Kiel

Working Papers

Kiel Institute
for the World Economy

Switching Rates and the Asymptotic Behavior of Herding Models

by Albrecht Irle, Jonas Kauschke, Thomas Lux, and Mishael Milaković

No. 1595 | February 2010

Web: www.ifw-kiel.de

Kiel Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany

Kiel Working Paper 1595 | February 2010

Switching Rates and the Asymptotic Behavior of Herding Models

Albrecht Irle, Jonas Kauschke, Thomas Lux, and Mishael Milaković

Abstract:
Markov chains have experienced a surge of economic interest in the form of behavioral agent-based models that aim at explaining the statistical regularities of financial returns. We review some of the relevant mathematical facts and show how they apply to agent-based herding models, with the particular goal of establishing their asymptotic behavior because several studies have pointed out that the ability of such models to reproduce the stylized facts hinges crucially on the size of the agent population (typically denoted by n), a phenomenon that is also known as n-dependence. Our main finding is that n-(in)dependence traces back to both the topology and the velocity of information transmission among heterogeneous financial agents.

Keywords: Markov chains, agent-based finance, herding, N-dependence

JEL classification: C10, D84, D85, G19

Albrecht Irle
Department of Mathematics
University of Kiel
Ludewig-Meyn-Straße 4
24098 Kiel, Germany
Phone: +49 431-880 4650
E-mail: irle@math.uni-kiel.de

Jonas Kauschke
Department of Mathematics
University of Kiel
Ludewig-Meyn-Straße 4
24098 Kiel, Germany
Phone: +49 431-880 4393
E-mail: kauschke@math.uni-kiel.de

Thomas Lux
Kiel Institute for the World Economy
24100 Kiel, Germany
Phone: +49 431-8814 278
E-mail: thomas.lux@ifw-kiel.de
E-mail: lux@bwl.uni-kiel.de

Mishael Milaković
Department of Economics
University of Kiel
Olshausenstraße 40
24118 Kiel, Germany
Phone: +49 431-880 3369
E-mail: milakovic@bwl.uni-kiel.de

Switching Rates and the Asymptotic Behavior of Herding Models*

Albrecht Irle
Department of Mathematics
University of Kiel, Germany

Jonas Kauschke
Department of Mathematics
University of Kiel, Germany

Thomas Lux
Department of Economics
University of Kiel, Germany
and Kiel Institute for the World Economy

Mishael Milaković
Department of Economics
University of Kiel, Germany

January 7, 2010

*We gratefully acknowledge financial support by the Volkswagen Foundation through their grant on "Complex Networks As Interdisciplinary Phenomena."

1 Introduction

Continuous-time Markov chains are a popular tool for stochastic modeling in a great variety of fields, ranging from population genetics to communication networks. The mathematical theory of such processes is well understood and provides powerful results for the handling of applied problems. An early economic application was Champernowne's model of income distribution [9] and more recently, inspired by Kirman's "ant model" [17], Markov chains have experienced a surge of economic interest in the form of behavioral agent-based models that aim at explaining the statistical regularities of financial returns. In this note, we shall review some of the relevant mathematical facts and show how they apply to these agent-based models, with the particular goal of establishing their asymptotic behavior, which has become an increasingly important concern in the agent-based literature.

The time series of financial returns reveal some ubiquitous statistical regularities across countries, time frequencies, and asset classes. Most notably, financial returns are leptokurtic and exhibit clustered volatility, and there are strong indications that both the distribution of large returns as well as the auto-correlation of transformations of returns are power-laws [10, 14, 19, 20, 21]. While traditional finance has paid very little attention to scaling laws and their economic origins, a series of stochastic behavioral asset pricing models with heterogeneous interacting agents have been able to account for the stylized facts [1, 2, 3, 13, 16, 18]. The unifying feature of these models is their conceptualization of agent interactions as a probabilistic herding process that is expressed through a pair of Markovian transition rates. Starting with [11], however, a number of studies have pointed out that the ability of such models to reproduce the stylized facts hinges crucially on the size of the agent population, typically denoted by N, a phenomenon that is also known as N-dependence.[1] In many cases the interesting properties of returns, namely their peculiar time-dependence structure and the power law decay of the return distribution, progressively disappear when N increases, instead leading to Gaussian price fluctuations and a weak degree of temporal dependence for an otherwise unchanged model parametrization. From a historical point of view such model behavior is also clearly unsatisfactory, because the collapse of the Bretton-Woods system in 1971 has led to more globalized financial markets with considerable increases in the number of agents and volumes, but certainly not to Gaussian returns.

