Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach

Simone Alfarano[*], Thomas Lux, Friedrich Wagner

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

Abstract

A growing body of recent literature allows for heterogenous trading strategies and limited rationality of agents in behavioral models of financial markets. More and more, this literature has been concerned with the explanation of some of the stylized facts of financial markets. It now seems that some previously mysterious time series characteristics like fat tails of returns and temporal dependence of volatility can be observed in many of these models as macroscopic patterns resulting from the interaction among different groups of speculative traders. However, most of the available evidence stems from simulation studies of relatively complicated models which do not allow for analytical solutions. In this paper, this line of research is supplemented by analytical solutions of a simple variant of the seminal herding model introduced by Kirman [1993, Ants, rationality, and recruitment. Quarterly Journal of Economics 108, 137–156]. Embedding the herding framework into a simple equilibrium asset pricing model, we are able to derive closed-form solutions for the time-variation of higher moments as well as related quantities of interest enabling us to spell out under what circumstances the model gives rise to realistic behavior of the resulting time series.
© 2007 Elsevier B.V. All rights reserved.

---

[*]Corresponding author. Tel.: +49 431 880 5136; fax: +49 431 880 4383.
  E-mail address: alfarano@bwl.uni-kiel.de (S. Alfarano).

1. Introduction

Until very recently, theoretical research in finance has largely ignored some of the really universal stylized facts of practically all available financial data. In fact, a glance at frequently used textbooks like the ones by O'Hara (1995) and Barucci (2003) shows that even their glossaries lack entries for some of the prevalent technical terms of the empirical finance literature. For example, while many developments in empirical finance are essentially motivated by the observation of non-Gaussian returns distributions with their 'fat tails' and temporal dependence of second moments leading to 'volatility clustering', these notions have been almost entirely absent from the theoretical literature.[1] While these phenomena have spurred the development of such seminal innovations like GARCH type and stochastic volatility models in empirical finance, their behavioral origins have apparently remained an almost unaccessible puzzle for a long time. Two reasons might be responsible for this neglect: first, the above features characterize the behavior of financial time series as a whole, while the interest in economic theory has typically been to spell out the effect of a change of one (endogenous) economic variable on other, exogenous variables. Even when allowing for an ensemble of traders, such a comparative statics approach is not appropriate for explaining universal conditional and unconditional stochastic properties. Second, the prevalent efficient market paradigm did, in fact, provide a very simple implicit answer to the question of the origin of *all* stylized facts of returns: since, in this framework, prices would reflect forthcoming news in an unbiased and immediate manner, any property of the returns distribution would simply reflect a similar feature of the distribution of new information items. As a corollary of the efficient market hypothesis, the 'news arrival process' would, therefore, have to come along with fat tails and clustering of important news. Unfortunately, this corollary can hardly be subjected to econometric scrutiny. On the other hand, enough evidence had been collected against the universal validity of the efficient market paradigm to motivate alternative, behavioral approaches which then mushroomed over the nineties.

First analyses of complex data-generating mechanisms based on interacting agents can be found in Kirman (1991, 1993) and DeGrauwe et al. (1993). While they did not focus then on the above stylized facts (not broadly acknowledged at that time among theoretical researchers), they both already showed that their models could mimic the random walk nature of asset prices and exchange rates although their data-generating processes were clearly different from a true random walk. Notably, both studies also investigated what might be described as secondary stylized facts: they applied certain popular econometric analyses to their simulated data and found similar behavior as with empirical records providing a possible explanation of, e.g. the forward premium puzzle of foreign exchange markets.

Evidence for volatility clustering as an emergent phenomenon of a multi-agent model appeared first in Grannan and Swindle (1994). While a large body of

---

[1]Perfectly the same holds for theoretical work on exchange rate determination where the same stylized facts have also remained unnoticed for a long time.

subsequent models studied artificial markets with heterogenous autonomous agents often endowed with some sort of artificial intelligence (classifier systems, genetic algorithms), any consideration of empirical stylized facts is also curiously absent in the first wave of such papers (Levy et al., 1994; Arifovic, 1996; Arthur et al., 1997). In fact, much of this early literature had been preoccupied with the question of convergence or not of their learning algorithms to the benchmark of rational expectations rather than considering empirical applications. However, subsequent research has shown that relatively simple models of interacting traders could produce realistic time series sharing the 'stylized facts' of fat tails and clustering volatility even up to numerical agreement with key quantities of empirical data (Lux and Marchesi, 1999, 2000). Similar investigations of the dynamic properties of alternative models revealed that many agent-based approaches share a certain tendency of generating fat tails and volatility clustering although their quantitative manifestations are not always identical to the very robust numbers obtained with empirical data (cf. LeBaron et al., 1999; Chen and Yeh, 2002; Kirman and Teyssière, 2002; Lux and Schornstein, 2005).

Often 'realistic' dynamics are identified as the consequence of some kind of critical or intermittent process. In Lux and Marchesi (1999, 2000) as well as in Giardina and Bouchaud (2003) and Lux and Schornstein (2005) the dynamics are characterized by endogenous changes in the population composition triggered by both systematic factors (i.e. some fitness criterion like profits or utility obtained by a particular trading strategy) and noise (idiosyncratic components or experimentation). However, the stability of the entire dynamic system also depends on the strategy configuration within the population and with its random dynamics the population composition switches between stable and unstable configurations leading to intermittent bursts of volatility. Similar mechanisms probably prevail in the artificial markets of Raberto et al. (2003) and Chiarella and Iori (2002).

A related branch of agent-based models with realistic time series properties has been launched by Cont and Bouchaud (2000) adapting the framework of percolation models from statistical physics. In these models, traders are situated in a lattice and form clusters of agents with the same 'strategy' (buy, sell, remain inactive). Contributions to this literature have been inspired by the well-known critical behavior of percolation models at a certain threshold of the connectivity level and have added various additional features for preserving the critical power-law fluctuations for a wider range of parameters or adding temporal dependency (cf. Bornholdt, 2001; Stauffer and Sornette, 1999; Iori, 2002). Another closely related approach is the discrete choice framework for the choice of agents' strategies adopted by Gaunersdorfer and Hommes (2007). Giving rise to a dynamic system with coexisting attractors (e.g. a stable fix point together with a limit cycle) they obtain realistic dynamics by allowing for additional noise that triggers intermittent switching between different basins of attraction.

While in most of these models the general appearance of simulated data seemed to be quite robust with respect to most of the underlying parameters, it often also turned out that the potential of generating stylized facts depends crucially on the system size (i.e. number of agents). Realistic dynamic patterns are typically

observed with the (probably natural) initial choice of a few hundred or thousand agents. However, with an increasing number of market participants one often experiences a fading away of the fat tails and volatility clustering beyond a certain threshold (cf. Egenter et al., 1999; Lux and Schornstein, 2005; Challet and Marsili, 2004).

The present paper attempts to shed light on both of the findings detailed above. Within a relatively simple type of herding model (broadly along the lines of Kirman, 1993) we derive closed-form solutions for auto-covariances of returns and their higher moments together with other statistics such as mean-passage times. Inspection of the results allows to infer in how far and under what conditions the model could mimic the empirical findings of fat tails and clustering of volatility. Investigation of different specifications of the model also allows us to point out why – in certain scenarios – increasing system size would lead to vanishing stylized facts. Our approach is broadly complementary to recent attempts at studying asymptotic properties of related agent-based models (Horst, 2004; Föllmer et al., 2005). These authors provide conditions under which the limiting distribution of the price process exists in models with both global and local interactions of agents (Horst, 2004) and models with feedback from the price process on the group dynamics (Föllmer et al., 2005). Since our model can be viewed as a special case of the class of models studied in Horst (2004) his results on convergence of equilibrium prices to a unique equilibrium distribution also apply in our case. However, instead of focusing on the properties of the price process alone we extend the analysis to returns to which the famous empirical regularities of fat tails and clustered volatility apply. We also go beyond asymptotic convergence results by working out various properties of the stationary distribution which are of interest in the light of empirical findings. The main results of our paper are:

- the derivation of closed-form solutions for conditional and unconditional moments which provide insights into the origins of leptokurtosis and heteroskedasticity in this model.
- the derivation of results on the dependency of their realistic features on system size. As it will turn out, whether or not we observe a fading away of stylized facts depends on rather subtle details of the formalization of transition rates (extensive vs. non-extensive transition rates, see below). While both variants can be clearly distinguished in the present framework, it might not be easy to determine whether certain more complicated models fall into the first or second category.

The paper is organized as follows: Section 2 introduces different variants of herding models inspired by Kirman's ant process (Kirman, 1993). Section 3 incorporates the recruitment mechanism of the ant model into an asset pricing framework. Section 4 details our analytical results for both the unconditional and conditional distribution of returns. In Section 5 we explain different techniques for simulating these agent-based models, while Section 6 points out under which conditions the interesting dynamics survives for large numbers of agents. Section 7 concludes. Technical details are relegated to various appendices.

