super types: *Innovators* rely solely on their own generation of ideas. However, Figures 13(b) and 13(c) show, respectively, that $\overline{\pi}^M$ and $\overline{\pi}^R$ depend on the mix: The performance of an average *imitator*, $\overline{\pi}^M$, is monotonically decreasing in the relative proportion of the *imitators* to the *innovators*, while the performance of an average *regular agent*, $\overline{\pi}^R$, is nonmonotonic in the ratio. Intuitively, an average *imitator* loses when the proportion of *innovators* in the system decreases, because the sources of new ideas are drying up (and, simultaneously, the amount of redundant ideas is increasing). This implies that the dominant drivers of the *imitators*' performance are the ideas directly supplied by the *innovators*. In contrast, an average *regular agent* initially benefits from having additional *imitators* in the system when there is a relatively small number of *imitators*. Once the proportion of *imitators* to *innovators* reaches a certain level, replacing more *innovators* with *imitators* tends to depress a *regular agent*'s performance. This is where there exist an insufficient number of *original* ideas and an excessive number of *redundant*

Figure 13 An Average Agent's Performance in Each Group
[$\sigma = 0.9$; $\rho = 1$]



ideas in the social system. *Regular agents*, who are the majority of the population, are then directly influenced by the presence and the prevalence of *imitators*.

Finally, it is a notable property of Figure 13 that *imitators* outperform not just *regular agents*, but also *innovators*.[24] Indeed, the differential between an *imitator* and an *innovator* is about the same size as that between an *innovator* and a *regular agent*. The superior performance of an *imitator* is consistent with the empirical finding of Sparrowe et al. (2001), that people with higher centrality have higher performance.

7. Concluding Remarks

This research was motivated by the finding in the organization science literature that interpersonal networks and the presence of connectors within such networks are instrumental to attaining superior organizational or social performance. In this paper, we proposed to take a step back from this accepted wisdom and ask a more fundamental question in the context of decentralized problem solving: Do connectors endogenously emerge in an organization or social system when a population of autonomous agents can individually choose to innovate (engage in solitary learning) or imitate (engage in social learning by establishing and developing network links to other agents)? Given that an individual's choice depends on his or her skills with regard to innovation and imitation, as well as the volatility of the environment, which impacts the relative returns to those two activities, we further posed the question of how the distribution of skills among agents and the stability of the environment determine the architecture of the network as well as the relative importance of connectors within them.

When the population contains superinnovative and superimitative individuals who are fully specialized (the baseline case), we find that the architecture of problem-solving networks evolves into a chain: *Innovators* generate ideas, *imitators* learn from *innovators*, and *regular agents* learn from *imitators*. The overall flow of knowledge then entails *imitators* acting as connectors between *innovators* and *regular agents*. When the superinnovative individuals have the capacity to imitate and the superimitative individuals have the capacity to innovate, we find strong mutual interactions between *innovators* and *imitators*. *Imitators* learn from *innovators*, but *innovators* (as wells as *regular agents*) learn from *imitators*. In both cases, the importance of *imitators* within the network comes from their ability to integrate dispersed knowledge in the social system. Their role as a repository of knowledge improves the efficiency of search by individual agents, thereby leading to their centrality in the emergent network.

One of the main parameters considered in this paper was the composition of the super-type group. Their importance in the social network was found to be directly affected by their relative scarcity in the population. The network architecture was also shown to be affected by the volatility in the task environment. A more stable task environment is more likely to lead to the emergence of a chain network structure with *imitators* as connectors. Finally, social performance was maximized when there was a heterogeneous mixture of *innovators* and *imitators*. This result directly confirms the complementary relationship between *innovators* and *imitators*—*innovators* as the generators of new ideas and *imitators* as the disseminators of those ideas.

Although the model is parsimonious and generic in design, the ensuing theory does offer some useful insights for organizations. First, prospective employees should be evaluated not only on their ability to come up with novel solutions (as with our *innovators*) but also on their ability to understand, distill, and communicate the solutions of others (as with our *imitators*). A high-performing network requires both agents who create and agents who communicate. Second, given the role for both of those attributes, it can be more disruptive to lose an employee who is a connector than to lose one who is an innovator. With the loss of any employee, the remaining agents must adapt their links in the network because one source of those links has disappeared. Given the chain structure of the network—imitators connect to innovators and most other agents connect to imitators—the loss of a connector requires the vast population of agents to reestablish their links, and it can be a slow process to form the new network. In contrast, the loss of an innovator means that connectors must reestablish their links and—this is the good news—connectors are effective at doing that, because they are effective in communicating with other agents. While no organization can expect to make long-run improvements without new ideas, our theory highlights the importance of retaining those who disseminate ideas rather than those who create them, when the priority is in avoiding short-run deterioration of performance. Finally, the importance of retaining connectors is greater when the organization faces a more stable environment. In the context of business, firms facing competition from rival firms in the market, a relatively steady market with few firm turnovers is likely to provide a stable environment for the agents in the firms. As shown in our study, this gives rise to the emergence of a more chainlike network structure within a firm. Because those individuals with superior skills in disseminating ideas are particularly important in relatively stable environments, firms need to pay special attention to retaining them more than those who are superior in generating ideas.