The formal problem of N-dependence can be traced to an apparently minor modification in the transition rates that is independent of the behav-

[1] Aoki [7, 8] utilizes the terms *(non) self-averaging* in lieu of N-(in)dependence.

ioral parameters [4], which in turn has motivated research into the network structure that describes the microscopic feasibility of agent interaction [5]. Topological aspects of heterogeneous agent interaction have very recently also been taken up in a similar model of herding in financial markets [22]. The motivation of our paper is to demonstrate the formal criteria for N-(in)dependence in Kirman-type herding models, thereby providing a more rigorous mathematical treatment of the results in [4]. Our strategy will be to consider two benchmark cases, labeled Model 1 and Model 2, and to carefully show how their asymptotic behavior depends on the detailed parametrization of the Markovian herding model. The main finding is that N-(in)dependence traces back to both the topology and the velocity of information transmission among heterogeneous financial agents.

2 Basic facts on continuous-time Markov chains

2.1 Generalities

We consider a continuous-time Markov chain $(X_t)_{t\in[0,\infty)}$ with discrete state space S, finite or countably infinite, having right-continuous paths. The Markov property states that

$$P(X_{t+s} = j|(X_r)_{0\leq r\leq s}, X_s = i) = P(X_{t+s} = j|X_s = i)$$

for all $t, s \geq 0, i, j \in S$, i.e. the future development, given the past and present, only depends on the present state i and is time-homogeneous. $(P_{ij}(t))_{i,j}$ are stochastic matrices and are called the transition matrices.

The infinitesimal characteristics are given by

$$q_{ij} = \lim_{t\downarrow 0} \frac{P_{ij}(t)}{t}, j \neq i,\ q_{ii} = \lim_{t\downarrow 0} \frac{P_{ii}(t) - 1}{t},$$

and we assume $0 < q_i = -q_{ii} = \sum_{j\neq i} q_{ij} < \infty$ for all states. The q_{ij} are the transition rates and the matrix $Q = (q_{ij})_{i,j}$ is the generator. Q uniquely determines the distribution of the process and we have the matrix differential equation

$$P'(t) = P(t)Q,\ \text{resulting in}\ P(t) = e^{Qt}.$$

A stationary distribution $\pi = (\pi_i)_i$ fulfills the equation $\pi^T Q = 0$. To Q corresponds the infinitesimal generator A which, for any bounded mapping $f : S \to S$, defines a mapping $Af : S \to S$ by

$$Af(i) = \lim_{t\downarrow 0} \frac{\sum_{j\in S} P_{ij}(t)(f(j) - f(i))}{t} = \sum_{j\neq i} q_{ij}(f(j) - f(i)).$$

It determines Q uniquely, hence also the distribution of the process, while the transition rates fulfill

$$\begin{aligned} P(X_t = j | X_0 = i) &= q_{ij}t + o(t), i \neq j \\ P(X_t = i) | X_0 = i) &= 1 - q_i t + o(t). \end{aligned}$$

2.2 Probabilistic representation

Starting from Q we can realize the process in the following way. Let

$$p_{ij} = \frac{q_{ij}}{q_i}, i \neq j, \ p_{ii} = 0.$$

This is the transition matrix of a discrete-time Markov chain, the so-called embedded chain. When the chain is in state i, it stays there for a random holding time, exponentially distributed with parameter q_i, and then jumps to j according to p_{ij}. An equivalent description is the following. Assume $\sup_i q_i \leq q < \infty$ and define

$$r_{ij} = \frac{q_{ij}}{q}, i \neq j, \ r_{ii} = 1 - \sum_{j \neq i} r_{ij}.$$

When the chain is in state i, it stays there for a random holding time, exponentially distributed with parameter q, and then makes a transition to j according to r_{ij}, so in this second realization the process may stay put in i with probability r_{ii}. This representation immediately shows how such a process may be simulated with exponentially distributed random variables and discrete transitions from i to j. It also shows the different roles of q_i and $\frac{q_{ij}}{q_i}$, the first determining the rate at which transitions occur and the second the probabilities according to which the new state is selected.

2.3 Representation using a random time transformation

We now assume that S is a subset of $\mathbb{Z}^d$; setting transition rates outside of S equal to zero, S may be assumed to be equal to $\mathbb{Z}^d$. The basic building block is the Poisson process which remains in any state $i = 0, 1, 2, \ldots$ with exponential holding time with parameter 1, and then jumps to $i + 1$. A continuous-time Markov chain $(X_t)_{t \in [0,\infty)}$ fulfills the following integral equation with $\gamma_l(i) = q_{i,i+l}$

$$X_t = X_0 + \sum_l l Y^l \left(\int_0^t \gamma_l(X_s) ds \right)$$

where the $(Y_t^l)_t$ are independent Poisson processes and we write Y_t^l as $Y^l(t)$; see [12], 6.4. Here the clock for the Poisson process Y^l runs with the random speed $\int_0^t \gamma_l(X_s)ds$. We shall see in the following how this representation immediately sheds light on the limiting behavior for such processes.

