

Should Network Structure Matter in Agent-Based Finance?

SIMONE ALFARANO MISHAEL MILAKOVIĆ

Economics Department, University of Kiel
Olshausenstr. 40, D-24118 Kiel, Germany

April 24, 2007

Abstract. We derive microscopic foundations for a well-known probabilistic herding model in the agent-based finance literature. Lo and behold, the model is quite robust with respect to behavioral heterogeneity, yet structural heterogeneity, in the sense of an underlying network structure that describes the very feasibility of agent interaction, has a crucial and non-trivial impact on the macroscopic properties of the model.

Key words: Herding, networks, mean-field approach, N -dependence.

“Investors spend a substantial part of their leisure time discussing investments, reading about investments, or gossiping about others’ successes or failures in investing.” (Robert J. Shiller)

1 Introduction

Financial time series are characterized by a number of robust statistical regularities that show up across different countries, assets, and time frequencies. The two most prominent features concern the fluctuations in the prices of financial assets, which exhibit heavy tails and clustered volatility, and there are strong indications that both the distribution of large returns as well as the autocorrelation of transformations of returns are power-laws (see, e.g., Ding et al., 1993; Gopikrishnan et al., 1998; Lobato and Savin, 1998; Lux and Ausloos, 2002; Pagan, 1996).

While traditional finance has paid little attention to scaling laws, a variety of artificial financial market models with heterogeneous interacting agents have been recently proposed in order to account for the observed scaling laws (see, e.g., Brock and Hommes, 1997; Cont and Bouchaud, 2000; Iori, 2002; Kirman, 1991; Lux and Marchesi, 1999). The central ingredient of agent-based approaches is typically the *behavioral* heterogeneity among agents, who follow various rules by forming expectations or beliefs on which they act. Popular examples are models in the spirit of Brock and Hommes (1997), where agents choose a particular trading strategy based on its past profitability, or the models inspired by Kirman (1991), where agents exhibit herding tendencies in their choice of trading strategies. *Structural* heterogeneity, in the sense of an underlying *network structure* that defines social and institutional relationships, and thereby the very feasibility of agent interaction, has played a minor role so far.¹ This is not overly surprising, given that behavioral heterogeneity initially went a long way in replicating the stylized facts of financial time series.

Following Lux (2006), we distinguish three broad model classes in agent-based finance: (i) dynamical systems models with switches between attractors, (ii) critical systems approaches, and (iii) herding models.

¹The model of Cont and Bouchaud (2000) is a notable exception; their results depend directly on the underlying network structure, in particular on a critically tuned random graph.

The models in the first category are inspired by the work of Brock and Hommes (1997) which, contrary to models in the other two categories, are not interaction-based since their heterogeneous and boundedly rational agents do not interact with each other, but merely through the price mechanism. The metaphor of a Walrasian auctioneer implicitly assumes that the underlying network structure is a ‘star,’ where the auctioneer is placed at the center of the network. The drawback of this class of models is that they need fine-tuning of their “signal-to-noise ratio” around unity in order to resemble the stylized facts.² In the second category, several authors have proposed to exploit the properties of well known critical systems from the statistical physics literature (see, e.g., Cont and Bouchaud, 2000; Iori, 2002; Bornholdt, 2001; Stauffer and Sornette, 1999). The underlying network structures in the criticality models are typically regular lattice structures that are routinely utilized in a branch of physics known as percolation theory (see, e.g., Stauffer and Aharony, 1994). These critical systems approaches, however, show an extreme dependence of the resulting power laws on carefully adjusted model parameters near criticality,³ and leave open how or why the market could self-organize into such a critical state. In the third category, herding approaches (see, e.g., Kirman, 1991, 1993; Lux and Marchesi, 1999; Alfarano et al., 2005) emphasize the processes of social interaction among agents who display a tendency to follow the crowd. Herding models implicitly assume either a fully connected network, where everyone can in principle directly interact with everyone else, or some regular structure that confines interactions to a local neighborhood. The drawback of the herding-based paradigm is that many of the proposed models are not robust with respect to an enlargement of system size, which usually coincides with the total number of agents, say, N . In many cases, the interesting properties of fluctuations in returns, namely their peculiar time-dependence structure and the power law decay of the return distribution, progressively disappear as soon as the number of agents is enlarged,

²The noise component that is responsible for switches between attractors in the deterministic ‘skeleton’ of these models needs to be superimposed on the macroscopic level. If it is too ‘small,’ the deterministic forces dominate; if it is too ‘large,’ the stochastic component dominates the system. Thus a fine-tuning of the variance in the added noise seems crucial for obtaining realistic time-series.

³While Cont and Bouchaud (2000) do not use a regular network structure, the above criticism applies as well to their model because their results depend on the critical tuning of a random graph.

showing instead Gaussian fluctuations and a weak degree of temporal dependence (see, e.g., Egender et al., 1999). Given the computational nature of most agent-based models, it is often not possible to determine analytically whether a specific model exhibits this N -dependence effect or not.⁴ It is clearly not our ambition to engage in arguments about the relative advantages or drawbacks of the various agent-based approaches. Instead we start by focusing our attention on the crucial impact that network structure has on the statistical equilibrium outcome of “type-(iii)” models, i.e. those that are built around a generic probabilistic herding mechanism.

Basically, our research is motivated by two observations. The first is that financial investment decisions carry an important social component, concisely summarized in the introductory quote by Shiller (1984), and it seems natural to model social relationships through the framework of various network structures. The second observation is inspired by a recent model of Alfarano et al. (2006), which manages to address the problem of N -dependence analytically, and hints at the importance of structural heterogeneity in type-(iii) models. We take the position that the underlying network structure is directly linked to the problem of N -dependence in this generic herding mechanism, and we investigate which network structures would be capable of overcoming the problem. Lo and behold, the network structures that are immune to an enlargement of system size are not an entirely unlikely way of organizing social relationships.

2 Probabilistic Herding Model

In a prototypical interaction-based herding model of the Kirman (1993) type, the population of traders of size N is divided into two groups, say, X and Y . Depending on the model setup, the two groups are typically labeled as fundamentalists and chartists, or optimists and pessimists, or buyers and sellers.

⁴The crucial dependence on system size also shows up outside of the herding-based models, e.g. in the genetic learning model of Lux and Schornstein (2005). Intuitively, the source of N -dependence is the central limit theorem, as already pointed out by Cont and Bouchaud (2000).

2.1 Extensive vs Non-Extensive Transition Rates

The interaction among these financial investors is formalized via the following transition rates

$$\pi(n \rightarrow n+1) = (N-n) \left(a + \beta \frac{n}{N} \right), \quad (1)$$

$$\pi(n \rightarrow n-1) = n \left(a + \beta \frac{(N-n)}{N} \right), \quad (2)$$

where n/N and $(N-n)/N$ are the fractions of agents in either state. The parameter a expresses idiosyncratic factors for switches between groups,⁵ and β captures the herding effect. A crucial property of the transition rates (1) and (2) is their *extensivity*, i.e. the fact that $\pi(n \rightarrow n') = N \pi(z \rightarrow z')$ where $z = n/N$ is shorthand for the concentration of agents in one of the two groups.

An apparently minor modification of the above transition rates would be to formulate them in the following *non-extensive* fashion

$$\pi^+ = \pi(n \rightarrow n+1) = (N-n) (a + bn), \quad (3)$$

$$\pi^- = \pi(n \rightarrow n-1) = n (a + b(N-n)), \quad (4)$$

where the herding term now does not depend on the concentration of agents in the opposite state, but rather on the total number (or occupation number) of agents in the other state. In this case the above transition rates no longer relate linearly in N when expressed in terms of concentrations, $\pi(n \rightarrow n') = N^2 \pi(z \rightarrow z')$, and hence we label them as non-extensive. It turns out that this modification has major implications for the behavior of the system when the number of agents is enlarged. Alfarano et al. (2006) demonstrate analytically that the emergence of realistic non-Gaussian dynamics is independent of system size in a simple Walrasian asset pricing model if transition rates are non-extensive, while an approach to Gaussian behavior occurs under extensive rates.⁶ They offer an interpretation of the

⁵Alfarano et al. (2005) have also investigated an asymmetric setting, where the idiosyncratic component a differs among the transition rates, allowing for a preferential direction of switches between the two states. But the impact of network structure on the generic herding model does not depend on the symmetry of a , since an asymmetric setting still leads to an equilibrium beta distribution (though with different parameters ϵ_1, ϵ_2).