2. The ant model and its financial interpretation

In his seminal 1993 paper, Kirman provides a simple stochastic formalization of information transmission inspired by macroscopic patterns emerging from entomologic experiments with ant colonies. The underlying scenario is one of foraging ants who have two identical sources of food at their disposal in the vicinity of their nest. Experimental settings show that, at any point in time, a majority of the ant population concentrates on exploiting one particular food source, but every once in a while a switch occurs from one to the other source. Thus, averaging over time, a bimodal distribution of the frequency of ants visiting one or the other manger would result. Kirman explains this phenomenon by a combination of pairwise interaction (exchange of information by pheromons) and an autonomous switching probability due to stochastic search. With an increase of the strength of the herding component, the resulting stationary probability distribution of the model bifurcates from uni-modality with an equal exploitation of both sources to a bi-modal one supporting the experimental findings.

In a long series of subsequent papers, this and similar mechanisms have been applied as formalizations of contagion effects in financial markets. Kirman (1991, 1993) replaces the binary choice of food sources by the ants by agents' choice of a particular rule for the formation of their expectations. In his adaption to foreign exchange market dynamics, agents can adopt chartist and fundamentalist forecasts of future exchange rate movements. With the herding mechanism governing agents' distribution on these two behavioral types, the market switches between a dominance of fundamentalist and chartist behavior with periods in which the later prevails giving rise to speculative bubbles. Market equilibria are determined from a standard monetary exchange rate model in which the assumption of rational expectations of the original framework has been replaced by 'non-rational' expectations computed as the average of the expectations of the two groups of chartists and fundamentalists. Most relevant to our purpose, Kirman and Teyssière (2002) have analyzed the temporal dependence of returns and volatility from simulated data using a variety of pertinent econometric tools and have found similar indication for long-term dependence in second moments as with real-life data.

A closely related variant of a herding model has been proposed by Lux (1995), who replaced the pair-interaction by an overall (mean-field) effect of the average influence of the whole ensemble of traders on each individual's choice of strategy. This setting allowed to derive a number of analytical results via application of the Master equation approach adopted from statistical physics (Lux, 1997, 1998). An extended version of this model (Lux and Marchesi, 1999) with switching between three different groups could be shown to give rise to realistic time series properties. Egenter et al. (1999) showed, however, that these results depended on the number of agents and that raising the number of market participants beyond several thousand individuals leads to a gradual fading away of all its 'interesting' dynamics. We will point out later what mechanisms are responsible for this loss of stylized facts and under which assumptions 'interesting' solutions can be preserved even in large markets.

To set the stage, let us first restate the mechanics of the herding model proposed by Kirman (1993). The market (or ant colony) is populated by a fixed number of agents N, each of them being either in state 1 or 2. The number of agents in the first state will be denoted by n, so that $n \in \{0, 1, \ldots, N\}$ defines the configuration of the system (the number of agents in the opposite state will obviously be $N - n$). The stochastic evolution of the system is governed by random meetings of pairs of agents, after which one of them might follow the companion, and autonomous switches, i.e. independent random changes of state. In the original set-up, the encounters of agents are formalized by pairwise interaction involving two ants/agents at a time, which implies that the density of ants in the given search area is low enough to exclude the possibility of multiple meetings. Additionally, *neither* the probability of following another companion *nor* the success in recruiting companions depend on the outcome of previous meetings. The lack of memory of the agents is the crucial assumption to formalize the population dynamics as a Markov process. It, then, evolves according to the probabilities of changing from state n at time t to some state n' at time $t + \Delta t_0$. Let these Markovian conditional probabilities be denoted by $\rho(n', t + \Delta t_0 | n, t)$ which are related to the transition rates per unit time, $\pi(n \to n')$ by $\rho(n', t + \Delta t_0 | n, t) = \pi(n \to n') \Delta t_0$ for small time increments Δt_0. Since in the limit of continuous time $\Delta t_0 \to 0$, multiple switches during one incremental time unit become increasingly unlikely, one can confine the analysis to $n' = n \pm 1$ with transition rates:

$$\pi(n \to n + 1) = (N - n)\left(a + \bar{b}\frac{n}{N}\right),$$

$$\pi(n \to n - 1) = n\left(a + \bar{b}\frac{(N - n)}{N}\right). \tag{1}$$

The term proportional to the number of agents in the second state models their random changes, while the quadratic term $\bar{b}((N - n)n/N)$ governs the recruitment mechanism, the parameter $\bar{b}$ measuring the strength of the influence of the companions.[2]

Some comments are in order here. First, note that the transition rates are invariant under the transformation $n \to N - n$. This reflects the fact that the recruitment mechanism is independent of the type of agents, and the random switching component a is equal in both states. The system, therefore, does not exhibit any preferential direction between the two states (in Alfarano et al., 2005, a more general asymmetric setting has been investigated). This underlying symmetry is responsible for the resulting symmetric shape of the unconditional distribution of ants' visiting time (see Section 4.3). Moreover, the above transition rates are non-linear because of the presence of the quadratic dependence on n. In general, non-linearity in the transition rates arises from the presence of interactions among agents. In Kirman's model, an individual in state 1 is indirectly affected by agents who are in the other state through the recruitment mechanism. Linear transition rates, conversely,

---

[2]In the original version $\bar{b} = 1 - \delta$, where δ is interpreted as the probability of an unsuccessful attempt at recruitment.

describe uncorrelated movements between groups due to idiosyncratic motivations and, in general, systems composed by independent individuals.

An additional crucial property is the *extensivity* of the transitions in (1), since they satisfy the following relation:

$$\pi(n \to n') = N\pi(z \to z'),\tag{2}$$

where $z = n/N$ stands for concentrations rather than the raw numbers of agents, n, as a variable to characterize the entire system. The extensivity property typically characterizes systems whose evolution is governed by interaction mechanisms involving a small number of neighbors as compared to the population size. As we will see more precisely later on, this 'local' nature of the interactions among individuals has remarkable consequences in the 'thermodynamic' limit (i.e. when the number of individuals N becomes very large): the equilibrium probability distribution $\omega_e(x)$ of the concentration of agents x becomes sharp around a macroscopic state variable, which might be defined by the mean[3] of $\omega_e(x)$. In other words, the fluctuations of the system can be neglected, and its dynamics can be described by a deterministic differential equation (see Section 6).

What happens if we relax the extensivity assumption (2) based on local interactions among individuals? In which way would the dynamics of the system change under the hypothesis of collective interactions? In order to study possible differences in these alternative scenarios, we introduce a new pair of transition rates, where the extensivity property (2) does not hold:

$$\pi(n \to n + 1) = (N - n)(a + bn),$$
$$\pi(n \to n - 1) = n(a + b(N - n)).\tag{3}$$

The transition rates (3) describe a scenario where the interaction among agents does not depend on the *fraction* of agents in the alternative states, but rather on the overall *number* of such agents. More precisely, an individual might change state due to a random event, as in the previous setting, or he might switch under the group pressure of *the overall number of the individuals* that belong to the opposite state. A closer inspection of Eqs. (1) and (3) shows that the main difference lies in the scaling properties of the herding parameter. In the first formalization we observe a dependence of the type $O(1/N)$, with the herding parameter vanishing in the limit $N \to \infty$. In the second setting, on the contrary, $b \sim O(1)$, which therefore remains bounded in this limit. One of the contributions of this paper is to point out the interesting consequences of this different formalization, and the implications in the characterization of the resulting dynamics in the two different scenarios, namely extensive versus non-extensive transition rates when we increase the number of individuals N.

Would it make *intuitive* sense to replace intensities by occupation numbers as we have done in Eq. (3)? First, let us consider the autonomous switching component. Increasing the number of agents and keeping the Poisson probability a of an

---

[3]Alternative definitions could be based on the mode, the median or other characterizations of the typical value of the distribution $\omega_e(x)$.

autonomous switch by any particular agent per time unit constant, larger n and $N - n$ would also lead to a higher probability of observing one such move within the time interval Δt. Since agents are independent in this autonomous component, the linearity in n and $N - n$ assumed in Eqs. (1) and (3), in fact, appears intuitively plausible. Replacing, for example, na by $(n/N)a$ in the second equation would amount to assuming that probability a does not apply to each individual agent herself but rather to a fraction of the population. One could argue in the same way with respect to the herding component: increasing the overall number of ants or agents (and assuming that they are distributed equally in the space or have the same probability of meeting other agents), the likelihood for agents to randomly bump into each other should also increase, presumably linearly, in the number of agents (if we were not to incorporate particular topological considerations into the model). Vice versa, a formulation in intensities, i.e. expressing the conversion probabilities by $b(n/N)$ and $b(N - n)/N$ in Eq. (1) would amount to a constant probability of meeting agents from the other group independent of the overall number of agents which – at least in the literal interpretation of the model – would be implausible. If we imagine to increase the number of ants N and to keep the area S constant, the resulting increase of the density of the colony would lead to a higher frequency of pairwise contacts, and finally, to a uniform mass of ants covering all the area S. However, if we use the entomological information transmission mechanism as a parabola for herding among humans, we are not anymore constrained to a bounded region due to the 'non-spatial' nature of modern communication systems. It is, therefore, not obvious which of the two approaches would be preferable in an economic setting. Using Eq. (1) to formalize the interactions, the extensivity property, then, constrains the information transmission to a *limited number of neighbors*. The alternative approach, formalized by Eq. (3), describes a scenario with connectivity increasing with the number of agents. It can be used to model, for instance, group pressure or interaction via mass media.