3 Basic facts on asymptotic behavior

Let us assume that for $N = 1, 2, \ldots$ we have continuous-time Markov chains $(Z_t^N)_t$ with state space S^N and transition rates q_{ij}^N, depending on N. In typical examples, N is the size of a population or of a network. For large N, the exact behavior is no longer tractable so one has to rely on approximations. In mathematical terms, the limiting behavior of suitably standardized versions $(X_t^N)_t$ of $(Z_t^N)_t$ as $N \to \infty$ has to be investigated. There are various mathematical methods to achieve this.

3.1 Convergence via infinitesimal operators

Suppose that $S^N \subseteq S$ for all N. The convergence of the process $(X_t^N)_t$ can be reduced to a study of the convergence of the infinitesimal operators A^N. This is a general fact and does not depend on the assumption of discrete state spaces which is discussed here.

Assume that we find an operator A, defined on a suitable set $\mathcal{D}$ of bounded functions $f : S \to \mathbb{R}$, where Af again is a function $Af : S \to \mathbb{R}$, such that the following condition holds

$$\lim_{N\to\infty} \sup_{y\in S^N} \left| A^N f(y) - Af(y) \right| \to 0.$$

If A generates a well-behaved Markov process $(X_t)_t$ in the mathematical sense of a Feller process, then

$$(X_t^N)_t \text{ converges to } (X_t)_t.$$

In exact mathematical terms, this convergence takes place in the sense of weak convergence in the function space $D_S[0, \infty)$, which includes the convergence of finite-dimensional distributions, e.g. the convergence of stationary distributions and distributions of first hitting times; we refer to [12], 1.6.1, 4.2.11. In many examples, A will be a differential operator on a certain set $\mathcal{D}$ of twice differentiable functions and the limit process will be a diffusion.

3.2 Convergence via the representation 2.3

For a concise discussion we let N denote the size of a population and $S^N = \{0, 1, \ldots, N\}$. Z_t^N describes the random number of a certain species within the population. The standardized version is given by the proportion of this species

$$X_t^N = \frac{Z_t^N}{N}$$

taking values in $\{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}$. Now assume that the transition rates for $(Z_t^N)_t$ fulfill

$$\gamma_l^N(i) = q_{i,i+l}^N = N\beta_l(\frac{i}{N})$$

for some bounded function β on $[0, 1]$, which may be generalized to $\gamma_l^N(i) = q_{i,i+l}^N = N(\beta_l(\frac{i}{N}) + O(\frac{1}{N}))$. This implies that the mean number of transitions in one unit of time is of the order N, equal in order to the size of the population. So we expect a law of large numbers to hold for X_t^N, hence a non-random limiting process $(X_t)_t$, and similarly a central limit theorem. This may be seen with the following arguments, and we refer to [12], 11.2 for a rigorous and complete discussion.

3.2.1 A law of large numbers

Let $\tilde{Y}_t = Y_t - t$ be a centered Poisson process. The law of large numbers for this process states that for all t

$$\sup_{s \le t} \left| \frac{1}{N} \tilde{Y}_{Ns} \right| \to 0 \text{ almost surely for } N \to \infty.$$

Using the representation from 2.3 for $(Z_t^N)_t$ we have, with $f(x) = \sum_l l\beta_l(x)$,

$$X_t^N = \frac{1}{N} Z_t^N = X_0^N + \sum_l l \frac{1}{N} \tilde{Y}^l \left(N \int_0^t \beta_l(X_s^N) ds \right) + \int_0^t f(X_s^N) ds.$$

So our random clock has a speed of order N. Assuming $X_0^N \to x_0$ and Lipschitz continuity of f, the law of large numbers for the centered Poisson process implies for all t

$$\sup_{s \le t} \left| X_s^N - X(s) \right| \to 0 \text{ almost surely for } N \to \infty.$$

where $X(t)$ is the non-random solution of the differential equation

$$\frac{d}{dt} X(t) = f(X(t)), X(0) = x_0.$$

3.2.2 A central limit theorem

The basic central limit theorem states that, for a sum of i.i.d. random variables X_i with finite mean μ and finite positive variance, the normalized sum $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(X_i-\mu)$ converges to a normal distribution. Writing