⁶Technically speaking, both sets of transition rates define non-linear Markovian *one-step processes*, which can be respectively approximated by a continuous diffusion process, gov-

various transition rates in terms of “local” versus “non-local” interactions (corresponding to extensive versus non-extensive transition rates), but they neither make this notion precise nor do they address how a prevalence of non-extensive rather than extensive interaction strengths might emerge in the underlying social processes among agents. Since social interaction in the form of herding tendencies is a crucial ingredient of these types of models, we believe that an answer to the above questions needs to be sought via the network structure that connects individual agents in their social interactions as market participants.

2.2 Microscopic Interpretation of Transition Rates

We start our inquiry into the influence of network structure by deriving microscopic foundations for the mesoscopic transition rates (1),(2),(3), and (4). Let n denote the number of agents in state X , so $N - n$ agents are in state Y . The basic idea is that agents change state for personal reasons or under the influence of their *neighbors*, with whom they interact during a given time period. The transition rate for agent i to switch from state X to state Y is

$$\rho_i(X \rightarrow Y) = a + \lambda \sum_{j \neq i} D_Y(i, j), \quad (5)$$

where a governs the possibility of self-conversion due to idiosyncratic factors, e.g. the arrival of new (private or publicly available) information, while the sum captures the influence of the neighbors. Notice that the influence of i 's neighbors on i is linear in (5), implying that the impact of each neighbor j carries an equal weight. The parameter λ governs the interaction strength between i and neighbor j , and for now we will assume that

erned by the so-called Fokker-Planck equation. (We provide the corresponding heuristics in the appendix; for more details, see Alfarano et al., 2006). It turns out that the drift functions are equal under both, the extensive and non-extensive transition rates, and depend linearly on the idiosyncratic parameter a . While both diffusion functions depend on the herding parameter (β and b , respectively), they differ substantially for extensive and non-extensive transition rates: in the former case, the diffusion is N -dependent, while in the latter case it is independent of N . This implies that fluctuations will become increasingly negligible for the system dynamics when the number of agents N is enlarged in the extensive case, and consequently the resulting time-series of returns will exhibit counter-factual Gaussian characteristics. In the non-extensive case, on the other hand, persistent fluctuations between opinions and along with them the empirical characteristics of returns, are preserved for any system size.

both parameters are constant across agents.⁷ The function $D_Y(i, j)$ is an indicator function serving to count the number of i 's neighbors that are in state Y ,

$$D_Y(i, j) = \begin{cases} 1 & \text{if } j \text{ is a } Y\text{-neighbor of } i, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

Analogously, the transition rates in the opposite direction are given by

$$\rho_i(Y \rightarrow X) = a + \lambda \sum_{j \neq i} D_X(i, j). \quad (7)$$

Defining $n_Y(i, j) = \sum_{j \neq i} D_Y(i, j)$ and $n_X(i, j) = \sum_{j \neq i} D_X(i, j)$, and using shorthands $\pi_i^- = \rho_i(X \rightarrow Y)$ and $\pi_i^+ = \rho_i(Y \rightarrow X)$, equations (7) and (5) can be compactly written as

$$\pi_i^+ = a + \lambda n_X(i, j), \quad (8)$$

$$\pi_i^- = a + \lambda n_Y(i, j). \quad (9)$$

The dependence on j indicates that the rates π_i^\pm still depend on the particular configuration of the neighbors.

2.3 Mean-Field Approximation

The dependence on the neighbors in (8) and (9) is difficult to handle analytically in the present form. We can employ a mean-field approximation in order to simplify the problem from a many-agent system to one with a sum of agents who are independently acting in an “external field” created by the other agents (see, e.g., Chap. 5 in Aoki, 1998). In other words, we assume that individual agents are influenced by the average opinion of their neighbors. The instantaneous probability for agent i to switch from X to Y is given by (9). As the neighbors' attitudes fluctuate, π_i^- fluctuates around its mean

$$\langle \pi_i^- \rangle = a + \lambda \langle n_Y(i) \rangle, \quad (10)$$

where the dependence on j gets lost if we assume that inhomogeneities among the different configurations of neighbors are solely due to the fluctuations. Next, we replace the number of Y -neighbors around each agent i

⁷We will relax this assumption in Section 5 when we allow for individual heterogeneity in the behavioral parameters a and λ .

with the average number of neighbors that agents are linked to, say, D . Then the quantity $\langle n_Y(i) \rangle$ can be expressed as

$$\langle n_Y(i) \rangle = D P_Y, \quad (11)$$

where P_Y is the ‘unconditional’ probability that an i -neighbor is in state Y , which we approximate via the relative frequency $(N - n)/N$ of agents in state Y , thus obtaining

$$\langle \pi_i^- \rangle = a + \lambda D \frac{N - n}{N}. \quad (12)$$

The quantity $\langle \pi_i^- \rangle$ is now independent of the particular configuration of neighbors, and it is equal for every agent i in state X . Symmetrically, the expression for agents currently in state Y is

$$\langle \pi_i^+ \rangle = a + \lambda D \frac{n}{N}. \quad (13)$$

Hence the mean field approximation reduces the many-agent system to a collection of independent agents who are “acting in the field” created by the other agents. Since we are interested in the probability of observing a *single* switch on the system-wide level (and not in switches of a particular agent i), we have to sum (13) over all agents in state Y in order to find the aggregate probability that *one* agent is switching from state Y to state X . We further assume that the time unit over which we look at the process is small enough to constrain the switch to a single agent. Summing (13), which is permissible since the agents are now independent, we obtain

$$\pi^+ = (N - n) \left(a + \frac{\lambda D}{N} n \right), \quad (14)$$

for a switch from Y to X , and

$$\pi^- = n \left(a + \frac{\lambda D}{N} (N - n) \right), \quad (15)$$

for the reverse switch, where (14) and (15) are generalized transition rates of Kirman’s ant model. Recalling the extensive transition rates (1) and (2), and the non-extensive rates (3) and (4), we see that $\beta = \lambda D = bN$.

Obviously, the particular network topology connecting the agents will

have an influence on the *average connectivity* D , i.e. the average number of neighbors in a given network structure. Non-extensive transition rates are only feasible if D increases on the same order of magnitude as the system size N . Kirman's original interpretation of random pairwise meetings corresponds to the special case where D is always equal to unity, which leads to extensive transition rates, and therefore suffers from the problem of N -dependence. A central issue of the generic herding model thus revolves around the question which network structures would lead to extensive or non-extensive transition rates. In light of the microscopic foundations of the herding model, we are particularly interested in the average connectivity of various network structures, and in comparing the accuracy of the mean-field approximation with microscopic implementations of the probabilistic herding model in these networks. In order to do so, we shall first clarify the concept of a statistical equilibrium outcome that is associated with the generic transition rates (14) and (15).

2.4 Equilibrium Distribution

At any time, the *state of the system* refers to the concentration of agents in one of the two states. We define the state of the system through the concentration $z = n/N$ of agents that are in state X . For large N , the concentration can be treated as a continuous variable. Notice that none of the possible states of $z \in [0, 1]$ is an equilibrium in itself, nor are there multiple equilibria in the usual economic meaning of the term. The notion of equilibrium rather refers to a statistical distribution that describes the proportion of time the system spends in each state. Utilizing the Fokker-Planck equation, we show in the appendix that for large N the equilibrium distribution of z is a beta distribution,

$$p_e(z) = \frac{1}{B(\epsilon, \epsilon)} z^{\epsilon-1} (1-z)^{\epsilon-1}, \quad (16)$$

where $B(\epsilon, \epsilon) = \Gamma(\epsilon)^2 / \Gamma(2\epsilon)$ is Euler's beta function, and the shape parameter of the distribution is given by $\epsilon = a/b = aN/\lambda D$. Since ϵ is a ratio of quantities that depend (i) on the time scale at which the process operates ($1/a$ and $1/\lambda$), and (ii) on the spatial characteristics of the underlying network (D and N), the parameter of the equilibrium distribution is a well-defined dimensionless quantity. If $\epsilon < 1$ the distribution is bimodal,

with probability mass having maxima at $z = 0$ and $z = 1$. Conversely, if $\epsilon > 1$ the distribution is unimodal, and in the “knife-edge” scenario $\epsilon = 1$ the distribution becomes uniform. The mean value of z , $E[z] = 1/2$, is independent of ϵ , but the system exhibits very different characteristics depending on the modality of the distribution. In the bimodal case, the system spends least of the time around the mean, mostly exhibiting very pronounced herding in either of the extreme states.

