Innovators, Imitators, and the Evolving Architecture of Social Networks*

Myong-Hun Chang
Department of Economics
Cleveland State University
Cleveland, OH 44115
216-687-4523, -9206 (Fax)
m.chang@csuohio.edu
academic.csuohio.edu/changm

Joseph E. Harrington, Jr.
Department of Economics
Johns Hopkins University
Baltimore, MD 21218
410-516-7615, -7600 (Fax)
joe.harrington@jhu.edu
www.econ.jhu/people/harrington

September 2, 2005

Abstract

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 by those with superior abilities to innovate, the observation and adoption of these ideas by those with superior abilities to communicate and imitate, and the endogenous development of social networks among heterogeneous agents. The objective is to explore the evolving architecture of social networks and the critical roles that the innovators and imitators play in the process. A central finding is that the emergent social network takes a chain-structure with the innovators as the main source of ideas and the imitators as the connectors between the innovators and the masses. The impact of agent heterogeneity and environmental volatility on the network architecture is also characterized.

Running Head: Evolving Architecture of Social Networks

Keywords: Innovator, Imitator, Connector, Social Networks, Network Architecture, Agent-Based Model

*The comments of Jon Harford, Bill Kosteas, and the conference participants at the NAACSOS 2005 Conference (Notre Dame) and the *Social Network Analysis: Advances and Empirical Applications Forum* (University of Oxford, UK, July 16-17, 2005) are gratefully acknowledged. We also thank Hannah Rozen for her capable research assistance throughout the project.

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 – e.g., 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), for instance, describes the extensive correspondence networks that were established and operated by a few human connectors during this period:

> [I]t was not without reason that the traditional heros of the Scientific Revolution – Copernicus, Kepler, Galileo, Descartes, Huygens, Newton – were so honoured and thus honoured first. Even a self-blinded scholar ... could not go untouched by such genius. There was, of course, more to the story, a broader context that cut across generation, class, temperament and traditional periodisations. Here we are reminded of the intelligence and industry of lesser lights, of a Mersenne, Hartlib, or Oldenburg. [Hatch (1998: 50-55)]

These social connectors, identified and studied by Hatch, include 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: 55)]. In fact, luminaries such as Galileo, Huygens, Dupuy, Mersenne, Oldenburg, and Fermat were all connected to Boulliau's network.[1]

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, and if they occur in close proximity, discovery is locally more important than dissemination. However, the more distant in space and time an undisseminated discovery is, the less important it is. As Lamarck (1809: 404) ruefully concluded his *Philosophie zoologique*,

[1] Boulliau was also a collector of manuscripts and a copyist of other peoples' letters. The Boulliau Archive contains "copies of letters of such contemporary and historical figures as Tycho Brahe" as well as those of Gallileo and Huygens, among others. [Hatch (1998)]

> "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: 376-377)]

It is clear that scientific progress requires individuals who are capable of generating new ideas. The real question is whether the connectors with their superior communication and networking abilities are important in this process and, if so, what it implies in terms of the emergent architecture of the social networks. More specifically, given the endowed skill differentials scattered among the individuals in the population, how does such heterogeneity feed into the private choices they make in allocating their efforts between individual learning (innovation) and social learning (imitation through social interactions)? Into what kinds of architecture do the social networks of these heterogeneous individuals ultimately evolve, and what are the social consequences of the interactions among such adaptive processes at the individual level? Does the social accumulation of knowledge always require social connectors as conduits between "genius" and "masses" or are there circumstances under which such connectors may be bypassed?

In order to address these issues, we develop a computational model of the process by which ideas are generated and diffused through evolving social networks. Our model entails a population of myopic, though adaptive, agents searching for a common optimum in the space of possible things that one can do. The agents choose whether to allocate their efforts to discovering new ideas – innovation – or to observing and copying the ideas of others – imitation. When they engage in imitation, agents decide from whom to learn, which takes the form of establishing links in a social network. These choices are made probabilistically and the probabilities are adjusted over time via reinforcement learning. This modeling structure allows us to examine the evolving architecture of the social network in terms of how observation probabilities are distributed across individuals.[2] The knowledge creation/diffusion process occurs in the context of a changing environment as represented by stochastic movement in the common optimum.

The success of an individual's innovation or imitation efforts depends on whether his inherent ability lies in generating new ideas or in imitating others via establishing communication links with agents in the population. We assume that the said agents

[2]Skyrms and Pemantle (2000) use a similar perspective in modeling repeated games between pairs of players who are chosen stochastically from a fixed population. Continually updating the probabilities of a player meeting other players through reinforcement learning, they focus on the limit probability matrix to which these probabilities converge as $t \to \infty$.

are heterogeneous in these capabilities.[3] We divide the population into three separate groups, based on their abilities to innovate and imitate. The first group, called *Innovators,* is composed of individuals who are super-innovative, but have modest ability to imitate. The second group, called *Imitators,* consists of those who are super-imitative, but have modest ability to innovate. The third group contains the rest of the population who are considered to be ordinary in their abilities to innovate and imitate. We refer to them as *Regular Agents.*

The goal of this paper is to explore the architecture of the social networks which develop over time as individuals in these three groups interact with one another through endogenous innovation and imitation activities. Once the structural properties of the social networks are identified, we examine how the distribution of the heterogeneous agents in the population as well as the parameters controlling the volatility of the environment affect the evolution of such network architecture. These positive analyses then lead the way to a normative analysis, in which we evaluate the *social* importance of *Innovators* versus *Imitators* in a concrete manner. Specifically, we ask "what is the socially optimal mix of these super-type agents?" Given that *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 Society better off by having some heterogeneous mixture of *Innovators* and *Imitators*? If we answer the latter question in the affirmative, what are the relevant environmental parameters that may affect this optimal mix?

2 The Model

2.1 Agents, Tasks, Goal and Performance

The social system consists of L individuals. Each individual $i \in \{1, 2, \ldots, L\}$ engages in an operation which can be broken down into H separate tasks. There are several different methods which can be used to perform each task. The method chosen by an agent for a given task is represented by a sequence of d bits (0 or 1) such that there are 2^d possible methods available for each task. In any period t, an individual i is then fully characterized by a binary vector of $H \cdot d$ dimensions. Denote it by $\underline{z}_i(t) \in \{0,1\}^{Hd}$ so that $\underline{z}_i(t) \equiv (\underline{z}_i^1(t), ..., \underline{z}_i^H(t))$ and $\underline{z}_i^h(t) \equiv (z_i^{h,1}(t), ..., z_i^{h,d}(t)) \in \{0,1\}^d$ is individual i's chosen method in task $h \in \{1, ..., H\}$.

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:

$$D(\underline{z}_i, \underline{z}_j) \equiv \sum_{h=1}^{H} \sum_{k=1}^{d} \left| z_i^{h,k} - z_j^{h,k} \right|. \tag{1}$$

In period t, the population faces a common goal vector, $\widehat{\underline{z}}(t) \in \{0,1\}^{Hd}$. The

[3] See Chang and Harrington (2005) for a version of this model in which the agents are homogeneous in their capabilities.

degree of turbulence in task environments is captured by intertemporal variability in $\widehat{\underline{z}}(t)$, the details of which are to be explained in 2.4.

The individuals are uninformed about $\widehat{\underline{z}}(t)$ *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 $\widehat{\underline{z}}(t)$, the period-$t$ performance of individual i is then measured by $\pi_i(t)$, where

$$\pi_i(t) = H \cdot d - D(\underline{z}_i(t), \widehat{\underline{z}}(t)). \tag{2}$$

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,

$$\widehat{\pi}(t) = \sum_{i=1}^{L} \pi_i(t). \tag{3}$$

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.[4] 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 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

[4]See Kitcher (1993: 60) for a similar view which treats innovation and imitation at the individual level as the two driving forces behind scientific change:

> Consensus practice changes in response to modifications of individual practices; individual practices alter as a result of changes in individuals' cognitive states. What drives these latter changes? ... Sometimes scientists modify their cognitive states as results of *asocial* interactions, sometimes they change their minds through *social* exchanges. The obvious exemplars for the former are the solitary experimentalist at work with apparatus and samples and the lone field observer attending to the organisms – although I shall also take encounters with nature to cover those occasions on which scientists reflect, constructing chains of reasoning that modify their commitments. Paradigm cases of conversations with peers are those episodes in which one scientist is told something by another (and believe it) or when a change in commitment is caused by the reading of a text. The point of the distinction is evidently to separate those episodes that (very roughly) consist in finding things out for oneself from those in which one relies on others.

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

decreasing the Hamming distance between the agent's new methods vector and the goal 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 chooses to imitate, how does he decide whom to imitate? We model this as a two-stage stochastic decision process with reinforcement learning.[6] 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 he chooses to innovate then, with probability μ_i^{in}, he generates an idea which is a randomly chosen task $h \in \{1, \ldots, H\}$ and a randomly chosen method, $\underline{z}_i^{h\prime}$, for that task such that the experimental method vector is $\underline{z}_i'(t) \equiv (\underline{z}_i^1(t), \ldots, \underline{z}_i^{h-1}, \underline{z}_i^{h\prime}, \underline{z}_i^{h+1}, \ldots, \underline{z}_i^H(t))$. μ_i^{in} is a parameter that controls the productivity of an agent's innovation. 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, $\widehat{\underline{z}}(t)$. Otherwise, it is discarded:

$$\underline{z}_i(t+1) = \begin{cases} \underline{z}_i'(t), & \text{if } D(\underline{z}_i'(t), \widehat{\underline{z}}(t)) < D(\underline{z}_i(t), \widehat{\underline{z}}(t)), \\ \underline{z}_i(t), & \text{if } D(\underline{z}_i'(t), \widehat{\underline{z}}(t)) \geq D(\underline{z}_i(t), \widehat{\underline{z}}(t)). \end{cases} \quad (4)$$

Alternatively, with probability $1 - \mu_i^{in}$ the individual fails to generate an idea, in which case $\underline{z}_i(t+1) = \underline{z}_i(t)$.

Now suppose individual i chooses to imitate in stage 1. Given that he decides to imitate someone else, he 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. Hence, μ_i^{im} measures the ability of the agent to communicate with others in the population. An agent that is connected then enters stage 2 of the decision process in which he 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 $\underline{z}_i''(t)$ $= (\underline{z}_i^1(t), \ldots, \underline{z}_i^{h-1}(t), \underline{z}_l^h(t), \underline{z}_i^{h+1}(t), \ldots, \underline{z}_i^H(t))$ be the experimental vector. Adoption or rejection of the observed method is based on the Hamming distance criterion:

$$\underline{z}_i(t+1) = \begin{cases} \underline{z}_i''(t), & \text{if } D(\underline{z}_i''(t), \widehat{\underline{z}}(t)) < D(\underline{z}_i(t), \widehat{\underline{z}}(t)), \\ \underline{z}_i(t), & \text{if } D(\underline{z}_i''(t), \widehat{\underline{z}}_i(t)) \geq D(\underline{z}(t), \widehat{\underline{z}}(t)). \end{cases} \quad (5)$$

If the agent fails to connect to the network, which occurs with probability $1 - \mu_i^{im}$, $\underline{z}_i(t+1) = \underline{z}_i(t)$.