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(X_i-\mu)=\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}X_i-\mu\right)$$

we see that the term $\frac{1}{N}\sum_{i=1}^{N}X_i-\mu$, which tends to zero according to the law of large numbers, has to be enlarged by a factor $\sqrt{N}$ to obtain a non-degenerate asymptotic normal distribution. This is a wide-spread phenomenon, so in our setting it is readily conjectured that the stochastic process $\left(\sqrt{N}(X_t^N-X(t))\right)_t$ converges to a non-degenerate process which is Gaussian. To see this we use the central limit theorem for the centered Poisson process. This states that for the process $(W_t^N)_t$ with $W_t^N=\frac{1}{\sqrt{N}}(Y_{Nt}-Nt)$

$$(W_t^N)_t \text{ converges to a standard Wiener process } (W_t)_t.$$

Set

$$\begin{aligned}V_t^N &= \sqrt{N}(X_t^N-X(t))=V_0^N+\sum_l lW^{l,N}(\int_0^t\beta_l(X_s^N)ds)+\int_0^t\sqrt{N}(f(X_s^N)-f(X(s))ds\\ &= V_0^N+\sum_l lW^{l,N}(\int_0^t\beta_l(X_s^N)ds)+\int_0^t\sqrt{N}\left(f(X(s)+\frac{1}{\sqrt{N}}V_s^N)-f(X(s))\right)ds.\end{aligned}$$

Letting N tend to infinity, with $V_0^N\to v_0$, the limiting equation is

$$V_t=v_0+U_t+\int_0^t f'(X(s))V_s ds,$$

where $U_t=\sum_l lW^l(\int_0^t\beta_l(X(s))ds)$ is a Gaussian process. It can be shown rigorously that in fact

$$(\sqrt{N}(X_t^N-X(t)))_t \text{ converges to a process } (V_t)_t$$

with $(V_t)_t$ being a solution of the limiting equation. Since $(U_t)_t$ is a Gaussian process, $(V_t)_t$ also is a Gaussian process. Gaussian processes are uniquely determined by their mean and covariance function. From the limiting equation one can obtain the following expressions:

Let $g(t,s)$ be a solution of

$$\frac{d}{dt}g(t,s) = f'(X(t))g(t,s), \;\; g(s,s) = 1.$$

Then the mean of V_t is given by $g(t,0)v_0$ and the covariance is given by

$$Cov(V_t, V_r) = \int_0^{\min\{t,r\}} g(t,s)g(r,s)\sum_l l^2\beta_l(X(s))ds.$$

So we have that the process

$$(X_t^N)_t \text{ is approximated by } X(t) + \frac{1}{\sqrt{N}}V_t,$$

hence

$$(Z_t^N)_t \text{ is approximated by } (NX(t) + \sqrt{N}V_t)_t.$$

The distribution of Z_t^N is approximated by a normal distribution with mean $Nx_0 + \sqrt{N}v_0$ and variance $NVar(V_t)$.

4 Agent-based models and their asymptotic behavior

In agent-based models we have a population of N agents that are interacting on a financial market through their opinion formation process. These agents are of two different types, e.g. optimists and pessimists. The number of agents of one type, say optimists, is described by a continuous-time Markov chain $(Z_t^N)_t$ with state space $\{0, 1, \dots, N\}$. Agents may switch from one type to the other, and the usual assumption is that $(Z_t^N)_t$ follows a birth and death process where

$$q_{i,i+l} = 0 \text{ for } |l| > 1.$$

In this context, birth represents the conversion of a pessimist to an optimists, while death corresponds to the opposite conversion. The birth and death rates are, respectively,

$$\lambda_i = q_{i,i+1} \text{ for } i = 0, 1, \dots, N-1, \;\; \mu_i = q_{i,i-1} \text{ for } i = 1, \dots, N,$$

with $\lambda_N = \mu_0 = 0$. For a birth-and-death process, the probability of multiple switches in time interval of length t tends to zero as t tends to zero.

There are two benchmark models that have been proposed in the literature, let us call them Model 1 and Model 2, corresponding to extensive and non-extensive transitions in the jargon of [4], who interpret the two benchmarks as cases of local vs non-local agent interactions. They argue that this difference should somehow reflect the different *intensities of interpersonal coupling* among agents, which initially inspired work on the topological features of agent interaction [5, 6]. In this paper, we demonstrate that the intensity of interpersonal coupling also depends on the velocity of information transmission between agents, adding an important aspect to the generic herding model. Put differently, interesting long-range correlations among agents that are prone to herding can arise both from the topology and the velocity of information transmission in the agent network.