Concerning the issue of N -dependence, the variance of z , $Var[z] = E[z^2] - E[z]^2 = 1/4(2\epsilon + 1)$, turns out to be a convenient summary measure of the model properties with respect to an enlargement of system size. If the variance of z remains constant when the system is enlarged, the power law scaling of returns will be preserved in a simple Walrasian market clearing scenario, while a decrease of the variance under an enlargement of the system leads to counter-factual Gaussian properties of returns (see Alfarano et al., 2006). The variance of z scales hyperbolically with N , so the inverse of the variance scales linearly with N ,

$$Var[z]^{-1} = 4 \left(\frac{2a}{\lambda} \frac{N}{D} + 1 \right). \quad (17)$$

After introducing various network structures in the next section, we first want to check whether the mean-field approach provides a good approximation to the predicted equilibrium distribution on the various networks. Then, by simulating a microscopic version of the mesoscopic transition rates (3) and (4) on different networks of increasing size, we would like to know which (if any) of the networks can overcome the problem of N -dependence. If the variance of z is invariant under increases in system size for some network, then this network would be able to provide the connectivity that is necessary in order to redeem the generic model from its N -dependence. Finally, we are going to relax the assumption that all agents act in identical fashion. The model would appear quite robust if we could allow for different ways of processing new information and for different herding intensities by varying the behavioral parameters a and λ across agents, and yet recovered equilibrium distributions that are in line with the analytical predictions of the mean-field approximation.

3 Prototypical Network Structures

We start our investigation by briefly reviewing the formal definition of a network. A network $G = \{V, E\}$ is defined by a set of *nodes* $V = \{1, \dots, N\}$, and a set of *links* E .⁸ In our case, the nodes represent agents, while a link between two agents $i, j \in V$ represents direct interaction or communication such that $\{i, j\} \in E$. We can denote links between agents by a binary variable $a(i, j) \forall i, j \in V$ such that $a(i, j) = 1$ if $\{i, j\} \in E$, and $a(i, j) = 0$ otherwise. Then the network can be summarized by an *adjacency matrix* $A_{N \times N}$; since we restrict our attention to *undirected* (or *bi-directional*) networks, the adjacency matrix will be symmetric, i.e. $a(i, j) = a(j, i)$. The *neighborhood* of i is the set $\Phi_i = \{j \in V : a(i, j) = 1\}$. A *path* that connects two agents $r, s \in V$ is defined by a set of pairs $\{(r, r_i), \dots, (r_j, s) : a(r, r_i) = \dots = a(r_j, s) = 1\}$; the *geodesic path* between two agents is the shortest path connecting them, and the *diameter* of G is the number of links of the longest geodesic path between any two nodes. Finally, each node in the network has a *degree* that corresponds to the number of links that are connected to the node. Let p_k denote the fraction of nodes that have degree $k \in K = \{0, \dots, N\}$, corresponding to the probability that a node chosen uniformly at random has degree k ; then the set $P = \{p_k, k \in K\}$ denotes the *degree distribution* of the network.

Since the equilibrium distribution of the probabilistic herding model depends on the average connectivity D , knowledge of the degree distribution enables us to compute the average connectivity of a given network, and to make theoretical predictions about the capability of various network structures to reproduce the stylized facts under an enlargement of system size.⁹

⁸This terminology is typically used by computer scientists, while mathematicians speak of *graphs* that are defined by a set of *vertices* connected through *edges*.

⁹Over the past decade, economists have started to investigate the interactions between network topologies and game-theoretic notions of *stability* and *efficiency* using the tools and jargon of graph theory. From our point of view, which mainly focuses on the statistical properties of network structures, it is noteworthy that the term *degree distribution* does not appear once in a recent survey of the economic literature on network formation (see Jackson, 2004), and it seems that the profession is only very recently becoming aware of the interdisciplinary research by statistical physicists and computer scientists that emphasizes the statistical features of empirically observed networks.

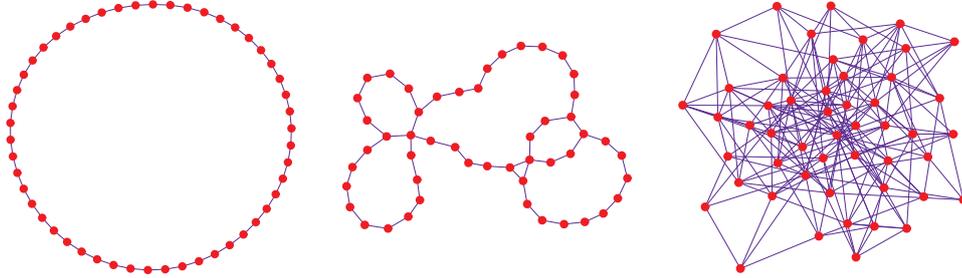


Figure 1: Three different network structures with $N = 50$ nodes (from left to right): a circle with neighborhood two; a small-world network built on the circle by randomly adding five shortcuts; and a random network with linking probability $p = 0.15$.

3.1 Regular Network Structures

In a regular structure, all agents are connected to a constant number \bar{k} of neighbors. The simplest case would be a circle that connects agents to their two nearest neighbors,¹⁰ illustrated in the left panel of Figure 3.1. Regular structures have a degenerate degree distribution since all nodes have degree \bar{k} .

3.2 Random Networks

For our purposes, a random network¹¹ will be constructed with the following algorithm: first we set the number of nodes equal to N , forming an adjacency matrix $A_{N \times N}$ with all entries zero. Since nodes cannot link to themselves, and because we consider undirected graphs, we only consider the $N(N - 1)/2$ elements off the main diagonal of A . Then we connect each of these among themselves with constant *linking probability* $p \in (0, 1]$; if i and j are linked, $a(i, j) = 1$ and $a(j, i) = 1$. An example of a random network is shown in the right panel of Figure 3.1. The resulting degree

¹⁰The circle is also called a one-dimensional lattice with periodic boundary conditions, and the number of neighbors is generally not restricted to the nearest ones, yet it is constant for each node.

¹¹The review article by Newman (2003) provides the historical background and a comprehensive summary of the many mathematical details of graph structures that we utilize here.

distribution is binomial with parameters N and p , and in the limit of large N it becomes Poissonian. Notice that the *average degree*, i.e. the average connectivity, will be Np .

3.3 Small-World Networks

Watts and Strogatz (1998) first proposed a model of small-world networks that is based on the idea that geographical proximity plays an important role in the formation of social networks. A well-known feature of observed social networks is that they show a high degree of *clustering*, yet a relatively small diameter.¹² There are various ways of constructing small-world networks, but the basic mechanism is always centered around the idea that one starts from a regular structure and then, with some probability p , one “rewires” nodes that are chosen randomly (for details, see e.g. Sec.VI of Newman, 2003). The rewiring probability p has to be chosen such that there is a low density of “shortcuts” in order to observe a high degree of clustering and a small diameter. If p is close to unity, the resulting network will closely resemble a random network. In our case, we simply add a small number of randomly placed shortcuts to the circle,¹³ illustrated in the central panel of Figure 3.1. Such a procedure implies that the degree distribution no longer degenerates to \bar{k} , but it will obviously be sharply peaked around \bar{k} .