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

[6] The description of the model in this section is identical to the one in Chang and Harrington (2005).

a version of the *Experience-Weighted Attraction (EWA)* learning rule as described in Camerer and Ho (1999). Using this 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 evolution of $A_i^{in}(t)$ and $A_i^{im}(t)$ follow the process below:

$$A_i^{in}(t+1) = \begin{cases} \phi A_i^{in}(t) + 1, & \text{if } i \text{ adopted a method through innovation in } t \\ \phi A_i^{in}(t), & \text{otherwise.} \end{cases} \tag{6}$$

$$A_i^{im}(t+1) = \begin{cases} \phi A_i^{im}(t) + 1, & \text{if } i \text{ adopted a method through imitation in } t \\ \phi A_i^{im}(t), & \text{otherwise.} \end{cases} \tag{7}$$

where $\phi \in (0,1]$. Hence, if the agent chose to pursue *Innovation* and discovered and then adopted his new idea, the attraction measure for *Innovation* increases by 1 after allowing for the decay factor of ϕ on the previous attraction level. If the agent chose to innovate but was unsuccessful (either because he failed to generate an idea, or because the idea he generated was not useful) or if he instead chose to imitate, then his attraction measure for innovation is simply the attraction level from the previous period decayed by the factor ϕ. Similarly, a success or failure in imitation at t has the identical influence on $A_i^{im}(t+1)$. Given $A_i^{in}(t)$ and $A_i^{im}(t)$, one derives the choice probability of innovation in period t as follows:

$$q_i(t) = \frac{\left(A_i^{in}(t)\right)^\lambda}{\left(A_i^{in}(t)\right)^\lambda + \left(A_i^{im}(t)\right)^\lambda} \tag{8}$$

where $\lambda > 0$. 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 (8) says that a favorable experience through innovation (imitation) raises the probability that an agent will choose to innovate (imitate) again in the future. In sum, a positive outcome realized from a course of action reinforces the likelihood of that same action being chosen again.

The stage-2 attractions and the probabilities are derived similarly. Let $B_i^j(t)$ be agent i's attraction to another agent j in period t. It evolves according to the rule below:

$$B_i^j(t+1) = \begin{cases} \phi B_i^j(t) + 1, & \text{if } i \text{ successfully imitated } j \text{ in } t \\ \phi B_i^j(t), & \text{otherwise.} \end{cases} \tag{9}$$

$\forall j \neq i$. The probability that agent i observes agent j in period t is adjusted each period on the basis of the attraction measures, $\{B_i^j(t)\}_{j \neq i}$:

$$p_i^j(t) = \frac{\left(B_i^j(t)\right)^\lambda}{\sum_{h \neq i} \left(B_i^h(t)\right)^\lambda} \tag{10}$$

$\forall j \neq i, \forall i$, where $\lambda > 0$.

[7] For analytical simplicity, we assume ϕ and λ to be common to all individuals in the population.

There are two distinct sets of probabilities in our model. One set of probabilities, $q_i(t)$ and $\{p_i^j(t)\}_{j \neq i}$, are endogenously derived and evolve over time in response to the personal experiences of agent i. Another set of probabilities, μ_i^{in} and μ_i^{im}, are exogenously specified and are 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

Central to the performance of a population is how it responds to an evolving environment or, if we cast this in the context of problem-solving, an evolving set of problems to be solved. It is such change that makes innovation and the spread of those innovations through a social network so essential. Change or turbulence is specified in our model by first assigning an initial goal vector, $\widehat{z}(0)$, to the population and then specifying a dynamic process by which it shifts over time.

Letting $s \in \{0,1\}^{Hd}$, define $\delta(s,\kappa) \subset \{0,1\}^{Hd}$ as the set of points that are exactly Hamming distance κ away from s. The set of points *within* Hamming distance κ of s is defined as

$$\Delta(s,\kappa) \equiv \bigcup_{i=0}^{\kappa} \delta(s,i) \tag{11}$$

$\Delta(s,\kappa)$ is a set whose "center" is s.

In period t, all agents in the population have the common goal vector of $\underline{\widehat{z}}(t)$. In period $t+1$, the goal stays the same with probability σ and changes with probability $(1-\sigma)$. The shift dynamic of the goal vector is guided by the following stochastic process. The goal in $t+1$, if different from $\underline{\widehat{z}}(t)$, is then an *iid* selection from the set of points that lie within the Hamming distance ρ of $\underline{\widehat{z}}(t)$. Defining $\Lambda(\underline{\widehat{z}}(t),\rho)$ as the set of points from which the goal in $t+1$ is chosen, we have

$$\Lambda(\underline{\widehat{z}}(t),\rho) \equiv \Delta(\underline{\widehat{z}}(t),\rho) \setminus \underline{\widehat{z}}(t). \tag{12}$$

Hence, $\Lambda(\underline{\widehat{z}}(t),\rho)$ includes all points in $\Delta(\underline{\widehat{z}}(t),\rho)$ except for $\underline{\widehat{z}}(t)$. Consequently,

$$\begin{cases} \underline{\widehat{z}}(t+1) = \underline{\widehat{z}}(t) & \text{with probability } \sigma \\ \underline{\widehat{z}}(t+1) \in \Lambda(\widehat{z}(t),\rho) & \text{with probability } 1-\sigma \end{cases} \tag{13}$$

The goal vector for the population then stochastically fluctuates while remaining within Hamming distance ρ of the current goal. This allows us to control the possible size of the inter-temporal change. The lower is σ and the greater is ρ, the more frequent and variable is the change, respectively, in the population's goal vector.

3 Design of Computational Experiments

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

We assume a population of fifty individuals: $L = 50$. The population is divided into three separate groups: *Innovators*, *Imitators*, and *Regular Agents*. Let N represent (and denote) the set of *Innovators* and M the group of *Imitators*. The group of *Regular Agents* is denoted as R. There are exactly ten 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 $(\mu_i^{in}, \mu_i^{im}) = (.25, .25)$ for all i in R. Later we will consider two extensions: 1) $(\mu_i^{in}, \mu_i^{im}) = (.75, .25)$ for all i in N, $(\mu_i^{in}, \mu_i^{im}) = (.25, .75)$ for all i in M, and $(\mu_i^{in}, \mu_i^{im}) = (.25, .25)$ for all i in R; and 2) $(\mu_i^{in}, \mu_i^{im}) = (.75, .25)$ for all i in N, $(\mu_i^{in}, \mu_i^{im}) = (.25, .75)$ for all i in M, and $(\mu_i^{in}, \mu_i^{im}) = (.5, .5)$ 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. 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 ten, 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 {.5, .7, .8, .9} and ρ from {1, 4, 9}.

Additional parameters are ϕ and λ, which control the evolution of the attraction measures. We assume that $\phi = 1$ and $\lambda = 1$. These values remain fixed over the relevant horizon. Finally, the initial attraction stocks are set at $B_i^j(0) = 1 \forall i, \forall j \neq i$, and $A_i^{in}(0) = A_i^{im}(0) = 1 \forall i$. Hence, an individual is initially equally attracted to innovation and imitation and has no inclination to observe one individual over another *ex ante*.

All computational experiments carried out here assume a horizon of 15,000 periods. The time-series of the performance measures are observed to reach a steady-state by the 2,000th period. 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 such that

$$\overline{\pi}_i = \frac{1}{5,000} \sum_{t=10,001}^{15,000} \pi_i(t). \tag{14}$$

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:

$$\overline{q}_i = \frac{1}{5,000} \sum_{t=10{,}001}^{15{,}000} q_i(t). \tag{15}$$

Finally, the endogenous steady-state imitation probabilities, denoted $\overline{p}_i^j$ are computed to be the average over the last 5,000 periods:

$$\overline{p}_i^j = \frac{1}{5,000} \sum_{t=10{,}001}^{15{,}000} p_i^j(t) \tag{16}$$

All of the experiments were based on 100 replications, each using a fresh set of random numbers.[8] Hence, the performance and probability measures reported in the paper are the averages over those 100 replications.

4 Evolving Architecture of Social 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 $(\mu_i^{in}, \mu_i^{im}) = (.25, .25)$ for all i in R. Hence, *Innovators* are true solitary geniuses who have no communication abilities. The *Imitators* are pure copycats with no ability to make independent discoveries. They rely exclusively on imitating someone else in the population through social networks. Finally, the *Regular Agents* have modest innate ability in both innovation and imitation.

In our model, the social network is defined in terms of the observation probabilities that the agents possess. As such, we must examine the steady-state probabilities, $\overline{p}_i^j$s, in order to analyze the evolving architecture of the network. Recall that $\overline{p}_i^j$ is the probability with which 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 $\sigma = .8$, $\rho = 1$ 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 through 4 are *Innovators* (group N), agents 5 through 10 are *Imitators* (group M), and agents 11 through 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. The six *Imitators* (agents 5 through

[8]Hence, the model is run for 1.5 million periods for each parameter configuration considered in this paper.

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 through 50) observe the *Imitators* with high probabilities, the *Innovators* with lower probabilities, and finally other *Regular Agents* with the lowest probabilities. This clearly suggests a *chain structure* to this social network: *Innovators* engage in individual learning without any reliance on social networks, *Imitators* learn mainly from *Innovators*, and *Regular Agents* learn mainly from *Imitators*.[9] 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.

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. Since the observation probabilities among agents belonging to the same group are likely to be 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 $g \neq g'$. Since an agent may also learn from other agents in his own group, we define two mean probabilities, f_{gg} and $f_{gg'}$, as follows:

$$f_{gg} = \frac{1}{|g|} \sum_{\forall i \in g} \left(\frac{1}{|g| - 1} \right) \sum_{\substack{\forall j \in g \\ j \neq i}} \overline{p}_i^j \tag{17}$$

$$f_{gg'} = \frac{1}{|g|} \sum_{\forall i \in g} \left(\frac{1}{|g'|} \right) \sum_{\forall j \in g'} \overline{p}_i^j \tag{18}$$

There are then nine different mean probabilities to be computed. Define a matrix F as the probability matrix showing all nine of them:

$$F = \begin{vmatrix} f_{NN} & f_{NM} & f_{NR} \\ f_{MN} & f_{MM} & f_{MR} \\ f_{RN} & f_{RM} & f_{RR} \end{vmatrix} \tag{19}$$

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 which are completely undeveloped – i.e., they observe others with equal probabilities – and they never get to develop the networks over the horizon.[10] However, the *Imitators*

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

[10] 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 $\mu^{im} = 0$ (as is the case for *Innovators* in the benchmark case), these probabilities are never adjusted over time.

in group M and the *Regular Agents* in group R do develop their social networks and the steady-state probabilities that ultimately emerge for these agents depend on all three parameters considered in this paper.