In the literature on agent-based economics, one often finds a formulation in which transition rates depend on intensities rather than on group occupation numbers, i.e. on n/N and $(N - n)/N$ instead of n and $N - n$ (Lux, 1995). Apparently, system size appears to become suppressed in such a formalization and one might argue that this would make the model outcome insensitive with respect to the market size (as it turns out, the opposite, in fact, happens). In order to investigate the effects of occupation number formalism vs. transition probabilities expressed in terms of concentrations later on, we can integrate this alternative approach into the framework of Eqs. (3) by replacing the constant b by a flexible form, $b = b_0 N_0/N$ in which b_0 and N_0 are the benchmark values for some initial scenario with constant parameters a and b_0 and benchmark population N_0. The changes brought about by an increasing population in the concentration formulation would, then, simply amount to multiplying the original herding constant by N_0/N which obviously leads to an effective decrease of the susceptibility of agents to contagion effects with increasing population size. For a sufficiently high value of N, this transformation triggers qualitative macroscopic changes in the resulting stationary distribution, from a bimodal equilibrium distribution to a 'trivial' Gaussian.

3. The financial market

In order to close the model, we embed the herding dynamics into a simple noise trader framework. In particular, we interpret groups 1 and 2 as optimistic and pessimistic noise traders, respectively, who expect the price to increase or decrease in the near future. Optimists are assumed to buy a certain number T_c of additional units of the asset while pessimists sell T_c units (T_c, thus, is a measure of average transaction volume of chartists). Besides these chartists or noise traders, whose belief dynamics are described by the above herding model, the model also hosts a second group of fundamentalist traders who buy/sell if the current price p is below/above their perceived fundamental value p_f. Assuming that this group has N_f members with average trading volume T_f and that they react on relative deviations between p and p_f, excess demand by this group amounts to

$$ED_f = N_f T_f \ln\left(\frac{p_f}{p}\right). \tag{4}$$

Since noise traders' excess demand mainly derives from a dominance of optimism and pessimism in this sub-population, we can conveniently trace it back to a quantity describing the *population configuration*: $x = 2n/N - 1$, which if zero signals a balanced disposition (zero excess demand) and if positive or negative amounts to optimistic or pessimistic majorities. Noise traders' excess demand, therefore, becomes $ED_C = NT_c x$. Invoking a standard Walrasian price adjustment mechanism, relative price changes would depend on overall excess demand:

$$\frac{1}{p}\frac{dp}{dt} = \beta\left(N_f T_f \ln\left(\frac{p_f}{p}\right) + NT_c x\right), \tag{5}$$

with β the assumed price adjustment speed. Except for the incorporation of the herding mechanism through x, this is a fairly standard asset price equation similar to what one finds in the legacy of models with chartist–fundamentalist interaction (e.g. Beja and Goldman, 1980; Day and Huang, 1990).

Setting (without loss of generality) $N_f T_f/NT_c = 1$ and assuming instantaneous market clearing ($\beta \to \infty$), we arrive at the equilibrium price driven both by fundamental information and the average mood of noise traders:

$$p(t) = p_f \exp(x(t)). \tag{6}$$

Hence, derivation of the limiting distribution of the configuration x together with a specification of the fundamentals process would allow us to characterize the distribution of equilibrium prices. However, we are interested in returns rather than prices themselves as only the later are known to be non-stationary in real-life markets. From Eq. (6), returns over *arbitrary* time horizons are computed as

$$r(t,\tau) = \ln p(t+\tau) - \ln p(t) = (x(t+\tau) - x(t)), \tag{7}$$

assuming that p_f is constant over time. Volatility can be obtained as $V(t,\tau) = r^2(t,\tau)$. Eqs. (1)–(6) define a simple noise trader/infection framework with heterogenous interacting agents which is quite similar in spirit but simpler than the formalization

of Kirman (1991), and Kirman and Teyssière (2002). Since the population configuration from the infection process enters directly on the demand side of the model it is, in fact, possible to infer the statistical properties of the returns and volatility processes under the influence of this ant-like contagion dynamics.

It is interesting to compare the above framework to the way in which the ant model is incorporated into a monetary exchange rate model in Kirman (1991) and Kirman and Teyssière (2002). In their papers, the contagion dynamics occurs on a much faster time scale than price formation in the foreign exchange market. In particular, they extract the population configuration at integer times t as a snapshot of the distribution of agents after a large number of pairwise meetings (for example, 10,000 meetings in Kirman, 1991) at the intra-daily level. The current distribution is then used to compute the weights of chartists and fundamentalists. Defining the market expectation as the average of current chartist and fundamentalist forecasts, the equilibrium exchange rate is derived from a discrete-time monetary model.[4] The separation of time scales allows to observe relatively large fluctuations of the average opinion at integer time steps, but it also impedes to take stock of the theoretical solution of the ant process for a formal analysis of the exchange rate dynamics. In contrast, in our approach, the time scales of the ant process and price formation are the same (although we will later distinguish between the continuous time dynamics of the process and discrete measurements thereof). With our Poisson probabilities, both occur in continuous time with changes of the configuration of noise traders coming along with changes in equilibrium prices.

4. Analytical results

The structure of the model allows to apply the Master equation formalism in order to derive macroscopic properties of our dynamical system (cf. Aoki, 2002). Note that our financial market consists of a large collection of interacting jump processes with state-dependent and non-linear transition rates. In this section we show how this formalism allows for a compact characterization of the aggregate dynamics of the system in terms of its underlying parameters.

4.1. Fokker–Planck approximation

Sticking with our definition of *non-extensive* transition rates in Eq. (3), the group dynamics is characterized by the transition probabilities for agents' moves between the two states. These transition probabilities define a Markovian stochastic process which belongs to the class of so-called non-linear 'one-step processes' (cf. Van Kampen, 1992). In the following, we approximate the Markovian process (3) by a continuous diffusion process governed by the Fokker–Planck equation (FPE hereafter). First note that the probabilities $\bar{\omega}(n,t)$ to have n agents in state

---

[4]There is, in fact, even a further intermediate step in which agents try to make an assessment of the prevailing majority opinion on the base of a noisy signal. The distribution of this observation rather than their own previous state is used in the computation of the weights of both groups.

one at time t, obey the so-called Master equation, which gives the probability flux between states

$$\frac{\partial \bar{\omega}(n,t)}{\partial t} = \sum_m (\pi_{n,m} \bar{\omega}(m,t) - \pi_{m,n} \bar{\omega}(n,t)),$$ (8)

with $\pi_{n,m} = \rho_{n,m}/\Delta t_0$ the transition rates from states m to states n. In the limit $\Delta t_0 \to 0$, 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 x and treating it, for large N, as a continuous quantity, we can derive a Fokker–Planck equation for the time change of the pertinent probability density $\omega(x,t)$ (cf. Appendix A.1 for details):

$$\frac{\partial \omega(x,t)}{\partial t} = \frac{\partial}{\partial x}(A(x)\,\omega(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(D(x)\omega(x,t)).$$ (9)

The drift and diffusion terms $A(x)$ and $D(x)$ in Eq. (9) are, respectively, given by

$$A(x) = -2ax, \quad D(x) = 2b(1-x^2) + \frac{4a}{N}.$$ (10)

The process is, therefore, characterized by a *linear* drift component with mean-reversion towards its unconditional mean value $E[x] = 0$. Additionally, the diffusion function, governing the properties of the fluctuations of the system, consists of two terms: one term is N-dependent, and becomes negligible for large values of N; the second term is a symmetric and quadratic function of x, which is independent of N. The two terms characterize *two different mechanisms* for the emergence of endogenous fluctuations. The N-dependent term governs the fluctuations due to the intrinsic 'granularity' of the system, which obviously has to vanish in the continuous approximation. The second term is directly related to the herding interactions among individuals. If we are not too close to the boundaries $x = 1$ and -1, the herding mechanism dominates, and the fluctuations are governed by the quadratic term, while the N-dependent term can be neglected.[5] The two components are comparable in magnitude in the vicinity of the boundaries only, so that for values of x close to ± 1 the N-dependent term should also be taken into account. Note that it is indeed a remarkable property of the system that the diffusion function does not depend on the number of agents N – if we neglect the granularity term. Asymptotic independence of N implies that the magnitude of the fluctuations of the opinion index x is also independent of the system size, and we observe qualitatively similar behavior for systems with any number of agents (see Fig. 1). The 'collective' nature of interactions among individuals, as it has been formalized in the non-extensivity of the transition rates (3), is the ultimate reason for this particular behavior.