3.4 Scale-Free Networks

Scale-free networks, i.e. networks with a power-law degree distribution, have been found in many social contexts, e.g. the citation network among scientific papers (Price, 1965; Redner, 1998), the World Wide Web and the Internet (see, e.g., Albert et al., 1999; Faloutsos et al., 1999), telephone call and e-mail graphs (Aiello et al., 2002; Ebel et al., 2002), or in the network of human sexual contacts (Liljeros et al., 2001).

¹²There are various clustering measures, which we are not going to detail here; intuitively speaking, a high degree of clustering (or *transitivity*) describes a situation in which many of i 's neighbors are connected among themselves. Put differently, if A is connected to B , and B is connected to C , there is an increased probability that A and C are connected as well. The small diameter property has been popularized through J. Guare's play that coined the now classic phrase of humankind's “six degrees of separation.”

¹³While enlarging the system, we keep the number of shortcuts S constant relative to the number of nodes, $s = S/N$. We arbitrarily chose $s = 0.1$.

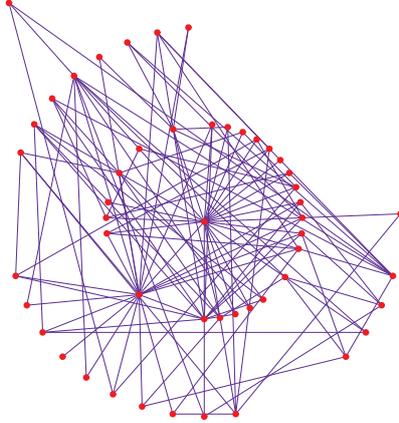


Figure 2: A “scale-free” graph with $N = 50$ nodes, constructed according to the preferential attachment algorithm of Barabási and Albert. Notice the presence of so-called *hubs* that is characteristic of scale-free networks.

A popular mechanism for the generation of scale-free networks is the *preferential attachment* of Barabási and Albert (1999), where one starts with m_0 initial nodes and then keeps adding new nodes with degree $m \leq m_0$; each of the m links is in turn connected to the existing nodes with a probability that is proportional to the degree of the already existing nodes, which is the reason why the mechanism is sometimes referred to as “the rich getting richer.” We can approximate the average connectivity in the scale-free network using a well-known result of Barabási and Albert (1999), who show that the *characteristic exponent* γ of the (cumulative)¹⁴ power-law degree distribution obeys $\gamma = 2$, irrespective of the choice of m and m_0 in the generating mechanism. Let $p(x)$ denote the power-law density of the degree distribution; then we have $p(x) = c x^{-\gamma-1}$ and the normalizing constant on the support $[m_0, N]$ is given by $c = \gamma m^\gamma / (1 - (m/N)^\gamma)$. Since $\gamma = 2$, the average degree will be $\int_m^N c x^{-2} dx = 2m(1 - (m/N)) / (1 - (m/N)^2)$, which for large N yields an average connectivity of $2m$.¹⁵

¹⁴Notice that a cumulative power-law distribution with exponent $\gamma = 2$ implies a power-law *density* with exponent $\gamma + 1 = 3$.

¹⁵Alternatively, one could also compute the average connectivity by realizing that the mechanism produces a total of $2(m_0 - 1) + 2(N - m_0)m$ links. The first term represents the

3.5 Whither N -Dependence?

Now that we have determined the average connectivity in the various network structures, we are finally in a position to make theoretical predictions about the ability of these networks to overcome the problem of N -dependence. According to the mean-field approximation for the variance of z in (17), the variance will only be independent of system size if the neighborhood D increases on the same order of magnitude as system size. But the only network that provides such connectivity is the random graph because its average connectivity is proportional to N . Actually, any network structure that keeps the *relative communication range* D/N constant under an enlargement of system size would overcome the problem of N -dependence, but among the prototypical structures that we consider it is only the random network that accomplishes this feat *within its generating mechanism*, i.e. when keeping the parameters of the respective generating mechanism fixed under an enlargement of system size. Hence we would expect that the regular, small-world, and scale-free networks cannot overcome the problem of N -dependence.¹⁶ That, however, would imply that network characteristics other than the mean degree—like the diameter, the average path length, various clustering measures, or any other network features one could think of—have no influence on the model’s ability to replicate the stylized facts independently of system size.

4 Model Implementation on Networks

Once we want to implement the stochastic herding model on various network structures, the immediate question becomes how to formulate microscopic transition probabilities for individual agents in the network that are consistent with the mesoscopic transition rates (14) and (15). In particular, we have to fix a time scale, to ensure that the transition probabilities are between zero and unity, and we have to choose values for the idiosyncratic parameter a and the herding intensity λ .

contribution of links from the starting m_0 nodes, while the factor of two shows up because we consider bi-directional networks, also showing that average connectivity is on the order of $2m$ for large N .

¹⁶The theoretical prediction that a scale-free graph is not capable of overcoming the problem of N -dependence is not entirely trivial, because some nodes (the “hubs”) have a connectivity that is on the same order of magnitude as system size.

The mean-field transition rates (14) and (15) describe the mesoscopic evolution of the system in the sense that during a time interval of length Δt we observe at most one switch between the two states, including the possibility of observing no switch on an aggregate level. This implies that all agents act on the same time scale, excluding behavioral heterogeneity in investment horizons among individual agents by assumption. Moving from mesoscopic transition rates per unit time Δt to microscopic transition probabilities therefore necessitates a choice of Δt that preserves this assumption.

At the individual level, contrary to the mesoscopic description of the system, an agent can either stay in its current state, or switch to the other state. An obvious way to implement individual transition probabilities, analogously to the transition rates (8) and (9), would be to posit the transition probability $p_i = (a + \lambda n(i, j))\Delta t$ for switching states on the individual level, where $n(i, j)$ counts the number of i 's neighbors that are in the opposite state. Since we need to ensure that $\forall i \ 0 \leq p_i \leq 1$, and that all agents act on the same time scale, this implies that $\Delta t \leq 1/(a + \lambda n_{max}(i, j))$, where $n_{max}(i, j)$ designates the number of neighbors of the node(s) with the highest degree in the network. Since an agent can be connected at most to all other agents, we simply implemented the following transition probability for switching states on the individual level

$$p_i = \frac{a + \lambda n(i, j)}{a + \lambda N}, \quad (18)$$

and correspondingly the probability to remain in the current state is $0 \leq 1 - p_i \leq 1$. We keep track of the microscopic configuration of the system by assigning a numerical value of unity if an agent is in state X , and a value of zero if it is in state Y . By convention, z will measure the concentration of agents in state X . For a given system size and network structure, we sequentially update the state of each agent according to the transition probabilities (18). One ‘‘sweep’’ of the system then corresponds to one round of sequential updating of all agents in the system, thus requiring N steps per sweep.

A straightforward way to assess the accuracy of the mean-field approximation for the various network structures is to compare the simulated empirical density of z with the theoretically predicted density given in (16). First, we constructed regular networks with $D = 10$ neighbors per node,

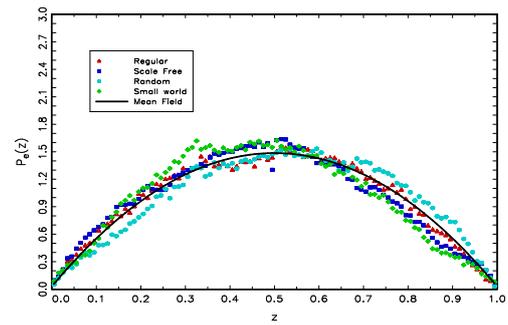
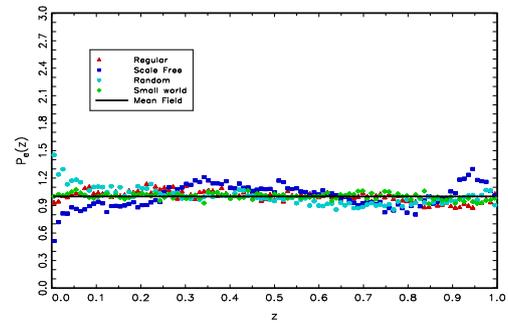
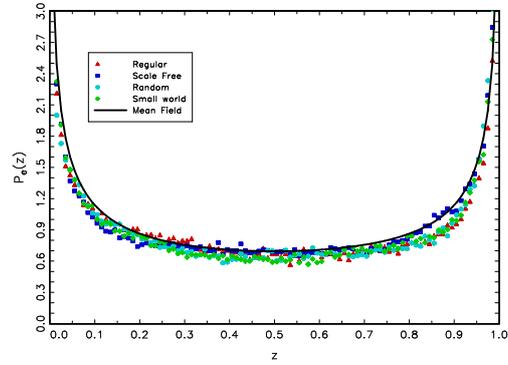