Acknowledgments
The authors thank Linda Argote, Rich Burton, and the three referees for their detailed comments. The paper has also been improved by the comments of Jason Barr, Jon Harford, Bill Kosteas, and the conference participants at the NAACSOS 2005 Conference (Notre Dame), the Social Network Analysis: Advances and Empirical Applications Forum (University of Oxford, UK, July 16–17, 2005), and the 2005 SEA Conference (Washington, D.C.). Hannah Rozen provided her capable research assistance throughout the project.

Endnotes
[1]"Our study offers evidence of at least three enduring relational characteristics that are predictive of the behavior of information seeking: (1) knowing what another person knows, (2) valuing what that other person knows in relation to one's work, and (3) being able to gain timely access to that person's thinking."

[2]In the context of organizational learning, Siggelkow and Rivkin (2005) utilize an agent-based modelling approach that is similar to ours. Their focus, however, is on the issue of *formal* organizational design in the presence of environmental turbulence and complexity.

[3]Our network is most closely related to what some scholars call an "advice network," though it is richer in that it allows for knowledge creation, not just knowledge sharing.

[4]The problem for an agent may be viewed as that of search over a landscape defined on an $(H \cdot d + 1)$ dimensional Euclidean space, with the final dimension indicating the performance of the agent. Our model then implies a nonlocal search on a single-peaked smooth landscape, because each experiment (a blind step on the landscape) via innovation or imitation entails changing up to d bits in one of the H tasks, and the performance depends only on the Hamming distance between the agent's methods vector and the goal vector. If one were to assume complex interactions among various tasks so that the performance contribution of a method chosen for a task depends on the methods used in other tasks, then the search landscape will be rugged with multiple local optima. (See Chang and Harrington 2006 for a survey of models using this approach.) Given the many diverse features of our model (including the environmental turbulence), how the results obtained in this paper will be affected by the degree of complexity (i.e., the ruggedness of the landscape) is an open question.

[5]There is a restriction in that an agent only has the option of adopting all d dimensions or none.

[6]Implicit in the adoption rules used in our model is the assumption about what agents know about their environment. We do not believe that they have complete information about the performance level associated with the proposed methods vector, but rather that they can experiment to get a reasonable estimate of it. Instead of modelling this process of experimentation, which would further complicate an already complex model, we implicitly assume that it is done instantly and without cost.

[7]For analytical simplicity, we assume that ϕ and λ have the same value for all agents in the population.

[8]For simplicity, we assume that the sensitivity parameter, λ, is the same for both $q_i(t)$ and $p_i^j(t)$.

[9]It should be noted that the network and the imitation behavior of an agent are mutually interdependent in this setting. A network in our model is defined by the probabilities with which an agent observes other agents in the population. To the extent that agent i has probability $p_i^j(t)$ of observing another agent j in period t, the network determines the imitating behavior of agent i in that period. However, another main feature of our model is that these imitation probabilities are updated on the basis of how successful agent i's attempt to imitate j $(j \neq i)$ was, so the imitating behavior of an agent also affects his or her network.

[10]Selecting the parameter values for our computational experiments is driven by the constraints imposed by the limited computational resources. $L = 50$ was the largest population size that we could realistically work with, given that each of these agents maintains a set of observation probabilities for 49 other agents, which are continually updated from one period to the next over a horizon of 15,000 periods. It should be noted, however, that we performed additional simulations for $L \in \{20, 30, 40\}$ and confirmed that the qualitative results reported in this paper for $L = 50$ have all been replicated for those values of L.

[11]As a robustness check, we also performed simulations assuming initial practices that are *heterogeneous*. All the qualitative results reported in the paper continue to hold even with this modification.

[12]By a *steady state*, we mean the state in which the *mean* value of the variable—that is, mean across multiple replications—is independent of time. This is to be contrasted with *transient* periods, in which the mean value of the variable changes over time (presumably on its way to converge on some steady state).

[13]A replication is the running of the model for 15,000 periods given a set of random numbers. For each parameter configuration considered in this paper, the model is then run for a total of 1.5 million periods (15,000 periods per replication × 100 independent replications).

[14]Note that the agents initially start out with uniform attraction stocks and, hence uniform probabilities of observing other agents such that $p_i^j(0) = p_i^k(0)\ \forall\, j, k \neq i$. When $\mu_i^{im} = 0$ (as is the case for *innovators* in the benchmark case), these probabilities are never adjusted over time.

[15]While the outputs reported in Figure 2 are the averages over 100 replications, the outputs from individual replications also display the same general pattern.

[16]This is still assuming that the environment is sufficiently dynamic. In a *completely* static environment, learning ceases to exist altogether once the goal is attained.

[17]We considered the population sizes of $L \in \{20, 30, 40, 50\}$ while holding fixed the total size of the super types at 10—i.e., $|N| + |M| = 10$.