If we employ the original Kirman setting as the underlying scenario for the artificial market, the previous property of independence of the level of the

---

[5]More precisely, the dominance of the herding term over the granular term is guaranteed if the inequality $2b(1-x^2) > 4a/N$ holds, which is equivalent to constraining the variable x to the interval $(2a/bN - 1, 1 - 2a/bN)$.
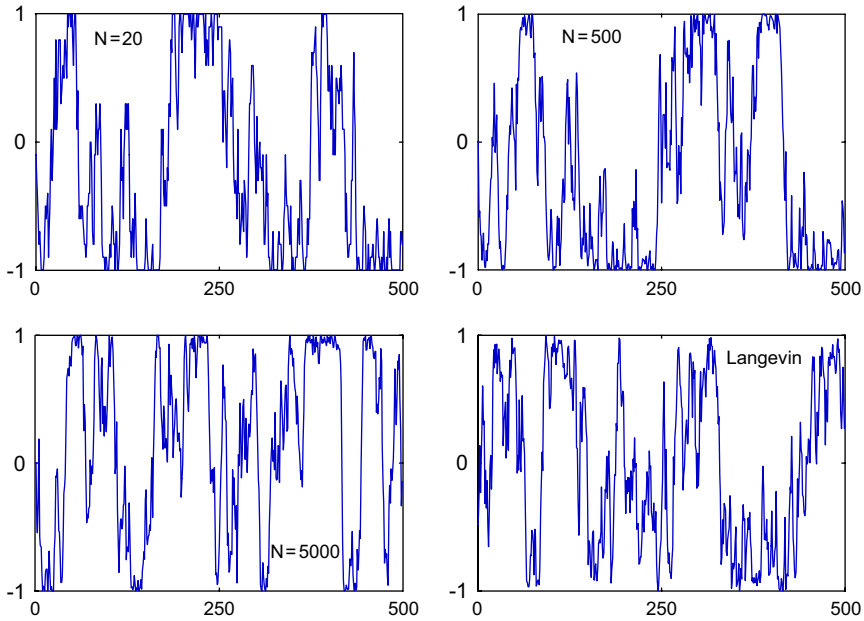
Fig. 1. The four panels show snapshots of the dynamics of the population index for increasing values of the number of agents *N*. The right bottom panel shows a trajectory computed via the Langevin approach given by Eq. (12). The other three panels show different examples of the simulation of the model with transition rates according to Eq. (3), as detailed in Section 5.1.

fluctuations of the system size does not hold any more. Let us assume that the underlying transition rates between the two states are now governed by Eq. (1) rather than by Eq. (3). In this setting we can also derive a diffusion approximation for the dynamics of the stochastic process (1), pretty much along the same lines as the derivation of the FPE (9) in the case of non-extensive transition rates (3). The main difference consists in replacing the parameter b in the diffusion function (10) by $\bar{b}/N$. We, then, find the following drift and diffusion functions:

$$A(x) = -2ax, \quad D(x) = \frac{2\bar{b}}{N}(1 - x^2) + \frac{4a}{N}. \tag{11}$$

While the drift function remains unchanged, the diffusion function, in this new setting, is *N*-dependent. This means that in the limit of large *N*, the fluctuations progressively attenuate, and eventually become completely negligible (see Section 6).[6]

4.2. Langevin equation

Keeping with the non-extensive formulation of Eq. (3), we might, as another way to analyze the dynamics of our model, derive a stochastic difference equation

---

[6]In the limit $N \to \infty$, convergence to the equilibrium distribution would take infinitely long time.

(Langevin equation) as a Gaussian approximation to the stochastic process over an appropriate time scale Δt. In Appendix A.6 we sketch the derivation of the Langevin equation as an approximation to the discrete process (3). As we show in the Appendix, we recover the drift and diffusion terms in the Fokker–Planck equation in the following Langevin equation:

$$x(t + \Delta t) = (1 - 2a\Delta t)x(t) + \sqrt{2b\Delta t(1 - x^2)}\eta(t). \tag{12}$$

In Eq. (12), $\eta(t)$ is an iid noise variable following a Normal distribution with mean zero. In Appendix A.6, we show how (12) can be derived from the Fokker–Planck equation if Δt is small enough to justify negligence of terms of order $(a\Delta t)^2$ and $(b\Delta t)^2$. Although for small step sizes Δt, the Langevin equation indeed provides a close approximation to the underlying agent-based model, it has the drawback that it might violate the built-in boundaries $x \in [-1, 1]$ of the population dynamics. To keep x within meaningful bounds in simulations of (12), reflecting boundaries at $x = \pm 1$ need, therefore, to be imposed by hand. Despite these difficulties, (12) has the important advantage that it not only facilitates numerical simulations, but also enables us to compute conditional and unconditional moments of the price process.

4.3. Equilibrium distribution

In the following, we derive some macroscopic statistical properties of the system governed by the transitions (3) and approximated by the FPE (9) with drift and diffusion given by (10).

Interestingly, the parameters a and b do not enter directly neither into the unconditional distribution $\omega_e(x)$ nor into the unconditional moments, but only via their ratio $\varepsilon = a/b$. As detailed in Appendix A.2, the following closed-form expression for $\omega_e(x)$ can be derived:

$$\omega_e(x) = \frac{\Gamma(2\varepsilon)}{2^{2\varepsilon - 1}\Gamma(\varepsilon)^2}(1 - x^2)^{\varepsilon - 1}. \tag{13}$$

We recover the main findings of Kirman (1993): $\varepsilon > 1$, i.e. dominance of autonomous switches, implies uni-modality with a peak at $x = 0$ and possibly small fluctuations of $\ln(p_t/p_f)$ around zero. Dominance of the herding component, $\varepsilon < 1$, however, leads to a bimodal distribution with probability mass reaching its maximum at ± 1, generating phases of overvaluation and undervaluation of the price compared to its fundamental value. In the knife-edge scenario $\varepsilon = 1$, the autonomous switching propensity is exactly balanced by the herding tendency, generating a uniform distribution of x. Note that the distribution (13) is symmetric around its mean value 0, which reflects the underlying symmetry between the two states of the system.[7]

It is interesting to show that for $\varepsilon \gg 1$, the equilibrium distribution converges to a Gaussian with mean zero and variance $1/2\varepsilon$. In this case, Eq. (13) can be

---

[7]Alfarano et al. (2005) derive closed-form solutions for an asymmetric herding model with different propensities for autonomous moves, a_1 and a_2, in one or the other direction. In this more general case, the resulting unconditional distribution is asymmetric, but covers Eq. (13) as a special case.

written as[8]

$$\omega_e(x) \rightarrow \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right], \tag{14}$$

where $\sigma^2 = 1/2\varepsilon$.

If one would like to investigate the case of pure herding, $\varepsilon = 0$ (i.e. $a = 0$), the boundary conditions would have to be modified in order to avoid absorbing states at $|x| = 1$. Following for example Lux and Marchesi (1999), reflecting boundary conditions can be imposed so that at least n_0 agents remain always optimistic or pessimistic. Let $\phi = n_0/N$, the boundary conditions then require the probability current to vanish at $|x| = 1 - \phi$. The parameter ϕ can become arbitrarily small for $N \rightarrow \infty$. Alfarano and Lux (2007) provide analytical results and numerical simulations of this scenario. They show that the stochastic switches between the vicinity of $x = 1$ and -1 also give rise to volatility clustering, a leptokurtic distribution of returns and persistence in the level of squared and absolute returns.

4.4. Unconditional moments of returns

Using the equilibrium distribution (13) we can proceed by deriving a compact formula for the calculation of moments (cf. Appendix A.2):

$$E[(1 - x^2)^k] = 2^{2k} \frac{\Gamma(2\varepsilon)}{\Gamma(2\varepsilon + 2k)} \left(\frac{\Gamma(\varepsilon + k)}{\Gamma(\varepsilon)}\right)^2. \tag{15}$$

Equipped with this result, we can invoke the Langevin equation (see Section 4.2) for the dynamics of x to characterize the ensuing dynamics of relative price changes. Neglecting changes of the fundamental value, we see that Eq. (7) simply leads to[9]

$$r(t, \Delta t) = x(t + \Delta t) - x(t) = -2a\Delta t x(t) + \sqrt{2b\Delta t(1 - x(t)^2)}\eta(t). \tag{16}$$

Note that an interesting consequence of Eq. (16) is that volatility is highest around $x = 0$ and decreases with an increasing absolute size of an emergent optimistic or pessimistic majority, $|x|$. While the unanimous subjective opinion of noise traders around $x = 1$ or -1 leads to minor fluctuations only in a situation with objective

---

[8]In the case $\varepsilon \gg 1$, the x-dependence in Eq. (13) can be written as

$$\exp[(\varepsilon - 1)\ln(1 - x^2)] \approx \exp[-\varepsilon x^2],$$

where the last step follows from the approximation

$$\ln(1 - x^2) \approx -x^2, \quad \text{for } |x| \ll 1.$$

The ε-dependent prefactor in Eq. (13) can be simplified using Stirling's approximation for the Gamma function

$$\Gamma(y + 1) \approx \sqrt{2\pi y}\, y^y e^{-y} \quad \text{for } y \gg 1.$$

[9]Note that the arbitrary interval of time τ of Eq. (7) is now replaced by the macro time step Δt. The difference will be explained in more detail in Section 5.1.

overvaluation or undervaluation, the higher uncertainty around $x = 0$ (and $p(t) = p_f$) is reflected in a higher level of volatility.