Figure 3: The empirical densities of the intensive variable $z = n/N$, which measures the concentration of agents in one of the two possible states, are in agreement with the mean-field prediction of a symmetric beta distribution $B(z; \epsilon, \epsilon)$, irrespective of the underlying network structure. We parameterized the simulations in such a way that the values of ϵ (from top to bottom) are 0.5, 1, and 2, respectively.

which we also use as a starting point for constructing small-world networks. In the generating mechanism for scale-free networks, we chose $2m = 10 = D$ for consistency. The random networks were constructed with a linking probability of $p = 0.1$. Then we simulated the model with these network parameterizations, and set the behavioral parameters a, λ in the microscopic transition rates in a way that should respectively yield bimodal, uniform, and unimodal Beta densities if the mean-field approximation provides a good description of the generic herding model.¹⁷ Casual inspection of Figure 3, where we plotted the empirical density after 100,000 sweeps for each of the network structures and the three parameterizations of $\epsilon = aN/\lambda D$, indeed confirms that the mean-field approximation performs very well for the different network structures.

A central issue of our investigation concerns the mean-field prediction that regular, small-world, and scale-free networks are not capable of overcoming the problem of N -dependence. In order to check the validity of the mean-field prediction, we start with $N = 50$ agents and keep enlarging the respective network with a step-size of $\Delta N = 500$ for each of the four network structures. Without loss of generality, we now parameterize the model with $a = 0.5$ and $\lambda = 1$ in the transitions (18). We compute the variance of z after 300,000 sweeps for each system size, and report the results in Figure 4. Given our parameterization of the average connectivity $D = 2m = 10$ for the regular, small-world, and scale-free structures, the mean-field approach predicts that $\text{Var}[z] = 1/(0.4N + 4)$. While the predictions of the mean-field approximation perform reasonably well for all network structures, we detect a slight deviation for the regular and small-world structures when plotting the inverse variance of z , shown in the inset of Figure 4. The inverse variance scales linearly none the less, and ordinary least-squares regressions yield a slope of 0.512 (standard error 0.031) for the regular network, and a slope of 0.507 (standard error 0.046) for the small-world network. The estimated slope for the variance of the scale-free network is 0.402 (standard error 0.022), consistent with the predicted slope. An OLS fitting of the random network data returns a slope that is not significantly different from zero, and the intercept is 46.81 (standard deviation 9.07), in line with the theoretical prediction of zero slope and an intercept

¹⁷Notice that there is nothing peculiar about the parameterizations we chose; we tried several other parameterizations that returned the same distributional characteristics.

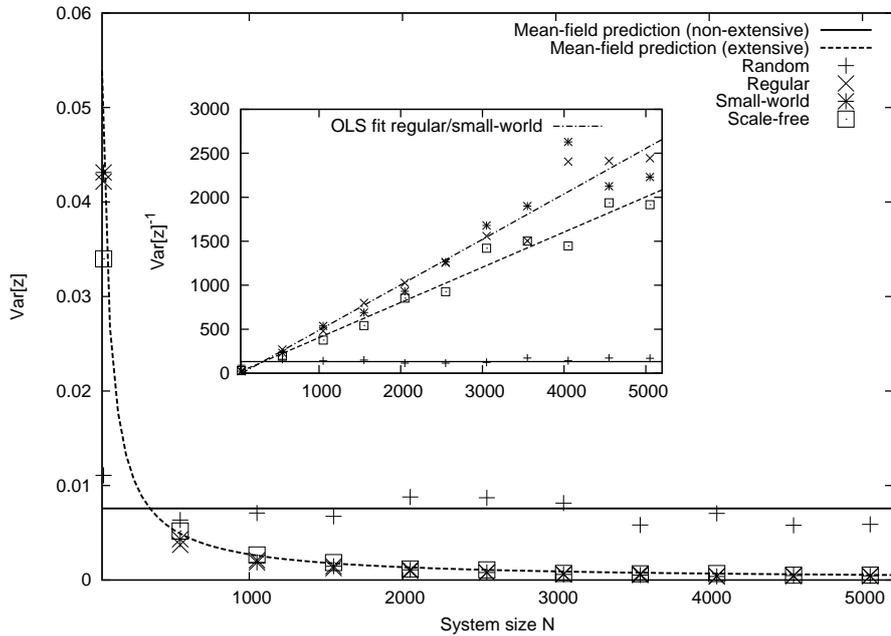


Figure 4: Simulation of the probabilistic herding model on various network structures shows that the variance of z is in line with the theoretical predictions of the mean-field approximation. The inset shows the inverse variance of z , which reveals a slight deviation from the mean-field approximation for regular and small-world networks, where we observe a steeper slope than the predicted one. Only the random network provides a variance of z that is constant, and is therefore able to preserve the stylized facts of financial returns in the herding model for any system size.

of 44 for the inverse variance of z .¹⁸

On the one hand, our simulations confirm that the mean-field approach provides a useful description of the probabilistic herding model, which is encouraging. On the other hand, however, the simulations also confirm that for constant behavioral parameters a and λ , the relative communication range D/N is really all that matters, and consequently the random network would appear to be the only structure capable of overcoming the problem of N -dependence in the herding model.

¹⁸For convenient plotting, we multiplied the variance of z by one over three, and the inverse variance by a factor of three in Figure 4.

5 Individual Heterogeneity

In the microscopic derivation of the transition rates (14) and (15), we assumed that agents exhibit the same sensitivity to news arrivals and an identical herding propensity. Intuitively, the mean-field approach would seem to allow for individual heterogeneity in the behavioral parameters, simply by replacing a, λ with a_i, λ_i in the microscopic transition rates (8) and (9) and, following the logic of the mean-field approach, with their respective mean values $\langle a \rangle, \langle \lambda \rangle$ in (14) and (15). Correspondingly, the model implementation with heterogeneous agents leads to the transition probabilities

$$p_i = \frac{a_i + \lambda_i n(i, j)}{a_{max} + \lambda_{max} N}, \quad (19)$$

which we simulated with 100,000 sweeps on random networks of size $N = 1000$. First, we varied a_i and λ_i by drawing randomly from Gaussian distributions with positive means. The resulting equilibrium distributions of z exhibited the same features as the ones shown in Figure 3. Second, in order to check if the mean-field would allow for more heterogeneity than in the case of a (sharply) peaked distribution, we chose a, λ from uniform distributions with respective supports $[\underline{a}, \bar{a}]$ and $[\underline{\lambda}, \bar{\lambda}]$, where we detected deviations from the predicted equilibrium distributions. Drawing a from a uniform with $\underline{a} = 0$ while keeping λ fixed, on the other hand, produced distributions that were in agreement with the mean-field prediction. Then we checked whether the presence of agents that are not herding, i.e. $\lambda_i = 0$ for some i , is responsible for the observed deviations by drawing λ_i uniformly at random with strictly positive $\underline{\lambda}$. The simulations confirmed that the mean-field approximation again predicts the equilibrium distribution of the heterogeneous agent system *as long as all agents have a strictly positive herding propensity*.