Shown in Figure 3 are the probability matrices, F, for $\sigma \in \{.9, .8, .7, .5\}$ when $\rho = 1$ and the $|N| : |M| = 4 : 6$. As expected, we get $f_{NN} = f_{NM} = f_{NR} (= \frac{1}{49} \approx .0204)$ 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 chain-like network structure is more pronounced in a more stable learning environment.[11]

In order to confirm the generality of the properties observed in Figure 3 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* since they do not develop social networks in the baseline case.

4.1 Steady-State Architecture and the Impact of *Innovator/Imitator* Mix

The collection of figures in Figure 4 displays (f_{MN}, f_{MM}, f_{MR}) for various σ and the $|N| : |M|$ mixes.[12] For each figure, the leftmost (rightmost) column of cells represents f_{MN} (f_{MR}) averaged over one hundred replications, while the center column of cells represents f_{MM}s. The first observation to be made from Figure 4 is that the cells in the leftmost column are brighter than those in the center column and the ones in the center column are brighter than those in the rightmost column for all considered values of σ and the $|N| : |M|$ mix. Hence, an *Imitator* observes an *Innovator* with the highest probability, another *Imitator* with a moderate probability, and a *Regular Agent* with the lowest probability: $f_{MN} > f_{MM} >> f_{MR}$.

Figure 5 contains similar information on (f_{RN}, f_{RM}, f_{RR}). The cells in the center column are brighter than those in the leftmost column, which, in turn, are brighter than those in the rightmost column. A *Regular Agent* observes an *Imitator* with a higher probability than he observes an *Innovator*: $f_{RM} > f_{RN} > f_{RR}$.

As can be seen from Figure 4 and Figure 5, these properties hold for all σ and the $|N| : |M|$ mixtures considered in our simulations. While not reported here, these properties also hold for all $\rho \in \{1, 4, 9\}$ as well as for various population sizes.[13]

[11] This is still assuming that the environment is sufficiently dynamic. Learning ceases to exist altogether in a *completely* static environment.

[12] For Figures 4 and 5, the visualizations of the probabilities captured in the sub-figures, (a)-(d), of a given figure are done on the basis of the same scale so that these sub-figures are visually comparable among themselves. For instance, Figure 4(a) and Figure 4(d) are visually comparable, but Figure 4(a) and Figure 5(a) are not.

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

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) $f_{RM} > f_{RN} > f_{RR}$.

From Figure 4, we note that f_{MN} monotonically decreases as $|N| : |M|$ ratio goes up from 1:9 to 9:1 and f_{MM} monotonically increases as $|N| : |M|$ ratio goes up from 1:9 to 8:2.[14] The intensity with which an *Imitator* focuses her attention on an *Innovator* tends to diminish as the number of *Innovators* relative to *Imitators* increases. The freed-up attention now goes to observing other *Imitators* who efficiently combine knowledge from an increasing number of *Innovators* – f_{MM} increases in $|N| : |M|$.

Similar observations can be made from Figure 5 which captures the probabilities of a *Regular Agent*: f_{RN} decreases and f_{RM} increases in $|N| : |M|$ ratio. It is clear that the importance of a super-type agent (an *Innovator* or an *Imitator*) to an average agent learning from him is positively related to the relative scarcity of the type in the system. When there is a decline (increase) in the relative availability of *Innovators* in the social system, an average *Imitator* observes an average *Innovator* with a higher (lower) probability. Likewise, when there is a decline (increase) in the relative availability of *Imitators* in the social system, an average *Regular Agent* observes an *Imitator* with a higher (lower) probability.

4.2 Impact of σ and ρ

One could infer the relationship between σ and the architecture of social networks from Figures 4 and 5 indirectly, but the inference can be made with greater clarity by considering the differential observation probabilities: $f_{MN} - f_{MM}$ and $f_{RM} - f_{RN}$. The former measures the extent to which an *Imitator* learns from an *Innovator* rather than another *Imitator*. The latter measures the extent to which a *Regular Agent* observes an *Imitator* over and above an *Innovator*. For these two measures, we focus on the *Innovators* and *Imitators* as the targets of observation, respectively, since they are the productive sources of performance improvements for *Imitators* and *Regular Agents*.

Figures 6(a) and 6(b) capture these differentials as functions of the $|N| : |M|$ mix and σ. First, notice that the differentials are always positive, thereby directly confirming Property 1. Second, $f_{MN} - f_{MM}$ monotonically declines and $f_{RM} - f_{RN}$ monotonically increases in $|N| : |M|$ ratio for all considered values of σ. This confirms our earlier observation that the relative importance of a super-type in the social network is directly related to his relative scarcity in the population. Finally, σ affects the two differential probabilities in different ways. First, $f_{RM} - f_{RN}$ increases monotonically in σ for all values of $|N| : |M|$ ratio. This implies that an increase in environmental volatility (a lower σ) induces a *Regular Agent* to learn directly from an *Innovator*, while a reduction in volatility induces a *Regular Agent* to learn through an *Imitator*. A volatile environment weakens the role of the *Imitators* as disseminators of knowledge in social networks: Knowledge takes time to go through the *Imitator*

[14]Note that $f_{MM} = 0$ for $|N| : |M| = 9 : 1$, since the population contains only one *Imitator*.

channel and it becomes obsolete faster with lower σ. Second, $f_{MN} - f_{MM}$ increases in σ for lower values of $|N| : |M|$ ratio and decreases in σ for sufficiently higher values of $|N| : |M|$ ratio. When there is a relative scarcity of *Innovators*, an increase (a reduction) in environmental volatility reduces (increases) the likelihood that an *Imitator* learns from an *Innovator* rather than from another *Imitator*. When there is a relative abundance of *Innovators*, an increase (reduction) in environmental volatility raises (lowers) the likelihood that an *Imitator* learns from an *Innovator* rather than from another *Imitator*.

Property 2: When *Innovators* are plentiful, an increase (decrease) in environmental volatility induces *Imitators* and *Regular Agents* to connect more (less) with *Innovators* rather than *Imitators*. When *Innovators* are relatively scarce, an increase (decrease) in environmental volatility induces *Regular Agents* to connect more (less) with *Innovators* but induces *Imitators* to connect less (more) with *Innovators.*

The non-monotonic impact of σ on $f_{MN} - f_{MM}$ requires a closer look. First, consider the case where $|N| : |M|$ is sufficiently high. When the number of *Innovators* is relatively large, they are viewed by the rest of the population as the main source of ideas. That $f_{MN} - f_{MM}$ decreases in σ in this case is based on the same intuition as that underlying the relationship between σ and $f_{RM} - f_{RN}$: As the environment is made more volatile, an *Imitator* finds it more effective to learn directly from an *Innovator* than through another *Imitator*. Next, consider the case where $|N| : |M|$ is sufficiently low. That $f_{MN} - f_{MM}$ *increases* in σ in this case is due to the subtle way in which the innovation activities of the *Regular Agents* respond to the limited availability of *Innovators* as well as the extent of the environmental volatility. The driving force here is that when there are fewer *Innovators* in the population, *Regular Agents* become a more attractive alternative source of ideas for *Imitators*. Their viability as a potential source of useful ideas depends on the following conditions: 1) the number of *Innovators* in the population is small so that the *Regular Agents* lack the external supply of ideas to copy; and 2) the environment is sufficiently volatile so that innovation is more effective than imitation. When these conditions are met, the *Regular Agents* will engage in innovation themselves and, in the process, become a useful source of ideas for *Imitators*. This is likely to be true to a greater extent when the environment is more volatile (i.e., lower σ).

That the fewness of *Innovators* in the population induces the *Regular Agents* to engage in more innovation is shown in Figure 7(a). Recall that $\overline{q}_i$ measures the endogenous probability with which an agent chooses to innovate (rather than imitate) in steady-state. We have collected and averaged the values of $\overline{q}_i$ for all $i \in R$. These steady-state probabilities are plotted in Figure 7(a) as a function of the $|N| : |M|$ ratio. Clearly, the agent chooses innovation with a higher probability when there are fewer *Innovators* in the population to supply the new ideas. As Figure 7(b) shows that the steady-state probability of a *Regular Agent* engaging in innovation monotonically decreases in σ, it validates the claim that *Regular Agents* play a more substantive role in generating ideas when the environment is more volatile. How this affects the imitation activities of the *Imitators* can be inferred from the probability matrices in Figure 3. When the environment becomes more volatile so that σ is lower, both f_{MN} and f_{MM} decrease in value, while f_{MR} increases. The crucial property is that f_{MN} decreases by more than f_{MM} when $|N| : |M|$ is low: the decrease in the value to an *Imitator* of observing an *Innovator* is likely to be bigger than the decrease in the value of observing another *Imitator*, since an *Imitator* still benefits from the increased innovations of *Regular Agents* and, hence, retains some of its usefulness as a potential target of imitation. Consequently, the differential probability, $f_{MN} - f_{MM}$, goes down when the environment is more volatile.

The varying influences of σ on the two differential probabilities, as captured in Figure 6, imply that an increase in σ (a more stable environment) strengthens the chain-like property of the social network when $|N| : |M|$ is sufficiently low – i.e., when there exists a relative shortage of *Innovators* in the system. When $|N| : |M|$ is sufficiently high, the result is mixed: An increase in σ raises the intensity with which a *Regular Agent* observes an *Imitator* over and above that with which he observes an *Innovator*, while it reduces the intensity with which an *Imitator* observes an *Innovator* over and above the intensity with which he observes another *Imitator*. Therefore, when $|N| : |M|$ is sufficiently high – i.e., when there exists a relative shortage of *Imitators* in the system – an increase in σ leads to a network structure with *Imitators* as the central hub. When $|N| : |M|$ is sufficiently low, an increase in σ leads to a more chain-like structure.

Another measure of the environmental volatility is ρ, which represents the magnitude of changes in the common goal. How does ρ affect the network architecture? We observe that the properties identified previously with respect to σ hold for ρ as well. In Figures 8 and 9, we report the findings on the observation probabilities, using the same methods that we employed in the previous section. Again, an increase in the degree of environmental stability (i.e., lower ρ) strengthens the pattern of chain-learning in the social network when $|N| : |M|$ ratio is sufficiently low, while it leads to the *Imitators* as the central hub when $|N| : |M|$ ratio is sufficiently high.