Model 1 (the extensive case) looks at the birth and death rates

$$\lambda_i = (N-i)(a + b\frac{i}{N}),\ \mu_i = i(a + b\frac{N-i}{N}),$$

Model 2 (the non-extensive case) looks at the birth-and-death rates

$$\lambda_i = (N-i)(a + bi),\ \mu_i = i(a + b(N-i))$$

where a describes the overall tendency to switch, often conceptualized as the impact of news arrival, while b describes the herding propensity among agents.

4.1 Analysis of Model 1

Referring to the discussion in Section 2, we slightly generalize the model to $\lambda_i = (N-i)(a_1 + b_1\frac{i}{N}),\ \mu_i = i(a_2 + b_2\frac{N-i}{N})$ and have with $y = \frac{i}{N}$

$$\begin{aligned} q_i &= \lambda_1 + \mu_i = N\left[(1-y)(a_1 + b_1 y) + y(a_2 + b_2(1-y))\right] \\ p_{i,i+1} &= \frac{\lambda_i}{\lambda_i + \mu_i} = \frac{1}{1 + [(1-y)(a_1 + b_1 y)]^{-1} y(a_2 + b_2(1-y))}, \\ p_{i,i-1} &= 1 - p_{i,i+1}. \end{aligned}$$

So we see that the expected number of switches, which may be loosely interpreted as being proportional to the expected number of encounters of agents in one unit of time, is of the order N. The probabilities $p_{i,i+1}$ and $p_{i,i-1}$ that describe in which direction switching occurs, only depend on the fraction of agents of one type in the population. The limiting behavior of the normalized process $X_t^N = \frac{Z_t^N}{N}$ follows readily from 3.2. Using the notation from this section, we have

$$\gamma_l^N(i) = q_{i,i+l} = N\beta_l(\frac{i}{N}),\ l = -1, 1,$$

with $\beta_1(x) = (1-x)(a_1+b_1x)$, $\beta_{-1}(x) = x(a_2+b_2(1-x))$ defined for $x \in [0,1]$. From this it is clear that 3.2 applies, and we have

$$f(x) = \beta_1(x) - \beta_{-1}(x) = (1-x)(a_1+b_1x) - x(a_2+b_2(1-x)).$$

The limiting process is the non-random solution of the differential equation, which is an inhomogeneous Bernoulli-differential equation,

$$\frac{d}{dt}X(t) = f(X(t)),\ X(0) = x_0.$$

As we have seen in 3.2, the standardized differences $\sqrt{N}(X_t^N - X(t))$ tend to a Gaussian process V_t. Assume $V_0 = 0$, the mean value function is zero. We have $f'(x) = b_1 - b_2 - a_1 - a_1 + 2(b_2 - b_1)x$, and the covariance function may be computed, at least numerically, from the linear differential equation

$$\frac{d}{dt}g(t,s) = f'(X(t))g(t,s),\ g(s,s) = 1,$$

and the resulting expression in 3.2. Let us now make the following standard assumption $b_1 = b_2 = b$. Then

$$f(x) = (a_1+a_2)\left(\frac{a_1}{a_1+a_2} - x\right),$$

and it follows that

$$X(t) = \frac{a_1}{a_1+a_2} - \left(x_0 - \frac{a_1}{a_1+a_2}\right)e^{-(a_1+a_2)t}.$$

This shows that $X(t)$ converges to $\frac{a_1}{a_1+a_2}$ as $t \to \infty$.
We have $f'(x) = -(a_1+a_2)$ and a solution of the above differential equation is given by

$$g(t,s) = e^{-(a_1+a_2)(t-s)}.$$

It follows that

$$\begin{aligned}
&Cov(V_t, V_r)\\
=\ &\int_0^{\min\{t,r\}} e^{-(a_1+a_2)(t+r-2s)}((2b - a_1 + a_2)s - 2bs^2 + a_1)ds\\
=\ &\frac{e^{-(a_1+a_2)(t+r)}}{2(a_1+a_2)}\int_0^{\min\{t,r\}\cdot 2(a_1+a_2)} e^{-s}(a_1 + \frac{2b-a_1+a_2}{2(a_1+a_2)}s - \frac{2b}{4(a_1+a_2)^2}s^2)ds\\
=\ &\frac{e^{-(a_1+a_2)(t+r)}}{2(a_1+a_2)}\Bigg[a_1 + \frac{2b-a_1+a_2}{2(a_1+a_2)} + 2\frac{2b}{4(a_1+a_2)^2}\\
&-e^{-\min\{t,r\}\cdot 2(a_1+a_2)}\Bigg(a_1 + \frac{2b-a_1+a_2}{2(a_1+a_2)}(1+\min\{t,r\}\cdot 2(a_1+a_2))\\
&+\frac{2b}{4(a_1+a_2)^2}(2 + 2(\min\{t,r\}\cdot 2(a_1+a_2)) + (\min\{t,r\}\cdot 2(a_1+a_2))^2)\Bigg)\Bigg]
\end{aligned}$$