The above findings suggest that the herding model is fundamentally influenced by the presence of agents that are not herding. On a purely topological level, this corresponds to a situation where certain links in the network become *unidirectional*, as illustrated in Figure 5. The fact that already a small fraction of independently acting agents leads to a large reduction in the variance of z is rather surprising, and casts doubts on the robustness of the model against this peculiar scenario. Basically, there are



Figure 5: Change of network topology in the presence of agents that are not herding. The left panel is a stylized representation of the bidirectional nature of interaction in the regular herding model, where each agents considers the state of its neighbors. The right panel illustrates that the presence of agents that are not prone to herding ($\lambda_i = 0$), designated by I in the graph, leads to an unidirectional flow of information. Because the independently acting agent does not care about the states of its neighbors any longer, the information transmission among herding agents becomes more difficult, or even impossible.

two possibilities why the simulated variance of z could differ substantially from the mean-field prediction. It could either be due to the presence of independently acting agents itself, or it could be due to changes in network topology. In order to disentangle and quantify both effects on the variance of z , we start by considering a benchmark case in which N agents with a strictly positive herding propensity are connected through a random network, and M agents are acting independently at random *outside* of the network. If the reduction in the variance of z is the same for the benchmark case and the case where independently acting agents are still part of the network, we would conclude that the variance reduction is simply due to the presence of independent agents.

Let $q = M/K$ denote the fraction of independently acting agents in the system, where $K = M + N$ stands for the total size of the system, and let $n + m = k$ denote the respective occupation numbers of agents in state X . Then the system-wide concentration of agents in the state is given by

$$z = \frac{k}{K} = \frac{N}{K} \frac{n}{N} + \frac{M}{K} \frac{m}{M} = (1 - q)z_H + qz_I. \quad (20)$$

Since the actions of the N agents in the *core* and the M agents in the *periphery* are independent of each other, the variance of z can simply be decomposed as,

$$\text{Var}[z] = (1 - q)^2 \text{Var}[z_H] + q^2 \text{Var}[z_I], \quad (21)$$

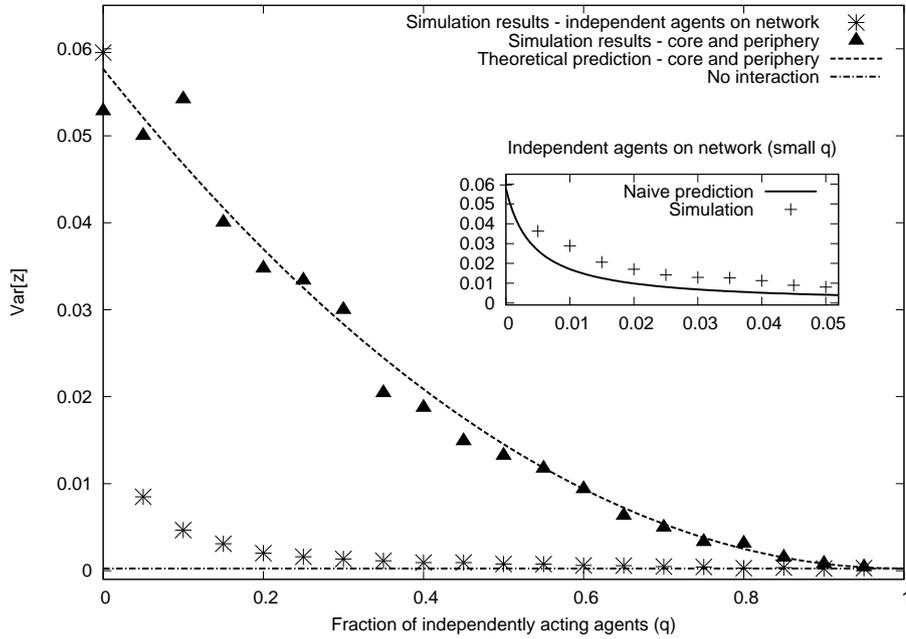


Figure 6: Simulation results in the presence of independently acting agents. In comparison to the core-periphery case, the variance reduction is much larger when independently acting agents are part of the network. The variance reduction is already very pronounced for a small fraction of independent agents on the network, illustrated in more detail in the inset.

where $Var[z_H]$ has already been established in (17) and $Var[z_I]$ is the variance of the noisy periphery,¹⁹ yielding

$$Var[z] = (1 - q)^2 \left(\frac{8a}{\lambda p} + 4 \right)^{-1} + \frac{q}{4K}. \quad (22)$$

Numerical simulations, averaged over 100,000 sweeps for each q , confirm that (22) is a valid description of the benchmark case, shown in Figure 6. In the second case, the presence of independent agents on the network influences the transition rates of agents that have a strictly positive herding propensity. Although an individual agent is not aware of whether its neighbors are herding or acting independently, we can employ exactly the same

¹⁹Since the M agents in the periphery are independent, and with equal probability in state 0 or 1, we have $E[m] = E[\sum_{i=1}^M s_i] = M/2$, and $Var[m] = E[m^2] - E[m]^2 = ME[m^2] - M(M-1)E[m]^2 - M^2/4 = M/4$. Therefore $Var[z_I] = Var[m]/M^2 = 1/4M$.

mean-field reasoning as in Section 2.3, such that (13) now transforms to

$$\langle \pi_i^+ \rangle = a + b \left(D_I \frac{m}{M} + D_H \frac{n}{N} \right), \quad (23)$$

where $D_I = qpK$ and $D_H = (1 - q)pK$ are the respective average numbers of neighbors that are herding or acting independently.²⁰ We know from the *Ehrenfest model* that the concentration of independently acting agents will fluctuate around one half; summing over the agents that are prone to herding therefore leaves us with the modified transition rates

$$\pi^+ = (N - n) [A + \lambda p n], \quad (24)$$

$$\pi^- = n [A + \lambda p (N - n)], \quad (25)$$

showing that the impact of independently acting agents in the network can be absorbed into the idiosyncratic parameter of the transition rates,

$$A = a + \frac{\lambda qpK}{2},$$

where the factor of 1/2 comes from the Ehrenfest model. Since the functional form of the modified transition rates (24) and (25) remains the same as in (14) and (15), the equilibrium distribution of z_H still corresponds to (16), while the parameter of the distribution now obeys

$$\epsilon = \frac{a}{\lambda p} + \frac{qK}{2}, \quad (26)$$

and correspondingly the variance of z_H is

$$\text{Var}[z_H] = \left(\frac{8a}{\lambda p} + 4 + 4qK \right)^{-1}. \quad (27)$$

Unfortunately, the variance of $z = (1 - q)z_H + qz_I$ can no longer be calculated according to (21) because z_H is now correlated with z_I , so $\text{Var}[z] = (1 - q)^2 \text{Var}[z_H] + q^2 \text{Var}[z_I] + 2(1 - q)q(E[z_H z_I] - E[z_H]E[z_I])$, and it is not obvious how to determine $E[z_H z_I]$ analytically. In a rough first approximation for small q , we could neglect the positive correlation between z_H and z_I , i.e. $\text{Var}[z | q \ll 1] \approx \text{Var}[z_H]$, and the inset of Figure 6 shows that

²⁰From the conceptual point of view of the mean-field approximation, we could equally well replace a and λ with $\langle a \rangle$ and $\langle \lambda \rangle$, which we refrain from for notational simplicity.

the pronounced variance reduction is reasonably well captured by (27), at least qualitatively, for small fractions of independently acting agents on the network.²¹

Therefore, the major difference between the two cases concerns the scale in the reduction of the system-wide variance. If K is large, the presence of only a few independently acting agents on the network is already sufficient to considerably decrease the extent of variations in the aggregate state of the system.²² Conversely, from our point of view, it is a non-trivial feature of the benchmark case that the variance of z will be dominated by a relatively small fraction of agents that are herding. Letting $c = 4 + 8a/\lambda p$, we see from (22) that the variance of z will be dominated by the contribution of agents prone to herding if $N > \sqrt{qKc}/2 = \sqrt{Mc}/2$. So suppose that q is close to unity and that K is on the order of 10^6 agents. Then N on the order of 10^3 already ensures that the major contribution to the variance of the system comes from the core of agents that are herding.²³ Hence it obviously makes a big difference for the aggregate properties of the herding model whether independently acting agents are part of the network or not.