Eq. (16) allows to determine the expectations of even powers of the returns, r^{2k}, from those of x^{2k}. Given the symmetry of the unconditional distribution, all odd moments of returns are zero. In leading order Δt we get

$$E[r^{2k}] = (2b\Delta t)^k E[(1 - x^2)^k \eta^{2k}]. \tag{17}$$

The second moment and kurtosis[10] κ, for example, are given by

$$E[r^2] = \frac{4a\Delta t}{2\varepsilon + 1}, \quad \kappa = \frac{3}{\varepsilon(2\varepsilon + 3)}. \tag{18}$$

Eq. (18) gives us some first insights into the potential explanatory power of this simple asset pricing model. Namely, one immediately infers that excess kurtosis is positive and vanishes only for $\varepsilon \to \infty$ so that even in the case of a uni-model distribution, the herding mechanism would lead to a leptokurtic shape of returns which is in agreement with the ubiquitous empirical deviation of relative price changes from Gaussian behavior. Since $\varepsilon \to \infty$ would imply dominance of stochastic search, due to very large a compared to b, we observe that *it is the introduction of interpersonal communication that leads to the emergence of leptokurtic returns.* Leptokurtosis is, thus, a generic feature of this framework. However, it comes along with an exponential decay for large returns. Note, in fact, that Eq. (17) guarantees the theoretical existence of all even moments of the return distribution; our model, therefore, does not permit a 'true' scaling for the tail of the distribution, as observed in real data.

If we move to the extensive formulation of transition rates, Eq. (1), the above moments formula still hold, but would contain a dependency on the number of agents via the relation $b = \bar{b}/N$. Hence, the parameter ε in Eq. (18) has to be replaced by an N-dependent parameter $\varepsilon_N = Na/\bar{b}$. As can be inferred immediately, the excess kurtosis of the unconditional distribution fades away with an increasing number of agents in this alternative set-up. Our results for the second moment indicates that volatility itself is gradually reduced with an increasing number of agents so that the system dynamics approaches its deterministic limit for $N \to \infty$. It is, however, worthwhile to emphasize that the approach towards the Gaussian is faster (excess kurtosis decreases by a factor of N^{-2}, while the second moment only has N^{-1} dependency). Prior to the quasi-deterministic convergence to the mean-value limit the fluctuations, therefore, change from a distinctly non-Gaussian format with possibly large outliers to the more moderate Gaussian shape. In contrast, the non-extensive case lacks this self-averaging behavior altogether and maintains its interesting behavior for any size of the population.

---

[10]We define the kurtosis as

$$\kappa = \frac{E[r^4]}{E[r^2]^2} - 3.$$

4.5. Autocorrelation functions

Conditional properties of the population index x and returns can be obtained in various ways. One approach could be based on the general solution to the Fokker–Planck equation, which can be derived as a series expansion of Jacobi polynomials (see Alfarano, 2006). We can, then, provide a complete characterization of conditional moments for any transformation of returns. Alternatively, correlation functions could be determined recursively using the Langevin approximation to the Fokker–Planck equation. Although this approach has severe limitations in dealing with non-linear transformations of returns, it can be employed to compute the autocorrelation for raw and squared returns. To this end, we consider the discrete time steps $t = n\Delta t$ and define the recursion:

$$F_n = \mathrm{E}[x_t \cdot x_0],$$

for which the Langevin equation provides us with the difference equation:

$$F_n = (1 - 2a\Delta t)F_{n-1}.$$

Backward iteration leads to

$$F_n = F_0(1 - 2a\Delta t)^n,$$

with $F_0 = \mathrm{E}[x_0^2]$, which can be used to determine the autocorrelation $C_x(t)$ of x:

$$C_x(t) = (1 - 2a\Delta t)^{t/\Delta t}. \tag{19}$$

In the limit $\Delta t \to 0$, Eq. (19) becomes $C_x(t) = \exp(-2at)$. This result is not surprising given the linear drift function of the FPE (see Gardiner, 2003; Van Kampen, 1992).

The correlation of r_t can similarly be expressed in terms of F_n:

$$\mathrm{E}[r_t \cdot r_0] = \mathrm{E}[(x_{t+\Delta t} - x_t) \cdot (x_{\Delta t} - x_0)] = 2F_n - F_{n-1} - F_{n+1},$$

which leads to the expression (see Appendix A.3):

$$C_r(t) = -a\Delta t(1 - 2a\Delta t)^{t/\Delta t - 1} \approx -a\Delta t \exp(-2at). \tag{20}$$

In principle, the small negative correlation is at odds with the informational efficiency and martingale nature of financial prices. However, the small mean-reverting tendency of the population index and returns can easily be blurred by the noise level generated by the finiteness of the data sample and would not easily be detected if a were not too large (realistic values would, then, be of the order of magnitude of $a \approx 10^{-2}$).

One can proceed similarly in deriving the autocorrelation functions of higher moments. In order to compute $C_{r^2}(t)$ one starts with the relation[11]:

$$\mathrm{E}[r_t^2 \cdot r_0^2] = 4b^2\Delta t^2(1 - \mathrm{E}[x_t^2] - \mathrm{E}[x_t^2\eta_0^2] + \mathrm{E}[x_t^2x_0^2\eta_0^2]). \tag{21}$$

---

[11]Additional details are to be found in Appendix A.3.

After some efforts, we finally arrive at

$$C_{r^2}(t) = \frac{1}{4\varepsilon^2 + 6\varepsilon + 3} \exp(-2bt(2\varepsilon + 1)). \tag{22}$$

Obviously, $C_{r^2}(t) > 0$ for all $t > 0$ so that we have shown that volatility persistence, as the second universal time series property of returns, is also a natural outcome of this model. Note that due to the presence of the multiplicative term $a\Delta t$, the correlations of raw returns are $O(\Delta t)$ while the correlations of squared returns are $O(1)$, i.e. the latter are roughly one order of magnitude larger than the former. This finding squares well with the ubiquitous finding of large autocorrelation of squared returns (as a measure of the volatility) in daily data and the confinement of slight autocorrelations in raw returns to short intra-daily time horizons of about 30 min. Unfortunately, the negative sign of the autocorrelations of raw returns is not in accordance with empirical results for stock markets where we rather find evidence of significantly positive returns over 5–10 min followed by a slight mean-reverting tendency at time lags of about 15 min (cf. Bouchaud and Potters, 2000, p. 60). While the positive part is usually explained by information dissemination (which is absent from our model), the more controversial negative part would have to be attributed to overshooting and subsequent correction of prices.[12] If anything our model has the potential of explaining the second feature, i.e. correction after large positive or negative changes. Because of the multi-modal distribution of $x(t)$ and $p(t)$ in the case of dominant herding, $\varepsilon < 1$, the small mean-reverting tendency of raw returns is not equivalent to mean-reversion of prices towards the fundamental value. The latter can be more appropriately quantified via the *mean first passage time* (MFPT) between the modes $x = 1$ and -1. The next section provides analytical results for this quantity.

Again, if we move to the extensive formulation (1), the parameter ε has to be replaced by its N-dependent counterpart:

$$\varepsilon_N = N\frac{a}{b}.$$

Obviously, volatility dependence gets lost along with leptokurtosis with an increasing population size, as it is apparent from Eq. (22). The vanishing heteroscedasticity is reflected in a convergence to zero of the autocorrelation of squared returns. The loss of volatility dependence corresponds to the approach of the unconditional distribution towards the Gaussian, as illustrated in Fig. 2. The stylized facts, namely leptokurtic distribution of returns and volatility clustering measured as the non-vanishing correlation in the squared returns, do not survive when we increase the number of agents in the market. This means that the herding mechanism depending on the fraction of agents in the other group, at least in the present setting, might not be a suitable mechanism to explain the stylized facts as emerging properties of interactions among the traders. On the contrary, the model with dependence on the overall group occupation numbers, as formalized in Eq. (3),