4.3 Robustness

All of the results presented here for the baseline case have also been replicated for alternative cases: 1) $(\mu_i^{in}, \mu_i^{im}) = (.75, .25)$ for all i in group N, $(\mu_i^{in}, \mu_i^{im}) = (.25, .75)$ for all i in group M, and $(\mu_i^{in}, \mu_i^{im}) = (.25, .25)$ for all i in group R and 2) $(\mu_i^{in}, \mu_i^{im}) = (.75, .25)$ for all i in group N, $(\mu_i^{in}, \mu_i^{im}) = (.25, .75)$ for all i in group M, and $(\mu_i^{in}, \mu_i^{im}) = (.5, .5)$ for all i in group R.

5 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: $(\mu_i^{in}, \mu_i^{im}) = (.75, .25)$

for all i in group N, $(\mu_i^{in}, \mu_i^{im}) = (.25, .75)$ for all i in group M, and $(\mu_i^{in}, \mu_i^{im}) = (.25, .25)$ 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.[15] How is the architecture of social networks affected by the availability of alternative learning mechanisms for the super-type individuals? As mentioned in 4.3, all of 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* in order to imitate their ideas, *Innovators* prefer to connect to *Imitators* rather than other *Innovators.* The latter property is clearly shown in Figure 10 which visualizes (f_{NN}, f_{NM}, f_{NR}) for various values of σ and $|N| : |M|$ ratio: The cells in the center column tend to be brighter than the cells in the left-most column, which, in turn, are brighter than those in the right-most column. 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 is connecting with fellow *Innovators.*

Property 3: When all agents can both innovate and imitate (though to varying degrees), *Innovators* tend to connect to *Imitators* (rather than other *Innovators*), while *Imitators* tend to connect to *Innovators* (rather than other *Imitators*).

The centrality of the *Imitators* in the social networks is further 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 $\sigma \in \{.9, .8, .7, .5\}$ and $\rho = 1$. Both $f_{NM} - f_{NN}$ and $f_{RM} - f_{RN}$ are positive for most values of $|N| : |M|$ and σ, but tend to decline as σ does.[16] This shows that the *Imitators* are central to social networks, even though the degree of their centrality diminishes as the environment becomes more volatile. Once a volatile environment exists, *Innovators* become more important. The impact of σ on $f_{MN} - f_{MM}$ is again mixed and dependent on the $|N| : |M|$ ratio: When $|N| : |M|$ is low (high), an increase in volatility reduces (increases) the intensity with which an *Imitator* observes an *Innovator* relative to that with which he observes another *Imitator.*

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

Having established the central role that *Imitators* play in the evolving social networks, we ask what the *socially optimal* mix of the super-type agents is. 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

[15] 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 – e.g., Mersenne Prime Numbers.

[16] In Figure 11(a), it appears that $f_{NM} - f_{NN}$ becomes negative for some $|N| : |M|$ ratios when $\sigma = .5$.

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?

For the baseline parameter configurations Figure 12 captures the steady-state aggregate performance, $\overline{\pi}$, as a function of the $|N|:|M|$ ratio for $\sigma \in \{.9, .8, .7\}$ and $\rho = 1$. It is clear from the figure that the aggregate performance is non-monotonic in the mix, with $|N|:|M| = 5:5$ emerging as the social optimum. The non-monotonicity of $\overline{\pi}$ is robust in that we observe the same property for all $\sigma \in \{.5, .7, .8, .9\}$ and $\rho \in \{1, 4, 9\}$. Furthermore, the optimal mix of $5:5$ appears to be invariant to changes in the values of σ and ρ.[17] While some parameter values may shift the optimum to $4:6$ or $6:4$, this rarely happens and, when it does, we can not reject the possibility that it is due to randomness. Observations made in Figure 12 then lead to the following general property.

Property 4: The aggregate performance is maximized when there is a heterogeneous mixture of *Innovators* and *Imitators* in the population.

The key insight here is that the maximal performance of a social system requires a mixture of *Innovators* and *Imitators*. The configurations of super-types consisting solely of *Innovators* (10 : 0) or of *Imitators* (0 : 10) produces strictly inferior aggregate performance.

Why is it that a heterogeneous 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 the group-M *Imitators* and the group-R *Regular Agents*. While the new ideas are generated by *Innovators*, these original ideas are scattered amongst 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 communicate with others, *Imitators* are easily able to observe and copy many of the valuable ideas that are distributed amongst the *Innovators*. When $|N| >> |M|$, 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 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 $|N| >> |M|$, the relative value of dissemination is important since the marginal social gain from an additional *Imitator* outweighs the marginal social

[17] We have also examined the aggregate performance as a function of the $|N|:|M|$ mix for a variety of population sizes, $M \in \{20, 30, 40, 50\}$, holding fixed the size of the super-types at ten. The optimal mix remains at 5:5 for all these population sizes.

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 performances at the *group* level are affected by the mixture of the super-types. To that end, we examine the average performance of an individual agent belonging to each type. Denote by $\overline{\pi}^k$ the average performance of a *single* agent in group $k \in \{N, M, R\}$:

$$\overline{\pi}^k = \frac{1}{|k|} \sum_{\forall i \in k} \overline{\pi}_i. \tag{20}$$

Figure 13 captures these group-level performances as functions of $|N| : |M|$ ratio, given $\sigma = .9$ and $\rho = 1$. Figure 13(a) shows that $\overline{\pi}^N$ tends to be independent of the mix of super types: an *Innovator* relies solely on his 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 non-monotonic in the ratio. Intuitively, an average *Imitator* (type-M agent) loses when the proportion of *Innovators* in the system decreases, since the sources of new ideas are drying out. 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 exists an insufficient amount of *original* ideas and an excessive amount of *redundant* ideas in the social system. The *Regular Agents*, who are the majority of the population, are then directly influenced by the presence and the prevalence of *Imitators*.

7 Concluding Remarks

When the population contains super-innovative and super-imitative individuals who are fully specialized (the baseline case), we find that the architecture of social networks evolves into a chain: *Innovators* generate ideas, *Imitators* learn from *Innovators*, and *Regular Agents* learn from *Imitators*. The overall flow of knowledge takes the form pictured in Figure 14(a), where *Imitators* act as connectors between *Innovators* and *Regular Agents*. When the super-innovative individuals have the capacity to imitate and the super-imitative 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*. Figure 14(b) captures the resulting flow of knowledge in the social system. In both cases, the importance of *Imitators* within the network is due to their ability to integrate dispersed knowledge in the social system. Their role as the 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 volatile (stable) task environment tends to support *Innovators* (*Imitators*) as the central target for social learning by the individual agents.

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.

References

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

[2] Chang, M.-H. and J. E. Harrington, Jr. 2005. Discovery and diffusion of knowledge in an endogenous social network. *American Journal of Sociology* **110** (4) 937-976.

[3] 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.

[4] 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.

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

[6] Kitcher, P. 1993. *The Advancement of Science: Science Without Legend, Objectivity Without Illusions.* Oxford University Press, New York.

[7] Skyrms, B. and R. Pemantle. 2000. A dynamic model of social network formation. *Proceedings of the National Academy of Sciences* **97** (16) 9340-9346.