If we consider both of the above cases in terms of a core (of agents that are prone to herding) and a periphery (of agents that are acting independently), the topological interpretation of the case where independent agents are part of the network corresponds to unidirectional links from the periphery to the core. But what about the opposite case, where unidirectional links emanate from the core to the periphery, i.e. when the agents in the periphery are not acting independently at all ($a_j = 0$ for all $j = 1, \dots, M$ agents in the periphery), but instead are herding with certainty on the state of an agent that they are linked to in the core? The simplest case would be to imagine that each agent in the core is linked to a constant number F of

²¹The observed deviations should be due to the positive correlation between z_H and z_I . We removed fluctuations in the states of independent agents by fixing half of them in one of the two states and half in the other, thereby preserving the mean value of the case with fluctuations, yet removing the correlation in z_H and z_I . Our simulations confirmed that the results correspond exactly to the naive prediction in this case.

²²We parameterized our simulations with $a = 0.5, \lambda = 1, p = 0.3$ and $K = 1000$ agents, and observe a reduction of the variance of z by a factor of two for $q=1\%$ and by a factor of seven for $q=5\%$.

²³Sticking with this numerical example, let us further suppose that $p \approx 0.1$; then the one thousand agents in the core would need to be connected on average to one hundred neighbors (which is not an entirely unrealistic scenario in some professional contexts) in order to dominate the variance of a system that is composed of a million agents.

followers such that $FN = M$, with a total of $M + N$ agents in the system. Then the concentration of agents in state X is

$$z = \frac{Fn + n}{M + N} = \frac{n(F + 1)}{N(F + 1)} = z_H, \quad (28)$$

which amounts to a mere relabelling of variables. Nevertheless, such a simple *hierarchical* structure that consists of opinion leaders on the one hand, who perpetually process new information but are prone to herding among themselves, and a possibly huge number of followers on the other, is capable of overcoming the problem of N -dependence for any parameterization of and any network structure in the core. This opens up new venues for future research in the direction of hierarchical network structures, for instance by considering more general cases in which opinion leaders are not connected to a constant number of followers, but instead to an essentially random number drawn from arbitrary distributions. Another potentially interesting extension in this line of research would be to consider several hierarchical “layers” instead of just two, allowing for different degrees of behavioral heterogeneity in each layer.

Although the assumption that each opinion leader has the same number of followers is overly simplistic, the idea of a hierarchical structure in opinion formation might actually not be too far off the mark in light of some recent empirical evidence on herd behavior among fund managers: Hong et al. (2005) document that money managers who work in close geographical proximity are prone to what they term “word-of-mouth” effects. In our simplistic scenario, the money managers would correspond to agents in the core, while individuals who invest in the various funds would represent followers in the periphery.

6 Conclusions

The probabilistic herding model that underlies our investigation incorporates individual heterogeneity as well as a tendency for social interaction, but it is not immune to an enlargement of system size, which we labeled as the N -dependence effect. Our derivation of microscopic foundations for the herding model reflects favorably on the mean-field approach as a formal way of economic modelling in the presence of individual hetero-

geneity, and it illustrates the crucial impact that structural heterogeneity has on the distributional outcome of the model. Yet we have shown that the random network, which is hardly a realistic way of organizing economic relationships, is the only prototypical network structure capable of overcoming the problem of N -dependence if we restrict the description of network characteristics to homogeneous links and nodes.

The structurally more heterogeneous and rather diametrically opposed cases that we encountered, where some agents are not herding at all, or where a vast majority of simpletons follows their respective opinion leaders, suggest that structural heterogeneity matters a great deal for the macroscopic properties of the probabilistic herding model. While we have demonstrated that the presence of very few independently acting agents in the network is sufficient to destroy correlations in the model, we have also hinted at the possibility that a hierarchical structure could overcome the problem of N -dependence irrespective of the network topology that connects the core of opinion leaders. It does not appear entirely unreasonable to us at this point that hierarchical principles of organization in complex market economies, such as the pervasive presence of institutional relationships and constraints, might play an important role in explaining how the problem of N -dependence could be overcome in a socially relevant and economically plausible network structure.

Finally, from a complex systems perspective, the probabilistic herding model warrants some interest beyond the context of financial markets. According to the central limit theorem, the sum of independent (or weakly dependent) random variables with finite means and variances will converge to a Gaussian distribution. Yet a plethora of social and economic phenomena are not distributed normally but instead show other distributional regularities that are robust in space and time. The power-law distribution of financial returns, which originally motivated the generic herding model that we have studied in the present paper, is just one economic example that comes to mind.²⁴ Since it is reasonable to exclude infinite means or variances in most descriptions of the constituent units of an economic system, we believe that correlations among these units could explain the observed

²⁴Other economic phenomena that do not exhibit Gaussian characteristics include the distribution of firm and city sizes, the distribution of personal income and wealth, and the distribution of firm growth rates and growth rates in national income, which have been established in the economic literature at different points during the past century.

deviations from normality, and it is not a far stretch of the imagination to consider agent interaction as the source of correlations in a social context. But the probabilistic herding model illustrates very clearly that interaction by itself is not sufficient to overcome the problem of N -dependence, i.e. the Gaussian distributional properties stemming from the central limit theorem. If we consider our results from this more general perspective, the question becomes whether structural factors in a network of agents could be capable of providing enough interaction to break away from the distributional implications of the central limit theorem.

References

- Aiello, W., F. Chung, and L. Lu. 2002. Random evolution of massive graphs. In *Handbook of Massive Data Sets*, eds. J. Abello, P. M. Pardalos, and M. G. C. Resende, 97–122. Kluwer: Dordrecht.
- Albert, R., H. Jeong, and A.-L. Barabási. 1999. Diameter of the world-wide web. *Nature* 401:130–131.
- Alfarano, S., T. Lux, and F. Wagner. 2005. Estimation of agent-based models: the case of an asymmetric herding model. *Computational Economics* 26:19–49.
- . 2006. Time-variation of higher moments in a financial market with heterogeneous agents: An analytical approach. Forthcoming in *Journal of Economic Dynamics and Control*.
- Aoki, M. 1998. *New Approaches to Macroeconomic Modeling*. Cambridge, UK: Cambridge University Press.
- Barabási, A.-L., and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286:509–512.
- Bornholdt, S. 2001. Expectation bubbles in a spin model of markets: Intermittency from frustration across scales. *International Journal of Modern Physics C* 12:667–674.
- Brock, W. A., and C. H. Hommes. 1997. A rational route to randomness. *Econometrica* 65:1059–1095.

- Cont, R., and J. P. Bouchaud. 2000. Herd behaviour and aggregate fluctuations in financial markets. *Macroeconomic Dynamics* 4:170–196.
- Ding, Z., R. Engle, and C. Granger. 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1:83–106.
- Ebel, H., L.-I. Mielsch, and S. Bornholdt. 2002. Scale-free topology of e-mail networks. *Physical Review E* 66. 035103.
- Egenter, E., T. Lux, and D. Stauffer. 1999. Finite-size effects in Monte Carlo simulations of two stock market models. *Physica A*, 268:250–256.
- Faloutsos, M., P. Faloutsos, and C. Faloutsos. 1999. On power-law relationships of the internet topology. *Computer Communications Review* 29:251–262.
- Gopikrishnan, P., M. Meyer, L. A. N. Amaral, and H. E. Stanley. 1998. Inverse cubic law for the distribution of stock price variations. *European Physical Journal B* 3:139–140.
- Hong, H., J. D. Kubik, and J. C. Stein. 2005. Thy neighbor’s portfolio: Word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance* 60:2801–2824.
- Iori, G. 2002. A micro-simulation of traders’ activity in the stock market: the role of heterogeneity, agents’ interactions and trade friction. *Journal of Economic Behavior and Organization* 49:269–285.
- Jackson, M. O. 2004. A survey of models of network formation: Stability and efficiency. In *Group Formation in Economics; Networks, Clubs and Coalitions*, eds. G. Demange and M. Wooders. Cambridge, UK: Cambridge University Press.
- Kirman, A. 1991. Epidemics of opinion and speculative bubbles in financial markets. In *Money and Financial Markets*, ed. M. P. Taylor, 354–368. Cambridge: Blackwell.
- . 1993. Ants, rationality, and recruitment. *Quarterly Journal of Economics* 108:137–156.
- Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg. 2001. The web of human sexual contacts. *Nature* 411:907–908.