---

[12]The literature on overshooting and mean-reversion in high-frequency stock market data is huge, cf. Lo and MacKinlay (1999, Chapter 5).
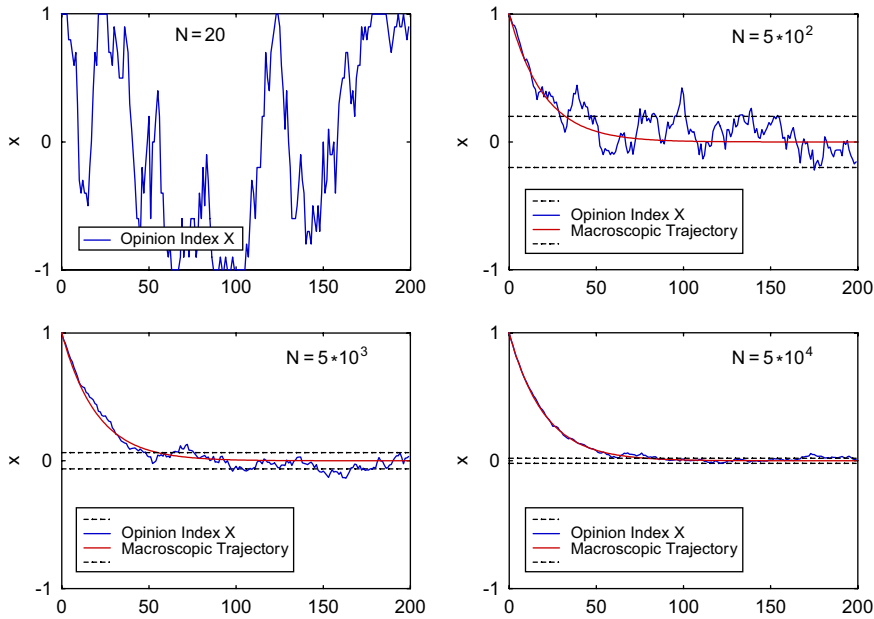
Fig. 2. The four panels show the emergence of the macroscopic skeleton of the system when governed by the extensive transition probabilities (1), for an increasing number of agents. Note the clear exponential decay of the trajectory of x for large population size, in accordance with the deterministic Eq. (31). The dashed lines represent the standard deviation of the equilibrium distribution of the variable x (see Appendix A.5). The underlying parameters are $N_0 = 20$, $\bar{b} = 1$ and $a_0 = 0.5/N_0$.

does not exhibit any relevant influence of the system size, and, therefore, gives identical results for markets with any number of participants.

4.6. Mean first passage time

The mean first passage time (MFPT hereafter) is defined as the time the system takes on average to switch from one state to another. Of major interest for the calculation of MFPTs are the modes and anti-modes $x = 0, -1$ and 1. As shown in Appendix A.4, in the case $0 < \varepsilon < 1$, the MFPT between the two modes $x = 1$ and -1 is given by

$$T(-1 \rightarrow 1) = T(1 \rightarrow -1) = \frac{1}{b} \frac{\pi}{1 - 2\varepsilon} \frac{\cos(\pi\varepsilon)}{\sin(\pi\varepsilon)}. \tag{23}$$

As can be seen, $T(\cdot)$ diverges for $\varepsilon \rightarrow 0$, which corresponds to the emergence of two absorbing states at $|x| = 1$ in the case of a vanishing autonomous switching probability. The lock-in at one of the extremal modes from which the dynamics would not return, since $a = 0$ and $\varepsilon = 0$, is mirrored in a divergence to infinity of its MFPT. For $\varepsilon \rightarrow 1$, divergence is obtained because in the uni-modal case obtained for $\varepsilon > 1, \pm 1$ are reached with a vanishing probability. Interestingly and maybe

running against the first intuition, the MFPT is not monotonic in ε, but has a minimum at $\varepsilon = 0.5$, i.e. $a = 0.5b$. At $\varepsilon = 0.5$, an increase of either the autonomous component or the herding propensity would lead to an increase of $T(\cdot)$. Again, plugging the N-dependent parameter ε_N into Eq. (23), we see that an increase of the size of the population eventually leads to the uni-modal case with $\varepsilon_N > 1$ triggering a divergence of $T(-1 \rightarrow 1)$.

Let us try a rough assessment of the magnitudes of the coefficients of our model that would lead to acceptable orders of magnitudes for the conditional and unconditional moments. To fix ideas, let us identify integer time steps by days in real-life markets. The dimension of both coefficients a and b is $[t^{-1}]$, i.e. 1/day. In order to find a plausible magnitude of kurtosis, we would have to choose values in the bi-modal regime, e.g. about $a = 0.1b$ for excess kurtosis in the range $\kappa \approx 10$. Choosing $a = 0.01$ or smaller and $b = 0.1$, autocorrelations of squared returns would be about 0.27 at daily horizons, while autocorrelation of raw returns would be almost invisible. The mean-passage time for switches between optimistic and pessimistic episodes would amount to $T = 120$ days (for comparison, for $a = 0.005$ and $b = 0.1$, we would obtain $\kappa = 19$, $C_{r^2} = 0.24$, and $T(-1 \rightarrow 1) = 220$ days, i.e. about one year). Roughly realistic leptokurtosis and volatility autocorrelation could, thus, come along with near random walk behavior of prices and long swings of overvaluation and undervaluation alternating at long time horizons. For a special case of our model with absence of autonomous switches ($a = 0$), Alfarano and Lux (2007) show that econometric analysis of simulated data would indicate random walk behavior for prices and apparently realistic scaling of volatility (Hurst coefficient significantly above 0.5) and the 'true' mean-reverting nature of the process would only become apparent at the time horizon T of the MFPT.

5. Different avenues for simulating the herding model

5.1. Occupation number approach

Following multi-particle simulations in statistical physics, different avenues exist for simulating the above model. The first, obvious choice would be a true microscopic simulation keeping track of the state of each individual agent and determining its switches over time by random number draws. Of course, the continuous-time framework would have to be simulated in discretized form. As an obvious restriction, these simulations would have to observe the condition that the sum of all transition probabilities has to remain smaller than unity, i.e. $\sum \pi(n \rightarrow n')\Delta t_0 \leqslant 1$ which translates into an upper bound for the usable micro time steps Δt_0 depending on the size of the population:

$$\Delta t_0 \leqslant \frac{2}{N(2a + bN)}, \tag{24}$$

since the herding component in both transition rates assumes its maximum at $n = N/2$. The maximum admissible time increment obviously decreases hyperbolically with the

population size, $\Delta t_0 \propto N^{-2}$. Microscopic simulations, therefore, become increasingly more time consuming with larger population size.

As an alternative, we could resort to simulating the aggregate outcome of the stochastic dynamics in terms of the population configuration which is summarized by the variable n, without taking into account the information on the history of every single individual. Rewriting (3) in terms of this state variable, one can see that transitions between states occur with probabilities:

$$\rho(n', t + \Delta t_0 | n, t) = \begin{cases} 2\dfrac{(N-n)(a+bn)}{N(2a+bN)} & \text{if } n' = n+1, \\[2mm] 2\dfrac{n(a+b(N-n))}{N(2a+bN)} & \text{if } n' = n-1, \\[2mm] b\dfrac{N^2 - 4n(N-n)}{N(2a+bN)} & \text{if } n' = n, \\[2mm] 0 & \text{otherwise.} \end{cases} \tag{25}$$

In (25), Δt_0 had been chosen such that it allows the highest efficiency of the macroscopic simulations, i.e. such that it minimizes the probability to observe no change in n. This is equivalent to using (24) as an equality. Of course, only the smallest possible change in n can be observed during the micro-step Δt_0 of the simulation. Similarly as in Kirman and Teyssière (2002), we may, therefore, introduce a distinction between micro time steps Δt_0 and macro time increments Δt in which many changes of n may be observed. The difference is, however, that we also allow the price process to operate along the micro time scale and that together with the change of the population configuration we also aggregate over many small price changes at the macro time scale Δt. To illustrate the dynamics and to provide a justification of a useful macro time scale, consider the following scenario: for $n = 0$, the system can evolve like follows: it may remain unchanged with probability $N/(N + 2\varepsilon)$, or it may change by one unit with the small probability $2\varepsilon/(N + 2\varepsilon)$, $\varepsilon = a/b$. Therefore, the average number of iterations of (25) needed to observe a move of one agent is approximately equal to N. To observe larger increments of n (as changes by one unit are negligible, in particular for the case of large N), we need a multiple of N to define the number of micro time steps that make up one macro interval. From Eq. (24), a sensible choice appears to be

$$\Delta t = \frac{b}{2} N^2 \Delta t_0. \tag{26}$$

In the simulations, one then iterates the process $(b/2)N^2$ times until one stores the current value of n as one realization at the macroscopic time scale Δt. The most interesting aspect of this approach is that it guarantees *invariance* of the dynamics of the macroscopic variable n with respect to the number of agents due to the flexibility of the relationship between the micro time scale and the macro time scale. Note that this does not imply that the overall time scale becomes arbitrary. Taking Eqs. (24) and (26) together, we see that we need a higher number of micro time steps Δt_0 for a larger population in order to capture the basic features of our model. Since we, then,

aggregate over a large number of micro steps according to Eq. (26) the chronological time captured in one macro step remains the same since the increase of the RHS of Eq. (26) exactly compensates for the decrease of Δt_0 in Eq. (24). This does not come as a surprise since we have already noticed that the drift and diffusion functions in the FPE (9) are independent of N. Fig. 1 provides an illustration of this feature in which we indeed observe no qualitative difference in the behavior of time series for different sizes of the population.

All in all, in order to have a meaningful micro-simulation algorithm for the Markov chain governed by Eq. (3) in the case of various population sizes, we have to introduce two different time scales. Note that the scaling of the macro-time with the number of individuals is non-linear. Again the non-linear dependence on N of the macro time scaling reflects the non-extensivity of the transition probabilities (3).

In the lower right-hand panel of Fig. 1, we also exhibit a simulation of the Langevin equation derived in Appendix A.6 for the non-extensive case. As we can observe, the Langevin dynamics (with added boundary conditions at the edges $x = -1$ and $x = +1$) is in good qualitative agreement with the 'true' micro-simulations shown in the other three panels.