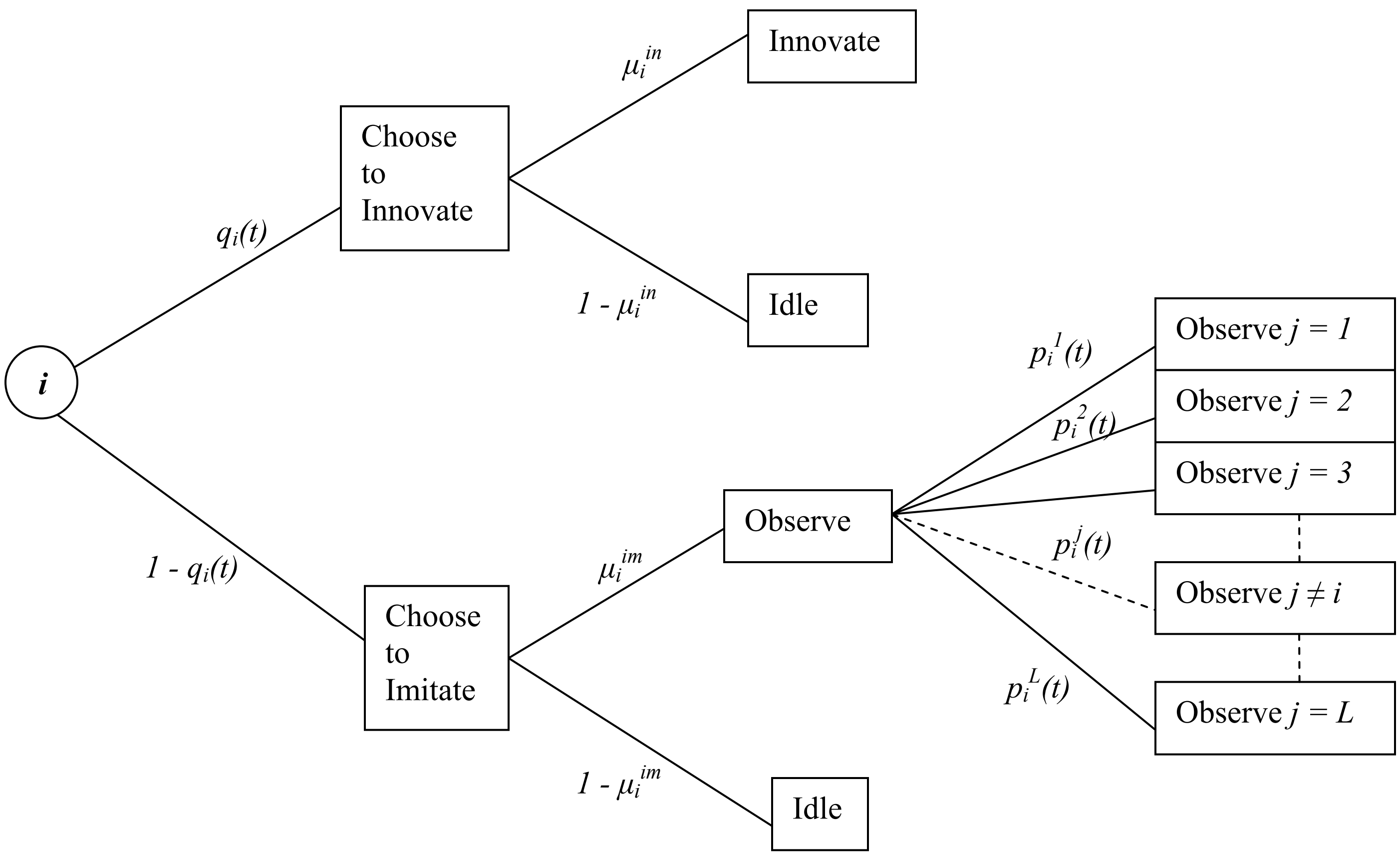


Figure 1: Decision Sequence of Individual *i* in period *t*

Figure 2: $\overline{p}_i^{\,j}$ for σ = .8 and |N|:|M| = 4:6

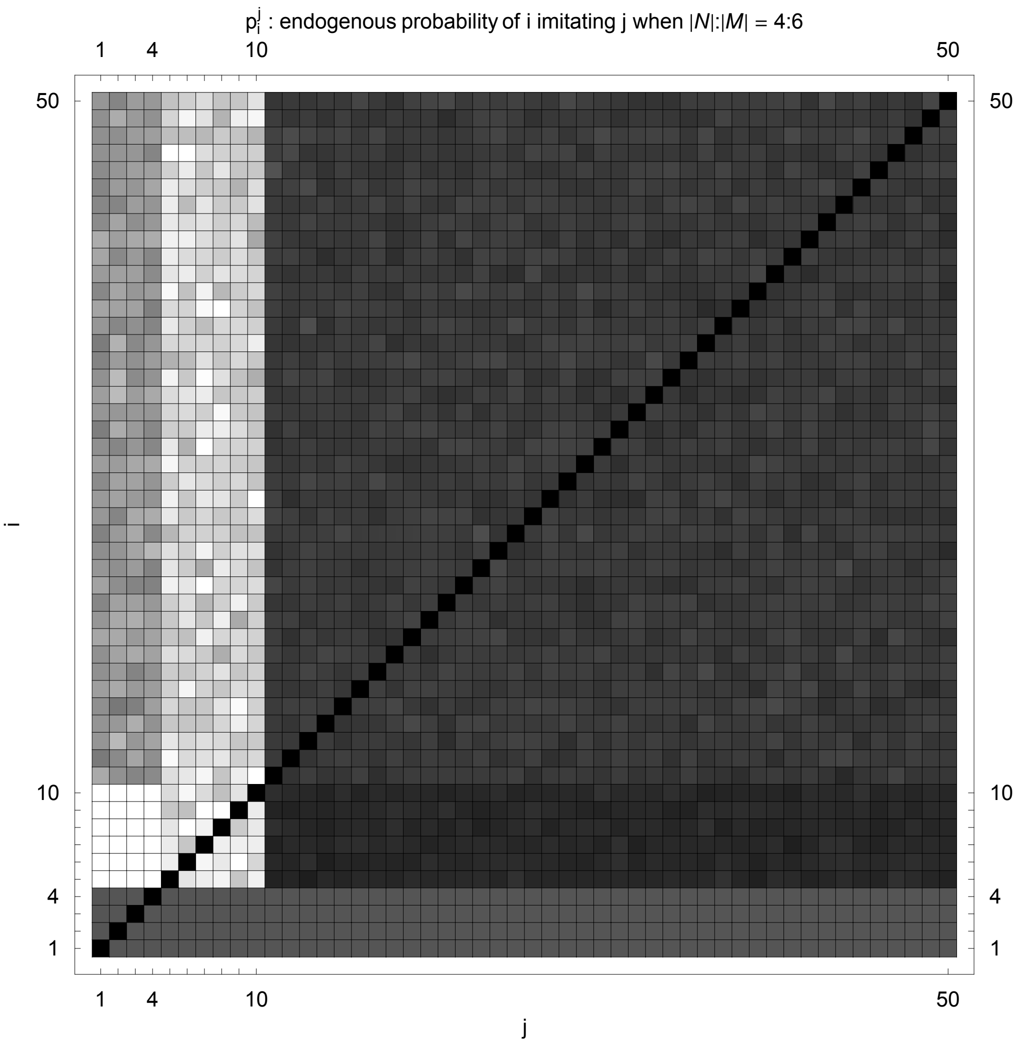


Figure 3: F matrix for $\rho = 1$ and $|N|:|M| = 4:6$

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

$$\text{(a) } \sigma = .9 : \begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0873 & 0.0582 & 0.0090 \\ 0.0356 & 0.0554 & 0.0135 \end{pmatrix}$$

$$\text{(b) } \sigma = .8 : \begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0843 & 0.0549 & 0.0097 \\ 0.0362 & 0.0512 & 0.0140 \end{pmatrix}$$

$$\text{(c) } \sigma = .7 : \begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0815 & 0.0523 & 0.0103 \\ 0.0355 & 0.0457 & 0.0150 \end{pmatrix}$$

$$\text{(d) } \sigma = .5 : \begin{pmatrix} 0.0204 & 0.0204 & 0.0204 \\ 0.0776 & 0.0462 & 0.0115 \\ 0.0338 & 0.0383 & 0.0163 \end{pmatrix}$$

Figure 4: Dependence of (f_{MN}, f_{MM}, f_{MR}) on $|N|:|M|$-mix [$\rho = 1$]

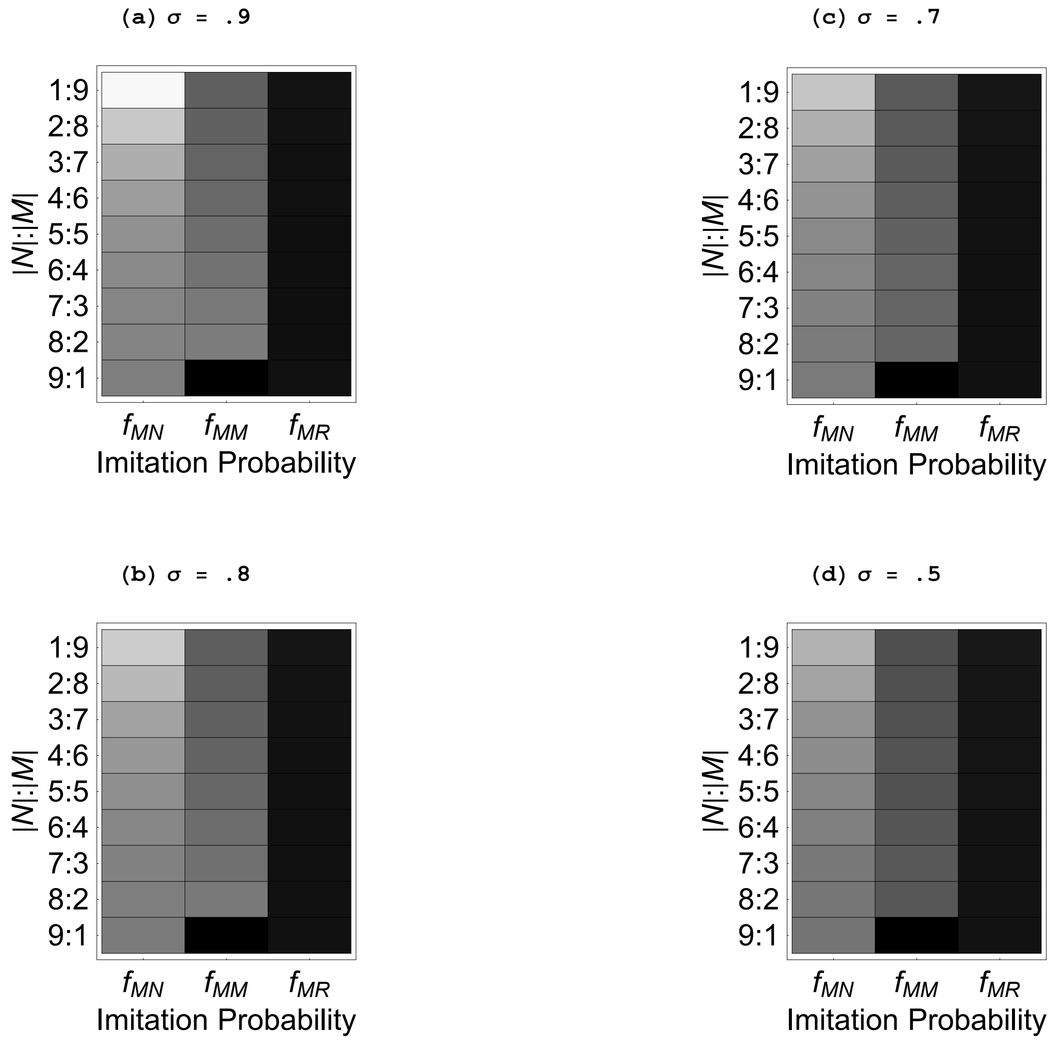


Figure 5: Dependence of (f_{RN}, f_{RM}, f_{RR}) on $|N|:|M|$-mix [$\rho = 1$]