- Lobato, I. N., and N. E. Savin. 1998. Real and spurious long-memory properties of stock market data. *Journal of Business and Economic Statistics* 16:261–283.
- Lux, T. 2006. Financial power laws: Empirical evidence, models, and mechanisms. In *Power Laws in the Social Sciences: Discovering Complexity and Non-Equilibrium Dynamics in the Social Universe*, ed. Claudio Cioffi. Cambridge, UK: Cambridge University Press. Forthcoming.
- Lux, T., and M. Ausloos. 2002. Market fluctuations I: Scaling, multiscaling and their possible origins. In *Theories of Disaster - Scaling Laws Governing Weather, Body, and Stock Market Dynamics*, eds. A. Bunde, J. Kropp, and H. J. Schellnhuber, 373–409. Berlin Heidelberg: Springer.
- Lux, T., and M. Marchesi. 1999. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* 397:498–500.
- Lux, T., and S. Schornstein. 2005. Genetic learning as an explanation of stylized facts of foreign exchange markets. *Journal of Mathematical Economics* 41:169–196.
- Newman, M.E.J. 2003. The structure and function of complex networks. *SIAM Review* 45:167–256.
- Pagan, A. 1996. The econometrics of financial markets. *Journal of Empirical Finance* 3:15–102.
- Price, D. J. de S. 1965. Networks of scientific papers. *Science* 149:510–515.
- Redner, S. 1998. How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B* 4:131–134.
- Shiller, R.J. 1984. Stock prices and social dynamics. *Brookings Papers on Economic Activity* 2:457–510.
- Stauffer, D., and A. Aharony. 1994. *Introduction to Percolation Theory*. London: CRC Press.
- Stauffer, D., and D. Sornette. 1999. Self-organized percolation model for stock market fluctuation. *Physica A* 271:496–506.

Van Kampen, N. G. 1997. *Stochastic Processes in Physics and Chemistry*. Amsterdam: Elsevier.

Watts, D. J., and S. H. Strogatz. 1998. Collective dynamics of ‘small-world’ networks. *Nature* 393:440–442.

Appendix

The derivation of the Fokker-Planck equation associated with the transition rates (14) and (15) follows the exposition in Alfarano et al. (2006), who consider the intensive variable $\eta = 2n/N - 1$. We focus on the concentration of agents in state X , i.e. $z = n/N$, that we have used throughout the text; since both, η and z are intensive variables, the results will be qualitatively the same. Assuming that N is large enough, the intensive variable z can be treated as a continuous quantity. Given the mesoscopic transition rates (3) and (4), the mean-field approximation in (14) and (15) implies that $b = \lambda D/N$, which we will use for notational convenience. The relation between the transition rates for n and z is

$$\pi_n^\pm = N^2 \pi_z^\pm,$$

where the transition rates (3) and (4) are now functions of z ,²⁵

$$\pi_z^+ = \pi(z \rightarrow z + 1/N) = (1 - z) \left(\frac{a}{N} + bz \right), \quad (\text{A1})$$

$$\pi_z^- = \pi(z \rightarrow z - 1/N) = z \left(\frac{a}{N} + b(1 - z) \right). \quad (\text{A2})$$

The transition probabilities associated with (3) and (4), or (A1) and (A2), are $\omega^\pm = \pi^\pm \Delta t$; they define a Markovian stochastic process that belongs to the class of so-called non-linear ‘‘one-step’’ processes. The probabilities $P(n, t)$ to have n agents in state X at time t obey the *Master equation*, which gives the probability flux between states

$$\frac{\partial}{\partial t} P(n, t) = \sum_m [\pi_{n,m} P(m, t) - \pi_{m,n} P(n, t)], \quad (\text{A3})$$

²⁵Since we are merely relabeling events, the probabilities are invariant under a transformation from n to z , i.e. $P(n) = P(z)$.

with $\pi_{n,m} = \frac{\omega_{n,m}}{\Delta t}$ designating the transition rates from states m to states n . In the limit $\Delta t \rightarrow 0$, multiple jumps occur with probability zero, so that one only has to consider jumps to neighboring states, i.e. $m = n \pm 1$. Replacing n by the intensive variable z , and switching from discrete probabilities to the probability density defined as the limit $p(z, t) = \lim_{\Delta z \rightarrow 0} P(z, t) / \Delta z$, we can derive a Fokker-Planck equation for the time change of the probability density $p(z, t)$. Adopting the notation of Van Kampen (1997), we introduce the “step operators” \mathbb{E} and \mathbb{E}^{-1} that act on an arbitrary function $f(n)$, respectively increasing or decreasing its integer argument n by one unit, i.e. $\mathbb{E}[f(n)] = f(n + 1)$ and $\mathbb{E}^{-1}[f(n)] = f(n - 1)$. With the aid of these operators, the Master equation for the one-step process can be compactly rewritten in terms of the variable z ,

$$\frac{\partial}{\partial t} p(z) = N^2 \left[(\mathbb{E} - \mathbf{1})[\pi_z^- p(z)] + (\mathbb{E}^{-1} - \mathbf{1})[\pi_z^+ p(z)] \right], \quad (\text{A4})$$

where we omitted the obvious time dependence of $p(z)$. Since the step operator acts on continuous functions in the limiting case, we can use the Taylor expansion

$$\mathbb{E}[f(z)] = f(z + \Delta z) = f(z) + \Delta z \frac{d}{dz} f(z) + \frac{1}{2} \Delta z^2 \frac{d^2}{dz^2} f(z) + o(\Delta z^2), \quad (\text{A5})$$

where $\Delta z = 1/N$. Thus the step operator can be approximated by

$$\mathbb{E} = \mathbf{1} + \Delta z \frac{\partial}{\partial z} + \frac{1}{2} \Delta z^2 \frac{\partial^2}{\partial z^2} + o(\Delta z^2). \quad (\text{A6})$$

The expansion for the operator \mathbb{E}^{-1} is obtained from the previous formula by replacing Δz with $-\Delta z$. Using the expansion (A6) for \mathbb{E} and \mathbb{E}^{-1} up to the second order, we end up with

$$\frac{\partial}{\partial t} p(z) = N^2 \left(-\Delta z \frac{\partial}{\partial z} [(\pi_z^+ - \pi_z^-) p(z)] + \frac{1}{2} \Delta z^2 \frac{\partial^2}{\partial z^2} [(\pi_z^+ + \pi_z^-) p(z)] \right).$$

The factor N^2 cancels out, and we obtain the Fokker-Planck equation

$$\frac{\partial}{\partial t} p(z) = -\frac{\partial}{\partial z} A(z) p(z) + \frac{1}{2} \frac{\partial^2}{\partial z^2} D(z) p(z), \quad (\text{A7})$$

where the *drift* is

$$A(z) = N (\pi_z^+ - \pi_z^-) = a - 2az = 2a \left(\frac{1}{2} - z \right), \quad (\text{A8})$$

and the *diffusion* is

$$D(z) = \pi_z^+ + \pi_z^- = 2(1-z)bz + \frac{a}{N}. \quad (\text{A9})$$

To determine the equilibrium distribution $p_e(z)$ that obeys $\frac{\partial}{\partial t} p(z) = 0$, we employ the well-known formula (see, e.g., Van Kampen, 1997)

$$p_e(z) = \frac{k}{D(z)} \exp \left(\int^z \frac{2A(z')}{D(z')} dz' \right). \quad (\text{A10})$$

Letting $\epsilon = a/b$, and neglecting the “granular” term a/N in the diffusion under the assumption that N is sufficiently large, the solution of the integral is $\epsilon \ln(z - z^2)$, and therefore the equilibrium distribution is $p_e(z) = k[(z(1-z))^{\epsilon-1}]$. The constant follows from the normalization of the density,

$$\frac{1}{k} = \int_0^1 z^{\epsilon-1} (1-z)^{\epsilon-1} dz = B(\epsilon, \epsilon),$$

where $B(\epsilon, \epsilon)$ is Euler’s beta function. Hence we see that the equilibrium distribution is given by equation (16) in the main text.