4.2 Analysis of Model 2

In Model 2 we have, with slight generalization and $y = \frac{i}{N}$,

$$\begin{aligned}\lambda_i &= (N-i)(a_1+b_1 i) = N^2(1-y)\left(\frac{a_1}{N}+b_1 y\right),\\ \mu_i &= i(a_2+b_2(N-i)) = N^2 y\left(\frac{a_2}{N}+b_2(1-y)\right)\end{aligned}$$

with

$$\begin{aligned}q_i &= \lambda_i+\mu_i = N^2\left[(1-y)\left(\frac{a_1}{N}+b_1 y\right)+y\left(\frac{a_2}{N}+b_2(1-y)\right)\right]\\ p_{i,i+1} &= \frac{1}{1+\left[(1-y)\left(\frac{a_1}{N}+b_1 y\right)\right]^{-1}\frac{i}{N}\left(\frac{a_2}{N}+b_2(1-y)\right)}.\end{aligned}$$

There are two changes compared to Model 1. The expected number of switches during one time period has increased to the order of N^2, and the overall tendency of switching has decreased to the order of $\frac{1}{N}$. This, of course, is a dramatic change since for, say, $N = 100$ the number of switches during one time unit is no longer of the order 100 but of the order 10 000. Due to the first change, it is clear that a law of large numbers, as stated in 4.1 for Model 1, can no longer hold. To analyze the asymptotic behavior of the system, we use the method of infinitesimal operators as described in 3.1. Let $X_t^N = \frac{Z_t^N}{N}$ as before, and denote the infinitesimal operator of $(X_t^N)_t$ by A^N. It follows for $y = \frac{i}{N}, 0 < \frac{i}{N} < 1$

$$A^N f(y) = \lambda_i(f(y+\frac{1}{N})-f(y)) + \mu_i(f(y-\frac{1}{N})-f(y)).$$

Let $\mathcal{D}$ denote the set of continuous mappings $f : [0,1] \to \mathbb{R}$ which are twice-continuous differentiable in the interior such that the derivatives have a continuous extension to 0 and 1. A Taylor expansion for $0 < y < 1$ shows that

$$A^N f(y) = \frac{1}{N}(\lambda_i-\mu_i)f'(y) + \frac{1}{2N^2}(\lambda_i+\mu_i)f''(y) + (\lambda_1+\mu_i)o(\frac{1}{N^2})$$

with error term $o(\frac{1}{N^2})$ uniform in y. Inserting the birth and death rates of Model 2, we find

$$\begin{aligned}A^N f(y) &= \frac{1}{N}N^2\left((1-y)(\frac{a_1}{N}+b_1 y) - y(\frac{a_2}{N}+b_2(1-y))\right)f'(y)\\ &+ \frac{1}{N^2}\frac{N^2}{2}\left((1-y)(\frac{a_1}{N}+b_1 y)+y(\frac{a_2}{N}+b_2(1-y))\right)f''(y)+o(1)\\ &= N\left((1-y)\frac{a_1}{N} - y\frac{a_2}{N} + b_1 y(1-y) - b_2 y(1-y)\right)f'(y)\\ &+ \frac{1}{2}\left((1-y)(\frac{a_1}{N}+b_1 y)+y(\frac{a_2}{N}+b_2(1-y))\right)f''(y)+o(1).\end{aligned}$$

From this it is clear that $A^N f$ neither converges for $b_1 \neq b_2$ nor for some fixed overall switching tendencies $\tilde{a}_1, \tilde{a}_2$ that are not of order $O(\frac{1}{N})$. So with the standard assumption $b_1 = b_2 = b$ we have

$$\begin{aligned} A^N f(y) &= ((1-y)a_1 - ya_2)f'(y) + \frac{1}{2}\left(2b(1-y)y + \frac{a_1}{N}(1-y) + \frac{a_2}{N}y\right) f''(y) + o(1) \\ &\to ((1-y)a_1 - ya_2)f'(y) + \frac{1}{2}2b(1-y)yf''(y) \text{ as } n \to \infty. \end{aligned}$$

If A is the differential operator on $\mathcal{D}$ defined by the last line, then we see that

$$\sup_{y \in S^N} |A^N f(y) - Af(y)| \to 0 \text{ as } N \to \infty.$$

The set of all functions in $\mathcal{D}$ is rich enough to yield convergence of the process according to 3.1, see [15]. So $(X_t^N)_t$ converges to a diffusion on $[0,1]$ with drift term $\mu(y) = (1-y)a_1 - ya_2$ and diffusion term $\sigma^2(y) = 2by(1-y)$.