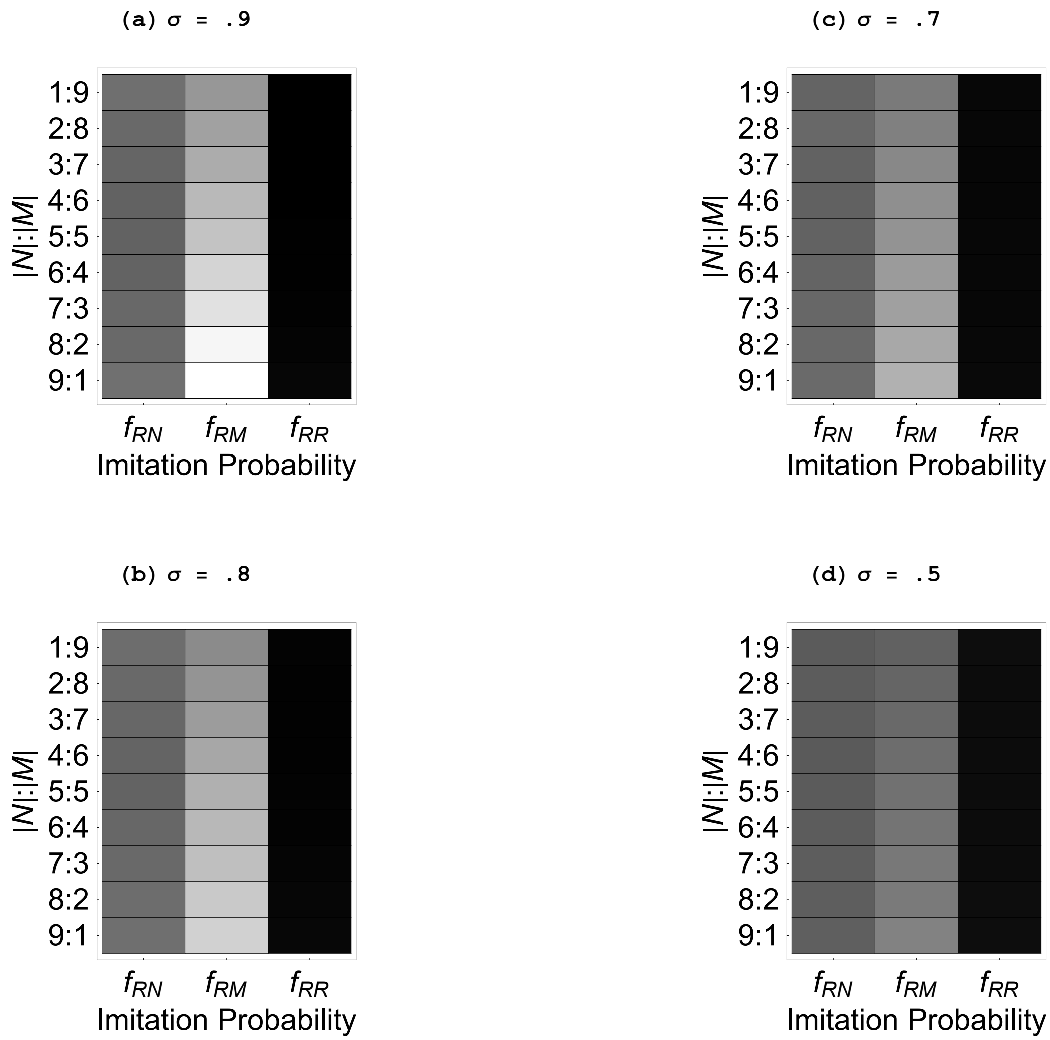
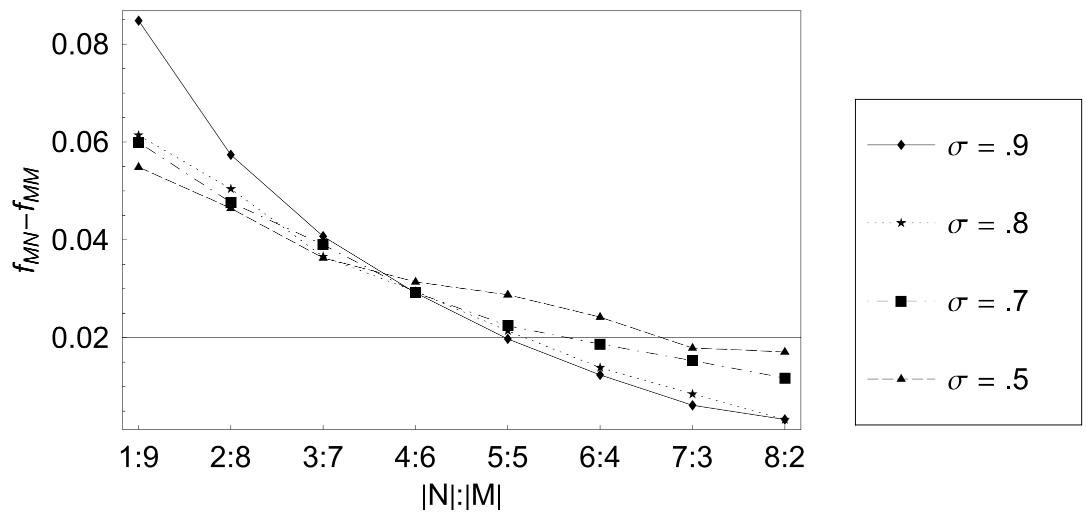


Figure 6: $f_{MN} - f_{MM}$ and $f_{RM} - f_{RN}$ [$\rho = 1$]

(a)

$f_{MN} - f_{MM}$



(b)

$f_{RM} - f_{RN}$

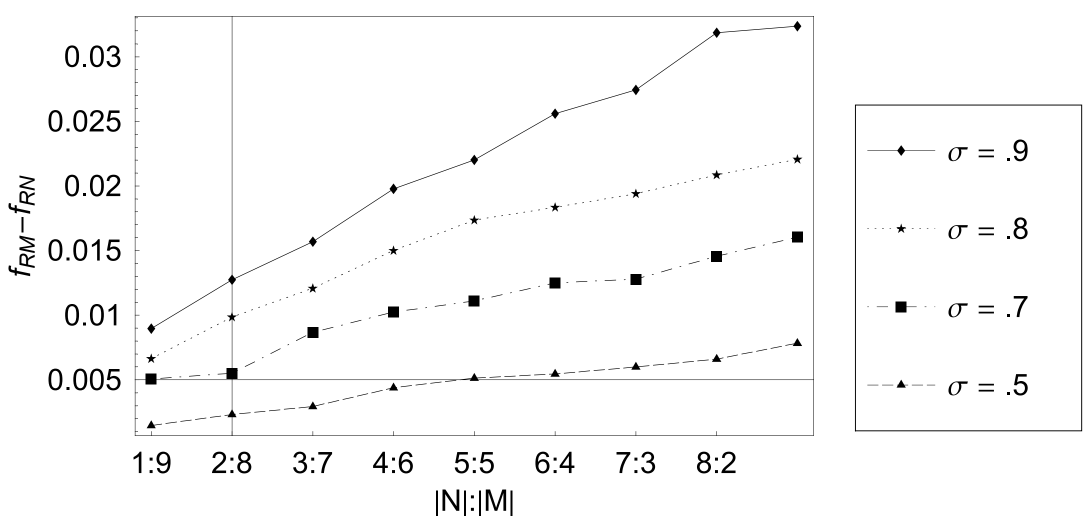
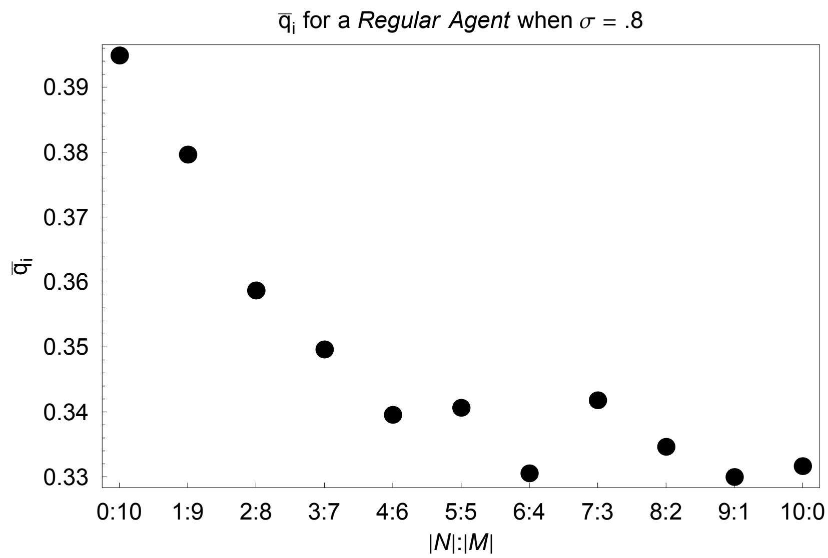


Figure 7: Steady-State $\overline{q}_i$ for a *Regular Agent*

(a) σ = 0.8



(b) $\sigma \in$ {0.5, 0.7, 0.8, 0.9}

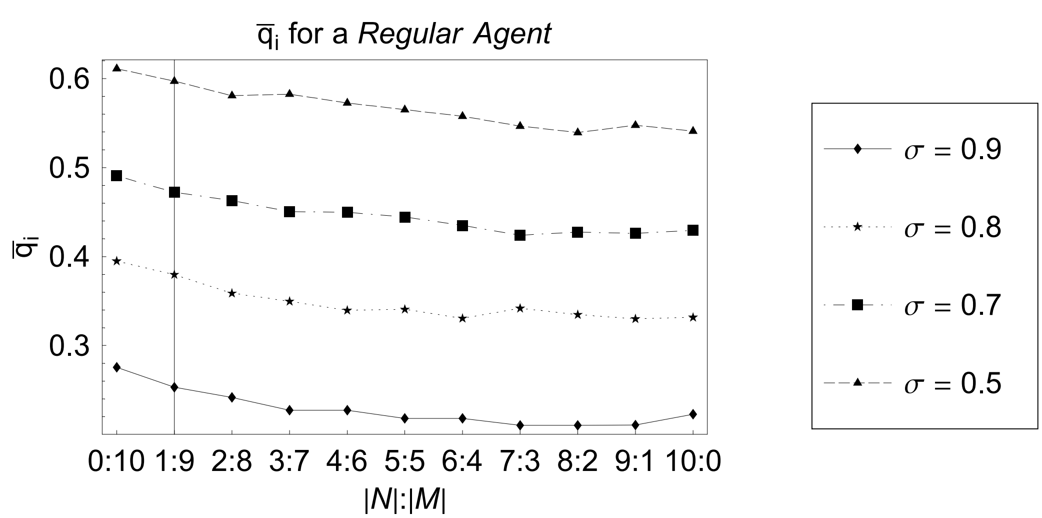


Figure 8: Impact of $|N|:|M|$ and ρ on (f_{MN}, f_{MM}, f_{MR}) and (f_{RN}, f_{RM}, f_{RR}) given $\sigma = .9$