[18]f_{RN} is nonmonotonic in σ for a wide range of $|N|{:}|M|$ values (see Figure 6(a)), even though it monotonically declines in ρ (see Figure 8(a)). Thus, we refrain from making a conclusive remark on the impact of volatility on f_{RN}.

[19]As described in §2, an idea is a sequence of d bits representing a method for a particular task.

[20]As shown in Figure 6(a), even though the probability with which a *regular agent* imitates an *innovator* rises when σ goes from 0.5 to 0.7, its impact is ambiguous when σ goes from 0.7 to 0.8 or from 0.8 to 0.9.

[21]This greatly understates the ability of Galileo as well as that of Mersenne. In fact, Galileo's success as a discoverer owes much to his extensive use of telescopes, which he initially learned of through his well-developed network connections. Similarly, Mersenne was a well-regarded scientist in his time, having certain discoveries to his name—for example, Mersenne prime numbers.

[22]In Figure 11(a), it appears that $f_{NM} - f_{NN}$ becomes negative for some $|N|{:}|M|$ ratios when $\sigma = 0.5$.

[23]In the previous sections, we focused on the positive question of *what actually happens to the network architecture*. In this section, we ask the normative question of *what the network architecture should be*. See Burton (2003) for an insightful discussion of the positive versus normative use of the computational modelling and analysis.

[24]This property was also displayed in Figure 9 for a wider variety of values for σ.

References

Borgatti, S. P., R. Cross. 2003. A relational view of information seeking and learning in social networks. *Management Sci.* **49**(4) 432–445.

Burton, R. M. 2003. Computational laboratories for organization science: Questions, validity and docking. *Comput. Math. Organ. Theory* **9**(2) 91–108.

Camerer, C., T.-H. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* **67**(4) 827–874.

Chang, M.-H., J. E. Harrington, Jr. 2005. Discovery and diffusion of knowledge in an endogenous social network. *Amer. J. Sociol.* **110**(4) 937–976.

Chang, M.-H., J. E. Harrington, Jr. 2006. Agent-based models of organizations. L. Tesfatsion, K. L. Judd, eds. *Handbook of Computational Economics II: Agent-Based Computational Economics*. North Holland, Amsterdam.

Cross, R., A. Parker. 2004. *The Hidden Power of Social Networks*. Harvard Business School Press, Boston, MA.

Cross, R., A. Parker, L. Sasson, eds. 2003. *Networks in the Knowledge Economy*. Oxford University Press, New York.

Cross, R., A. Parker, L. Prusak, S. P. Borgatti. 2001. Knowing what we know: Supporting knowledge creation and sharing in social networks. *Organ. Dynam.* **30**(2) 100–120.

Goyal, S. 2005. Learning in networks. G. Demange, M. Wooders, eds. *Group Formation in Economics: Networks, Clubs, and Coalitions*. Cambridge University Press, Cambridge, UK.

Haas, M. R. 2006. Acquiring and applying knowledge in transnational teams: The roles of cosmopolitans and locals. *Organ. Sci.* **17**(3) 367–384.

Hatch, R. A. 1998. Between erudition and science: The archive and correspondence network of Ismaël Boulliau. M. Hunter, ed. *Archives of the Scientific Revolution: The Formation and Exchange of Ideas in Seventeenth-Century Europe*. The Boydell Press, Woodbridge, NY.

Hull, D. L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. The University of Chicago Press, Chicago, IL.

Hunter, M. 1998. *Archives of the Scientific Revolution: The Formation and Exchange of Ideas in Seventeenth-Century Europe*. The Boydell Press, Woodbridge, NY.

Jackson, M. O. 2006. The economics of social networks, Ch. 1. R. Blundell, W. Newey, T. Persson, eds. *Advances of Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society,* Vol. I. Cambridge University Press, Cambridge, UK.

Klein, K. J., B.-C. Lim, J. L. Saltz, D. M. Mayer. 2004. How do they get there? An examination of the antecedents of centrality in team networks. *Acad. Management J.* **47**(6) 952–963.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organ. Sci.* **2**(1) 71–87.

Podolny, J. M., K. L. Page. 1998. Network forms of organizations. *Annual Rev. Sociol.* **24** 57–76.

Rogers, E. M. 2003. *Diffusion of Innovations*, 4th ed. Free Press, New York.

Siggelkow, N., D. A. Levinthal. 2003. Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. *Organ. Sci.* **14**(6) 650–669.

Siggelkow, N., J. W. Rivkin. 2005. Speed and search: Designing organizations for turbulence and complexity. *Organ. Sci.* **16**(2) 101–122.

Sparrowe, R. T., R. C. Liden, S. J. Wayne, M. L. Kraimer. 2001. Social networks and the performance of individuals and groups. *Acad. Management J.* **44**(2) 316–325.

Uzzi, B., S. Dunlap. 2005. How to build your network. *Harvard Bus. Rev.* **83** 53–60.