5.2. Concentration approach

A very different behavior is observed when the system is described by extensive transition rates (1) in the limit $N \to \infty$. Let us illustrate the outcome using the reference parameters $a_0 = a \cdot N_0$ (we may assume that $a_0 = 0.5$) for the autonomous switching term, and $\bar{b} = 1$ for the herding parameter; let N_0 be the (small) reference value for the population size. The effective herding parameter (the constant in front of the quadratic term) is $\bar{b}/N$, for a market populated by N traders. The ratio between the autonomous term and the herding parameter is, then, defined by

$$\varepsilon_N = \frac{N}{N_0} \frac{a_0}{\bar{b}} = \frac{N}{2N_0}. \tag{27}$$

Assume that when the number of agents N coincides with the reference size N_0, the market falls into the regime dominated by herd behavior, with a bimodal probability density function for x. If one now increases the number of agents beyond the starting level N_0 and, at the same time, leaves unchanged the other parameters, the equilibrium distribution will undergo a transition from a bimodal shape to a narrowly peaked distribution around the average value $E[x] = 0$. The linear dependence of ε_N on the number of agents N, cf. Eq. (27), brings about a convergence of the equilibrium distribution $\omega_e(x)$ to a Gaussian. Since $\varepsilon_N \to \infty$ for $N \to \infty$, the results derived in Eq. (14) can be applied. However, because of the dependence of b on N of the diffusion term, the N^{-1} term needs to be taken into account when deriving the stationary distribution (cf. Appendix A.5). Due to the N-dependence of ε_N, we also find that the variance of the Gaussian approximation depends inversely on N:

$$\sigma^2 = \frac{1}{2(N/2bN_0 - 1)} \approx \bar{b} \frac{N_0}{N}. \tag{28}$$

Convergence to the Gaussian with vanishing fluctuations is illustrated in Fig. 2. The transition from a bimodal to a unimodal distribution implies the existence of a critical number of agents N_C, which we might conveniently define by the relation[13]

$$\varepsilon_{N_C} = \frac{N_C}{N_0}\frac{a_0}{\bar{b}} = 1, \quad N_C = N_0 \frac{\bar{b}}{a_0}. \tag{29}$$

The critical size of the market, N_C, might be used to distinguish between two regimes. An interacting-agent regime, where the random meetings among agents have a relevant impact on the opinion index dynamics, with a resulting bimodal distribution, and an individualistic regime, where the market behaves like a collection of independent agents and the interactive component has a negligible influence. The resulting Gaussian probability distribution for the case $N \gg N_C$ is, in fact, the asymptotic distribution for a collection of independent agents. In the case of extensive transition rates and a large number of components the system can be described by a deterministic equation for the mean and a 'small' *superimposed* noise perturbation (see Gardiner, 2003 and, in particular, Aoki, 1996, 2002 for several applications of this formalism to economics). It is well known that the following equation holds for the mean value $\bar{x}(t)$ of a stochastic process with linear drift function:

$$\frac{d}{dt}\bar{x}(t) = A(\bar{x}(t)), \tag{30}$$

where $A(\cdot)$ is given by Eq. (11) in our case. Eq. (30) simply states that the mean of $\omega(x, t)$ follows a differential equation of first order, whose unique and fully deterministic solution is given by

$$\bar{x}(t) = \bar{x}_0 \exp(-2\varepsilon bt), \tag{31}$$

where $\bar{x}_0$ is the mean value of the starting distribution $\omega_0(x, 0)$. As a result, $\bar{x}$ converges to its unconditional value $E[x] = 0$ with an exponential rate. This macroscopic equation predicts no evolution for $\bar{x}$, if the starting value $\bar{x}_0$ coincides with the unconditional expected value (which, in fact, was the case in all the simulations reported in Fig. 1). Nevertheless, we observe a non trivial dynamics of opinion index x in our simulations with non-extensive transition rates for every size of the population, which gives rise to the interesting statistical properties of our artificial market. The contrast between the macroscopic approximation of Eq. (30) and the more complete description in terms of the FPE is striking. In this case, endogenous fluctuations of the system can be neglected only at the cost of a meaningless description of its time-evolution. On the contrary, the approach to the equilibrium distribution for the model governed by intensities rather than occupation number is well described by Eq. (31), as illustrated in Fig. 2. In the case of extensive transition probabilities and a large number of agents, the aggregate behavior of the system can be described by a deterministic equation for the mean and vanishing fluctuations around $\bar{x}(t)$ for $N \to \infty$.

---

[13]Note that in the case $\varepsilon_N = 1$, the equilibrium distribution is uniform.

6. Local vs non-local interactions

Extensive transition rates describe systems whose evolution is driven by 'local' mechanisms. In order to justify this assertion, namely the identification of the extensive transition probabilities with the local character of the interactions, we will provide some qualitative arguments. Let us start considering a system composed of a large number of independent components. To describe the asymptotic behavior of such a system, we may invoke the CLT, expecting Gaussian fluctuations around the mean of some aggregate quantities. The ratio of the amplitude of the fluctuations relative to their mean, then, will converge to zero. More precisely, we expect a dependence of the type $O(1/\sqrt{N})$. The convergence to a well-defined mean with vanishing fluctuations is the basic justification for a deterministic approach to the description of models composed of independent 'particles' in terms of macroscopic observables.

This scenario can be generalized to include some sort of interactions among the constituents, which typically generate correlations among the different particles. If the correlations are 'weak' enough, we can still apply the CLT and arrive at the Gaussian regime for the fluctuations as for independent particles, with a well-defined mean for aggregate quantities, which can still be described by deterministic laws. Here the term 'weak' stands for a limited number of constituents involved in the interaction mechanism. On the contrary, the mechanism has a 'non-local' or, in extreme cases like our simple herding model, global character if the range of the correlations involves a macroscopic fraction of elements. The long-range correlations prevent the use of the CLT and, therefore, we cannot expect the fluctuations to eventually converge to an asymptotic Gaussian regime. Spatial or temporal correlations in the level of fluctuations might be observed for any system size. Extensive transition rates guarantee convergence to the Gaussian regime, while non-extensive systems *might* or *might not* exhibit convergence to the Gaussian regime.

The somewhat counterintuitive result here is that, on one hand, pairwise meetings and interactions in line with the original philosophy of random meetings in space – Eq. (3) – lead to a global character of the system with the dynamics preserving the interesting features independent on the system size. Dependence of the transitions on overall fractions of groups within the population, on the other hand, lead to a more 'local' character of the dynamics and loose all interesting features with an increasing number of agents, ultimately converging to a Gaussian regime. The solution of this puzzle is that in the non-extensive formalization the number of neighbors per constituent increases with an increasing population. It means that every agent is coupled with the same fraction of other agents for any system size, which guarantees that the interesting properties of the fluctuations are independent on N. Hence, the *intensity of the interpersonal coupling* is preserved via Eq. (3) for varying numbers of agents N. In the extensive framework of Eq. (1), in contrast, the herding component looses its importance vis-à-vis the autonomous switches when increasing the number of agents, because the coupling intensity effectively decreases with system size.

The results on the striking difference between extensive and non-extensive transition probabilities in our herding model might also give us a clue to the dependency on population size of other models. With hindsight one may argue that

certain models that are designed in the limit for large N might have built-in long-range coupling of agents corresponding to our non-extensive framework. Percolation models are a case in question since they are usually designed so that the average number of other agents to whom an agent is connected has a finite limit. However, in the models of Cont and Bouchaud (2000) and Stauffer and Sornette (1999) the returns distribution is derived from the distribution of clusters in the random network, which already implies a kind of aggregation over agents. The object of interest is the cluster and every agent is somewhat identical within a cluster. Moreover, these models show interesting properties just at the critical connectivity probability (which is N-dependent) or in a region very narrowly centered around it. Using our terminology, every agent will be connected with a macroscopic fraction of the entire population right at the percolation threshold, which somehow corresponds to our non-extensive formalization of the transition rates.

The Cont and Bouchaud framework, therefore, somehow establishes non-extensivity by its assumption of a level of connectivity in the vicinity of the percolation threshold (over which an infinite cluster spanning the entire lattice appears). What the effect of the increase of the population would be in the models of Iori (2002) and Bornholdt (2001) is not clear. In contrast to our framework, interactions in these models are spatially organized with agents' behavior being influenced by that of other agents within a well-defined *neighborhood*. Whether the degree of interactions remains the same with an increase of the population, then, depends on how the definition of the neighborhood is adapted and how the network of connections is designed. Current research by the authors attempts to explore what type of network design would be compatible with the non-extensive case of overall mean-field effects analyzed in the previous sections.