5 Mean field analysis of agent networks

If we conceptualize agent interaction with a communication network, we can model transition rates through a term that represents a general switching tendency and another term that depends on the particular network structure among agents. So for any agent α, the transition rate to switch from a state k to a state l might be modeled as a function

$$t_\alpha(k, l; N_\alpha(k), N_\alpha(l), N),$$

where N is the total number of agents in the network, $N_\alpha(k)$ is the number of neighbors of agent α in state k, and $N_\alpha(l)$ is the number of neighbors in state l. More complicated models are possible of course, involving neighboring agents in different states, but already a model of the above type is analytically intractable on the individual agent level. So one may resort to an analysis of the aggregate number of agents in a particular state k, with transition rates depending on the average number of neighbors and their states; see [5] for a corresponding mean-field derivation in Kirman-type models. Going back to two states, e.g. optimistic and pessimistic agents in the network, the number Z_t^N of optimistic agents may be viewed as a birth-and-death process. We can model the birth and death rates with $y = \frac{i}{N}$ as

$$\begin{aligned} \lambda_i &= h(N)(1-y)(a_1(N) + b_1(N)y), \\ \mu_i &= h(N)y(a_2(N) + b_2(N)(1-y)). \end{aligned}$$

Here $a_i(N)$ represents the aggregate switching tendency in the network, $b_i(N)$ describes the herding intensity that depends on the average opinion of neighbors in the network, and $h(N)$ gives the order of the mean number of switches within a time unit, loosely interpreted as being of the order of the mean number of encounters in the network. All the parameters $h(N), a_i(N), b_i(N)$ can depend on the network topology.

Suppose that any agent is linked to a fixed proportion γ of the other agents in the network, and that the mean number of encounters in one unit of time is of the order δN. Then, with $a_i(N) = a_i$ and $b_i(N) = \gamma b_i$, we recover Model 1 of Section 4

$$\begin{aligned}\lambda_i &= \delta N(1-y)(a_1 + \gamma b_1 y),\\ \mu_i &= \delta N y(a_2 + \gamma b_2(1-y)).\end{aligned}$$

Now assume that the mean number of encounters during one time unit increases to δN^2. Then, with $a_1(N) = \frac{a_1}{N}$ and $a_2(N) = \frac{a_2}{N}$, we arrive at

$$\begin{aligned}\lambda_i &= \delta N^2(1-y)(\frac{a_1}{N} + \gamma b_1 y),\\ \mu_i &= \delta N^2 y(\frac{a_2}{N} + \gamma b_2(1-y)),\end{aligned}$$

which corresponds to Model 2 of Section 4.

Further assume for any agent that the number of neighbors in the network grows slower than the total agent number N. Then we may model the proportion of neighboring agents as $\gamma f(N)$ where $f(N)$ tends to 0 as N tends to infinity. So we arrive at

$$\begin{aligned}\lambda_i &= h(N)(1-y)(a_1(N) + \gamma f(N) b_1 y),\\ \mu_i &= h(N) y(a_2(N) + \gamma f(N) b_2(1-y)).\end{aligned}$$

By a suitable choice of $a_i(N)$ and $h(N)$, however, we can again recover either Model 1 or Model 2. From a socio-economic point of view, the choice of these parameters fixes the speed of information transmission in the system. Speeding up the agent encounters, for instance to $h(N) = \frac{N^2}{f(N)}$ with overall switching tendency $a_i(N) = \frac{a_i f(N)}{N}$, we arrive at the N-independent Model 2

$$\begin{aligned}\lambda_i &= N^2(1-y)(\frac{a_1}{N} + \gamma b_1 y),\\ \mu_i &= N^2 y(\frac{a_2}{N} + \gamma b_2(1-y)).\end{aligned}$$