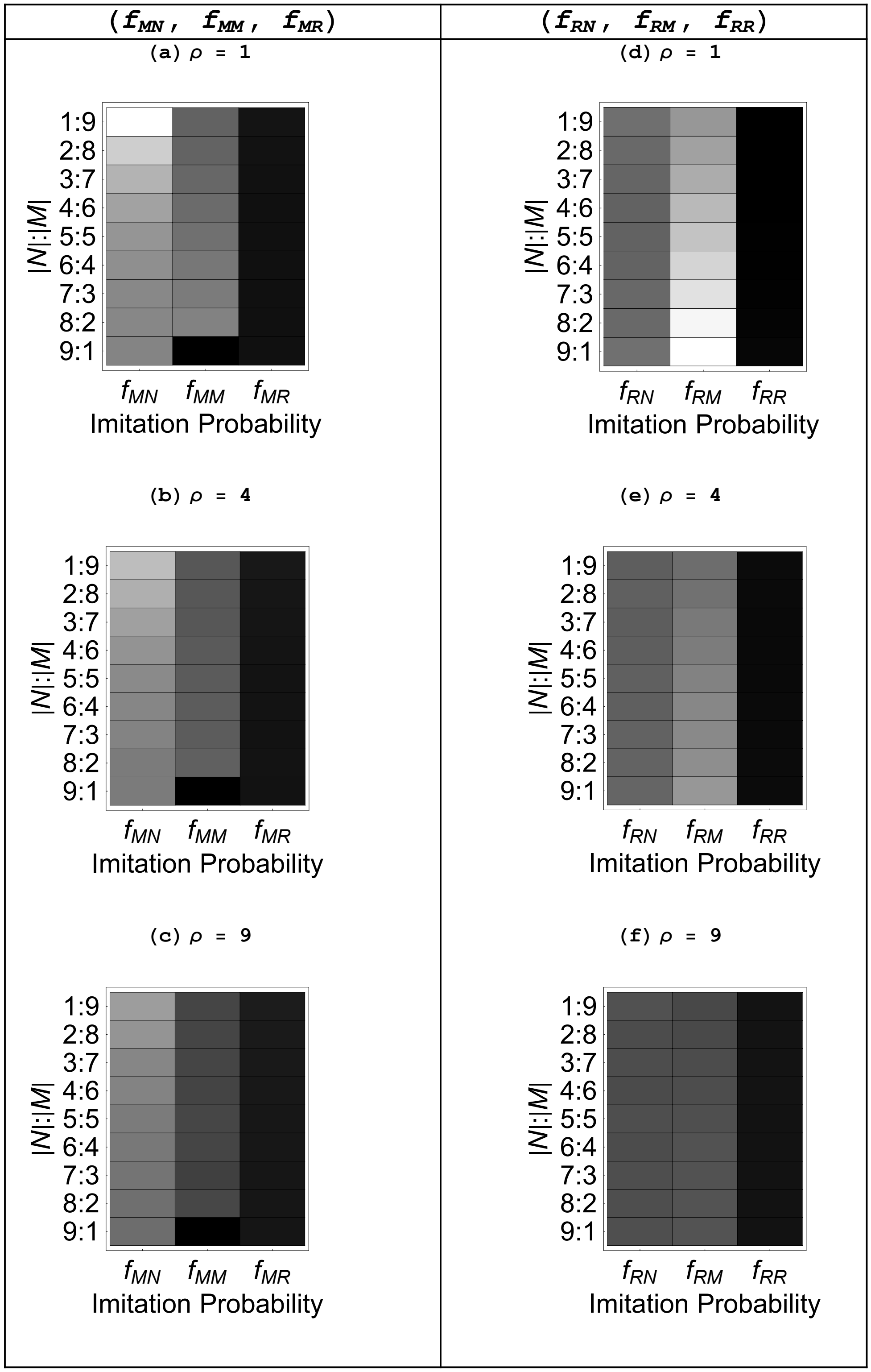
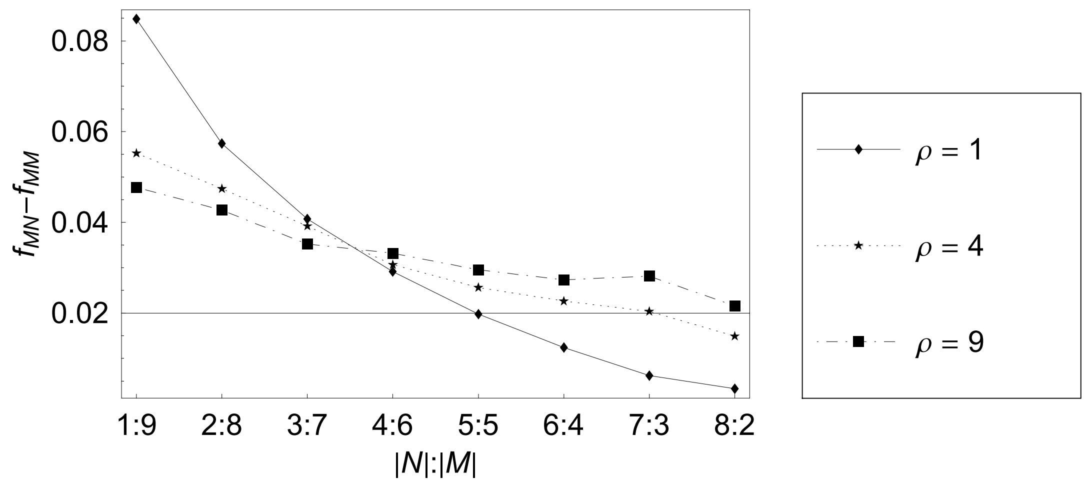


Figure 9: $f_{MN} - f_{MM}$ and $f_{RM} - f_{RN}$ [σ = .9]

(a)

$f_{MN} - f_{MM}$



(b)

$f_{RM} - f_{RN}$

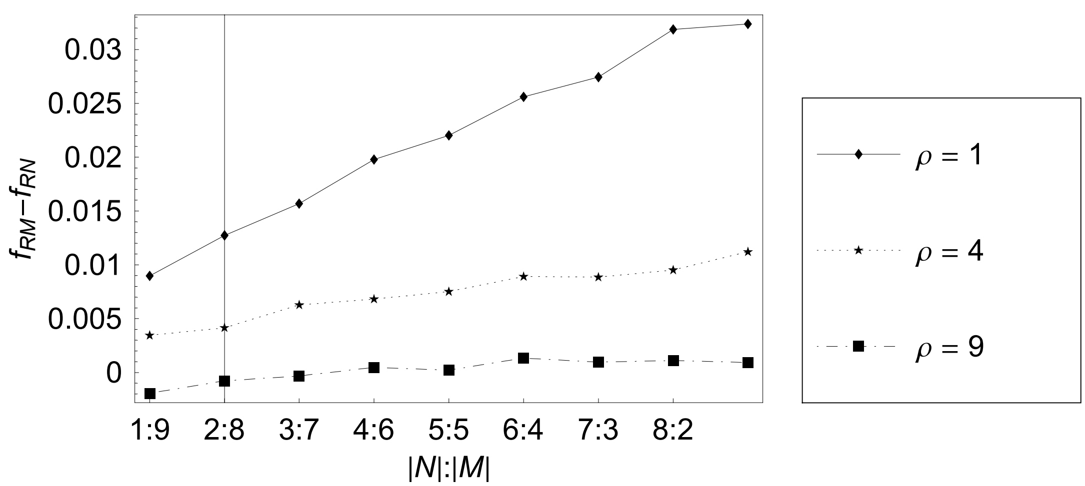


Figure 10: (f_{NN}, f_{NM}, f_{NR}) with $\rho = 1$

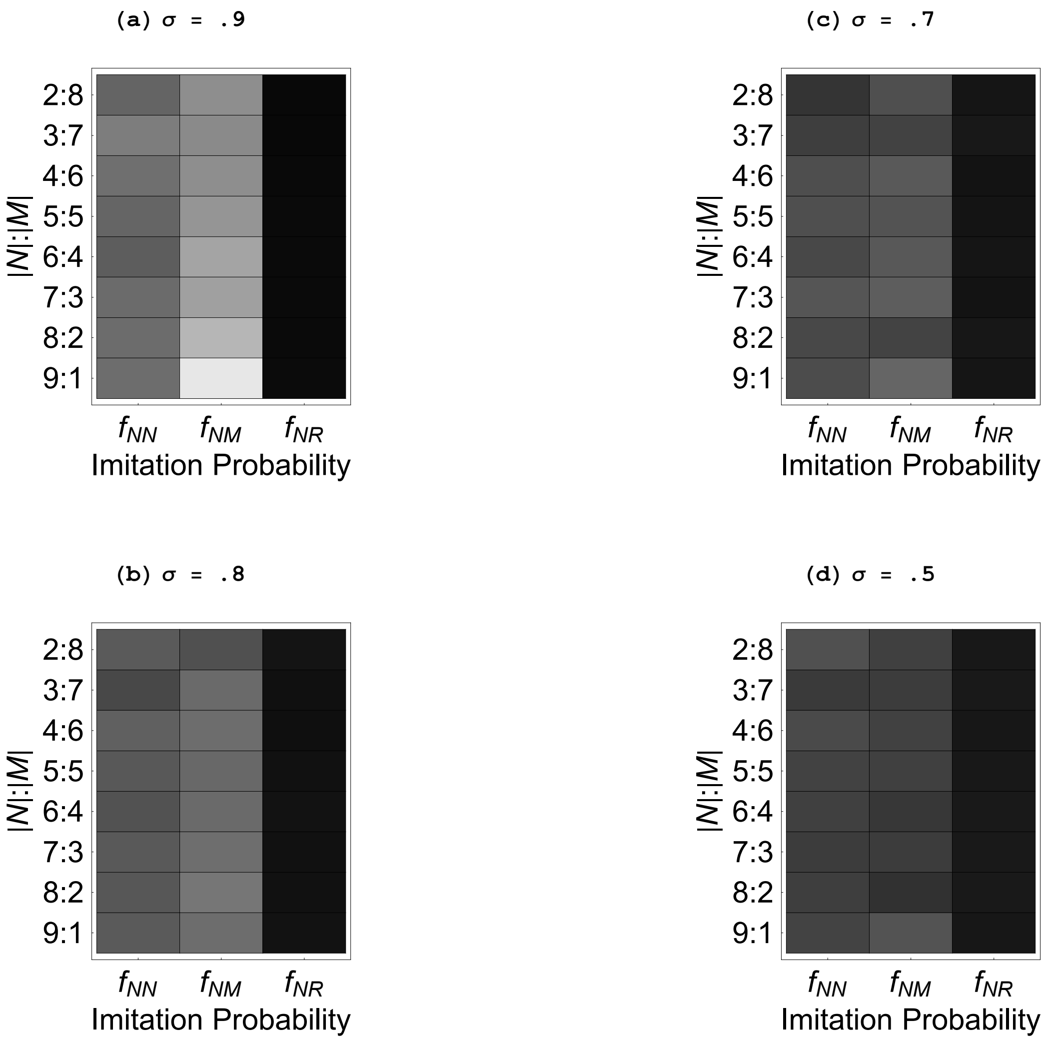


Figure 11: $f_{NM} - f_{NN}$; $f_{MN} - f_{MM}$; $f_{RM} - f_{RN}$ [$\rho = 1$]