6 Conclusion

In summary we see that topological network features, like the average number of neighboring nodes, are only one possible source of N-(in)dependence in the class of agent-based herding models that we have discussed here. Our main finding is that link characteristics in the network, for instance the speed of information transmission along the links, are equally important for the central issue of N-(in)dependence. Put differently, herding models in the Kirman tradition can reproduce the stylized facts of financial markets for any system size by way of an appropriate combination of network topology and communication speed that will lead to a non-trivial limiting behavior of the stochastic processes that describe agents' opinion (or strategy) formation. Interestingly, casual observation would suggest that the decades since the collapse of Bretton-Woods were not only accompanied by an increase in the number of market participants, but also by technological advances that have vastly increased the speed of communication among agents. An emerging empirical question for future research is whether one can identify and disentangle the two effects in financial markets, which would obviously have profound implications for the understanding of financial fluctuations and the mitigation of risks that stem from them.

References

[1] S. Alfarano and R. Franke. A simple asymmetric herding model to distinguish between stock and foreign exchange markets. *Working Paper*, 2007.

[2] S. Alfarano and T. Lux. A noise trader model as a generator of apparent financial power laws and long memory. *Macroeconomic Dynamics*, 11(S1):80–101, 2007.

[3] S. Alfarano, T. Lux, and F. Wagner. Estimation of agent-based models: the case of an asymmetric herding model. *Computational Economics*, 26(1):19–49, 2005.

[4] S. Alfarano, T. Lux, and F. Wagner. Time-variation of higher moments in a financial market with heterogeneous agents: An analytical approach. *Journal of Economic Dynamics and Control*, 32:101–136, 2008.

[5] S. Alfarano and M. Milaković. Network structure and N-dependence in agent-based herding models. *Journal of Economic Dynamics and Control*, 33(1):78–92, 2009.

[6] S. Alfarano, M. Milaković, and M. Raddant. On network hierarchy and risk in Kirman's ant model. *Working Paper*, 2009.

[7] M. Aoki. *New Approaches to Macroeconomic Modeling.* Cambridge University Press, Cambridge, UK, 1998.

[8] M. Aoki. Thermodynamic limits of macroeconomic or financial models: One- and two-parameter Poisson-Dirichlet models. *Journal of Economic Dynamics and Control*, 32(1):66–84, 2008.

[9] D. G. Champernowne. A model of income distribution. *Economic Journal*, 63:318–351, 1953.

[10] Z. Ding, C. W. J. Granger, and R. F. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

[11] E. Egenter, T. Lux, and D. Stauffer. Finite-size effects in Monte Carlo simulations of two stock market models. *Physica A,*, 268:250–256, 1999.

[12] S. N. Ethier and T. G. Kurtz. *Markov Processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley Sons Inc., New York, 1986. Characterization and convergence.

[13] H. Föllmer, U. Horst, and A. Kirman. Equilibria in financial markets with heterogeneous agents: A probabilistic perspective. *Journal of Mathematical Economics*, 41(1-2):123–155, 2005.

[14] P. Gopikrishnan, M. Meyer, L. A. N. Amaral, and H. E. Stanley. Inverse cubic law for the distribution of stock price variations. *European Physical Journal B*, 3:139–140, 1998.

[15] A. Irle and J. Kauschke. Diffusion approximation of birth and death processes with applications to financial market herding models. *Working Paper*, 2008.

[16] A. Kirman. Epidemics of opinion and speculative bubbles in financial markets. In M. P. Taylor, editor, *Money and Financial Markets*, pages 354–368. Blackwell, Cambridge, 1991.

[17] A. Kirman. Ants, rationality, and recruitment. *Quarterly Journal of Economics*, 108:137–156, 1993.

[18] A. Kirman and Teyssière. Microeconomic models for long memory in the volatility of financial time series. *Studies in Nonlinear Dynamics and Econometrics*, 5(4-A3), 2002.

[19] I. N. Lobato and N. E. Savin. Real and spurious long-memory properties of stock market data. *Journal of Business and Economic Statistics*, 16:261–283, 1998.

[20] T. Lux and M. Ausloos. Market fluctuations I: Scaling, multiscaling and their possible origins. In A. Bunde, J. Kropp, and H. J. Schellnhuber, editors, *Theories of Disaster - Scaling Laws Governing Weather, Body, and Stock Market Dynamics*, pages 373–409. Springer, Berlin Heidelberg, 2002.

[21] A. Pagan. The econometrics of financial markets. *Journal of Empirical Finance*, 3:15–102, 1996.

[22] G. M. Schütz, F. P. A. Prado, R. J. Harris, and V. Belitsky. Short-time behaviour of demand and price viewed through an exactly solvable model for heterogeneous interacting market agents. *Physica A*, 388:4126–4144, 2009.