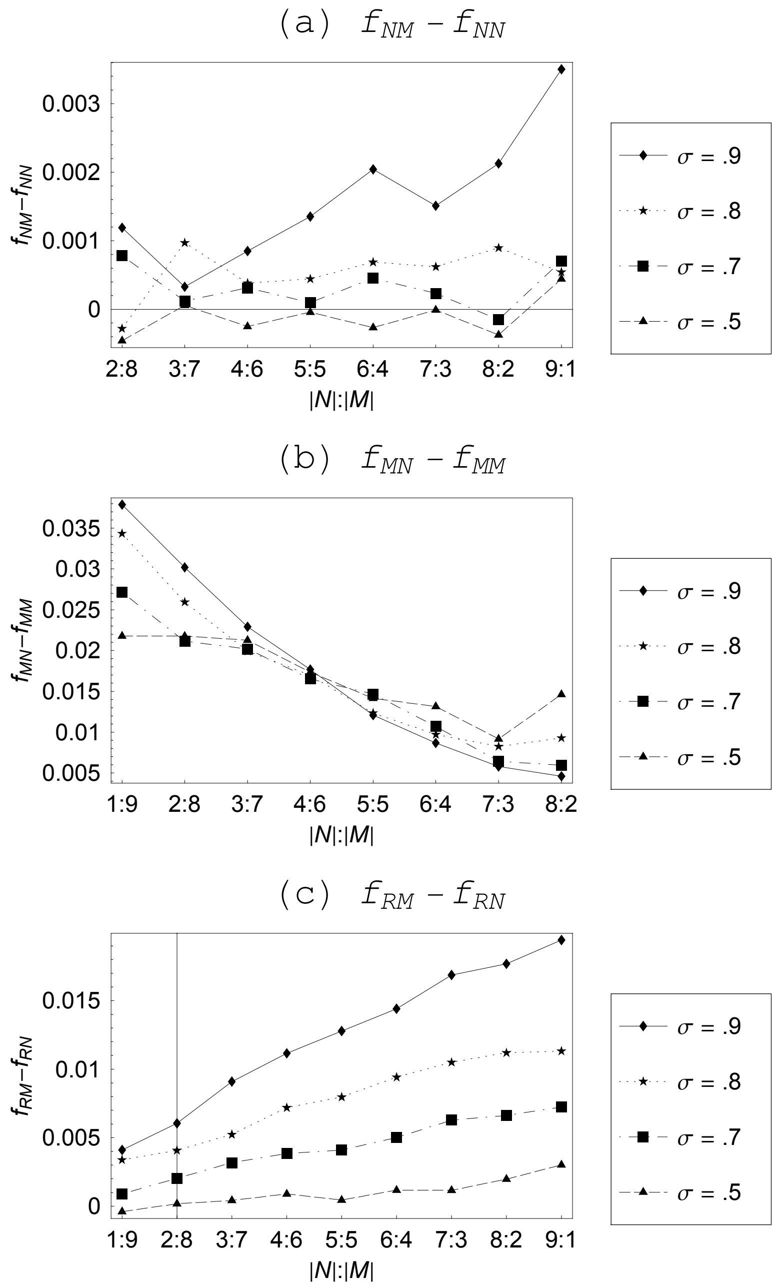


Figure 12: Aggregate Performance for ρ = 1

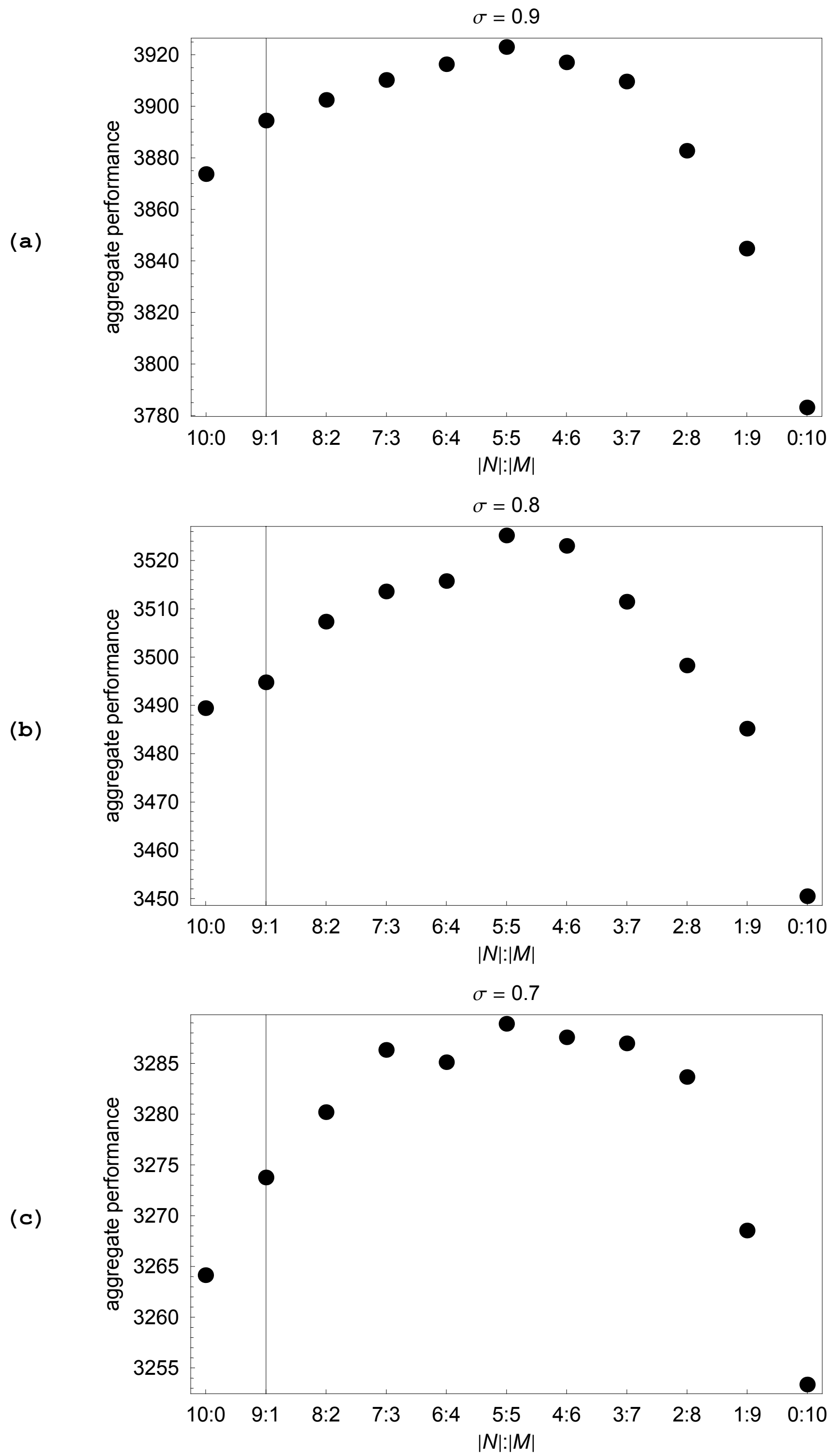


Figure 13: An Average Agent's Performance in Each Group
[σ = .9; ρ = 1]

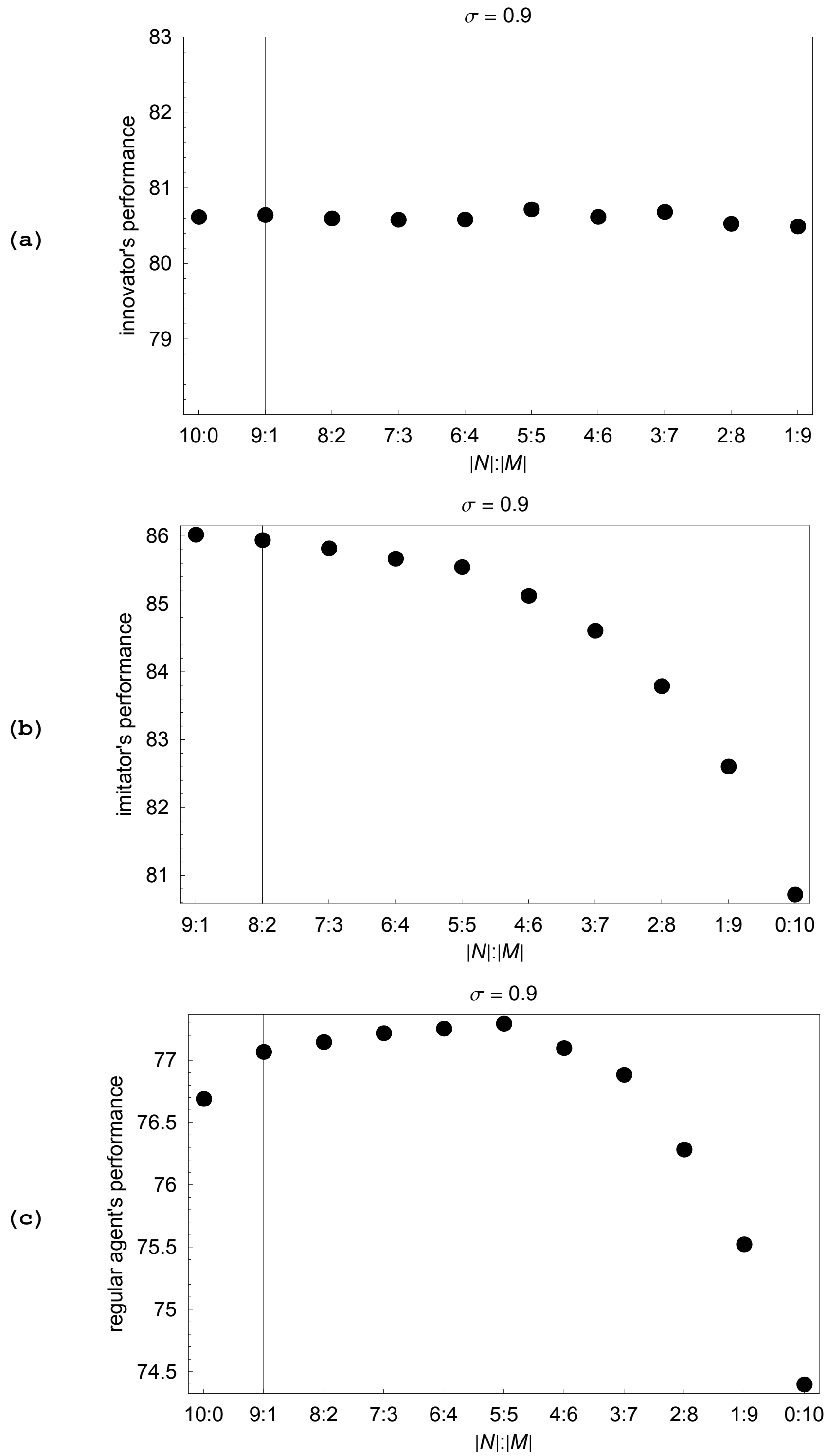
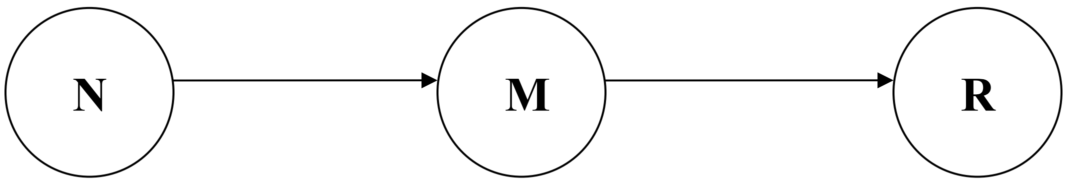


Figure 14: Flow of Knowledge

(a)



(b)