In other type of models, it is even more obscure how they should be classified within the extensive/non-extensive dichotomy. For example, according to the results reported in Lux and Schornstein (2005), artificial markets with strategies evolved by genetic algorithms might fall into the class of extensive models. With an increasing number of agents, the genetic operators of selection, recombination and mutation might, therefore, amount to a reduced intensity of interpersonal coupling. How to modify this structure in a way to preserve non-extensive coupling is not obvious and would probably amount to a new design of the GA algorithms.[14] Since the difference between extensive and non-extensive formalization is maybe rather opaque in broad classes of models, our findings can also be seen as a caveat for agent-based models: 'interesting dynamics' as the emergence of stylized facts could simply be a finite size effect. While simulations with small samples of agents might look impressive, the behavior of the same system with a larger number of agents could turn out to be quite disappointing. On the other hand, the ubiquity of the stylized facts with their non-self-averaging features in empirical data speaks in favor of prevalence of non-extensive mechanisms of interaction in real-life markets.

---

[14]Note that a higher correlation intensity of individual GA chromosomes might not be desirable from the point of view of their original applications in optimization of complex functions. After all, it is the *independent* simultaneous exploration of different parameters sets that makes GAs a successful tool.

7. Conclusions

The aim of the present paper was to contribute to a better understanding of the statistical properties of agent-based models of financial markets. To this end, we attempted a thorough analysis of different types of herding models in the tradition of Kirman's seminal ant model. Our novel insights are the following:

(i) We were able to derive closed-form solutions for various moments as well as the autocorrelations of raw and squared returns. Our results demonstrate that both leptokurtosis of returns and temporal dependence of volatility are generic features of this model and are intimately linked to the herding component. This underscores the importance of boundedly rational behavior as a potential explanation of the stylized facts. Admittedly, our model did not produce power laws (hyperbolic decay) of the cumulative distribution nor of the autocorrelation function. However, in a companion paper (Alfarano and Lux, 2007), we also show that pseudo-empirical analysis of time series for such a model could easily produce 'spurious' or 'apparent' power laws, which is a typical outcome within a pre-asymptotic regime with a length of the time series below some threshold value.

(ii) We have also investigated the dependency of the 'interesting' dynamics and stylized facts on system size, revisiting earlier findings of trivial large N limits in similar models. As it turns out, trivial or non-trivial behavior of large markets depend on the exact formalization of transition rates. Ironically, a literal interpretation of pairwise interactions leads to some sort of 'global' coupling of agents, which preserves all interesting features of the model for any system size. On the contrary, a more conventional formalization of the transitions rates in terms of concentrations leads to a convergence to uninteresting (and unrealistic) Gaussian dynamics in the limit of large markets.

Appendix A

A.1. Derivation of the Fokker–Planck equation

In this Appendix, we derive the Fokker–Planck equation (9) as a second-order Taylor approximation of our population dynamics governed by Eq. (3). In order to do so, we introduce the intensive variable

$$x = 2n/N - 1, \tag{A.1}$$

that can be treated as a continuous quantity, if the total number of agents N is large enough. The relation between the transition rates (3), expressed in terms of the variables n or x, is given by the following formula:

$$\pi_n^{\pm} = \frac{N^2}{4}\pi_x^{\pm}.$$

Note that the probabilities are invariant under the transformation (A.1), therefore, $\bar{\omega}_n = \bar{\omega}_x$. The transition rates (3) are now functions[15] of x:

$$\pi_x^+ = (1 - x)(2a/N + b(1 + x)),$$
$$\pi_x^- = (1 + x)(2a/N + b(1 - x)).$$

Let us introduce the 'step' operators $\mathbf{E}$ and $\mathbf{E}^{-1}$, using the notation of Van Kampen (1992). Their effects on an arbitrary function $f(n)$ are, respectively, to decrease or increase its integer argument n by one unit. Formally

$$\mathbf{E}[f(n)] = f(n + 1), \quad \mathbf{E}^{-1}[f(n)] = f(n - 1).$$

With the aid of the operators $\mathbf{E}[\cdot]$ and $\mathbf{E}^{-1}[\cdot]$, the Master equation (8) for the one-step process can be compactly rewritten in terms of the variable x:

$$\frac{\partial}{\partial t}\bar{\omega}_x = (\mathbf{E} - 1)\left[\pi_x^- \frac{N^2}{4}\bar{\omega}_x\right] + (\mathbf{E}^{-1} - 1)\left[\pi_x^+ \frac{N^2}{4}\bar{\omega}_x\right], \tag{A.2}$$

where we skip the obvious time dependence of $\bar{\omega}_x$. Considering the continuous limit for x, the probability $\bar{\omega}_x$ has to be 'converted' into a probability density:

$$\bar{\omega}_n = \bar{\omega}_x \equiv \omega_x \frac{2}{N},$$

where ω_x is a continuous function of x. The probability density might be alternatively defined by the limit:

$$\omega_x = \lim_{N \to \infty} \frac{\bar{\omega}_x}{\Delta x} = \lim_{N \to \infty} \frac{N}{2}\bar{\omega}_x.$$

Eq. (A.2) can be rewritten as

$$\frac{\partial}{\partial t}\omega_x = \frac{N^2}{4}(\mathbf{E} - 1)[\pi_x^- \omega_x] + \frac{N^2}{4}(\mathbf{E}^{-1} - 1)[\pi_x^+ \omega_x].$$

Since in the limiting case, the step operator acts on continuous functions, we can perform the Taylor expansion:

$$\mathbf{E}[f(x)] = f(x + \Delta x) = f(x) + \Delta x \frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{1}{2}\Delta x^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2}f(x) + \mathrm{o}(\Delta x^2), \tag{A.3}$$

---

[15]The new transitions have to be read as $\pi_x^{\pm} = \pi(x \to x \pm 2/N)$.

where $\Delta x = 2/N$. The step operator can be approximated by the following expansion:

$$\mathbf{E} = 1 + \Delta x \frac{\partial}{\partial x} + \frac{1}{2} \Delta x^2 \frac{\partial^2}{\partial x^2} + o(\Delta x^2). \tag{A.4}$$

Note that the derivatives in Eq. (A.3) are replaced by the partial derivatives in Eq. (A.4), since the function f, which the operator is applied to, might have more than one variable. The expansion for the operator $\mathbf{E}^{-1}$ is simply obtained from the previous formula replacing Δx with $-\Delta x$. Using the expansion (A.4) for $\mathbf{E}$ and $\mathbf{E}^{-1}$ up to second order,[16] we end up with

$$\frac{\partial}{\partial t} \omega_x = \frac{N^2}{4} \left\{ -\Delta x \frac{\partial}{\partial x} [(\pi_x^+ - \pi_x^-)\omega_x] + \frac{1}{2} \Delta x^2 \frac{\partial^2}{\partial x^2} [(\pi_x^+ + \pi_x^-)\omega_x] \right\}.$$

The $N^2/4$ factor in front of the equation disappears, and we arrive finally at the Fokker–Planck equation:

$$\frac{\partial}{\partial t} \omega_x = -\frac{\partial}{\partial x} A(x)\omega_x + \frac{1}{2} \frac{\partial^2}{\partial x^2} D(x)\omega_x, \tag{A.5}$$

where the drift and diffusion functions are, respectively, given by

$$A(x) = \frac{N}{2}(\pi_x^+ - \pi_x^-) = -2ax, \tag{A.6}$$

and

$$D(x) = \pi_x^+ + \pi_x^- = 2b(1 - x^2) + \frac{4a}{N}. \tag{A.7}$$

For finite b the term proportional to a/N can be neglected.

It should be emphasized that the Fokker–Planck equation is an approximation to the continuous limit of our intrinsically discrete stochastic process governed by the transition probabilities (3), up to terms proportional to $1/N^2$. Negligence of the granularity term in Eq. (A.7) would only be innocuous as long as $\Delta x \ll |1 \pm x|$, i.e. not too close to the boundaries. Closer to the boundaries, the discreteness of the variable x cannot be neglected any longer and the term $4a/N$ should be taken into account. More precisely, the granularity term can be neglected if $2b(1 - x^2) \gg (4a/N)$, which is equivalent to the constraint $\varepsilon/2N - 1 \ll x \ll 1 - \varepsilon/2N$. However, given the dependence of Δx on N, this 'critical' region might be set to any arbitrarily small interval by simply increasing the total number of agents N.

A.2. Unconditional distribution of x and returns

The following textbook formula allows to compute the equilibrium distribution $\omega_e(x)$, given in Eq. (13) in the main text, if we know the diffusion and

---

[16]Higher terms would be $o(1/N^2)$, and are, therefore, negligible. The reason to include the second order will be clarified later.

drift functions:

$$\omega_e(x) = \frac{K}{D(x)} \exp\left(\int^x \frac{2A(y)}{D(y)} \, \mathrm{d}y \right),$$ (A.8)

(see for instance Van Kampen, 1992). The first step consists in solving the following integral:

$$\int \frac{2A(x)}{D(x)} \, \mathrm{d}x = -\frac{2a}{b} \int \frac{x}{1-x^2} \, \mathrm{d}x,$$

where we take the drift and diffusion functions from Eqs. (A.6) and (A.7), neglecting the last term in $D(x)$. With the aid of $\varepsilon = a/b$ and a little effort, the solution is

$$\int \frac{2A(x)}{D(x)} \, \mathrm{d}x = \varepsilon \ln(1-x^2),$$

that plugged into (A.8), gives the equilibrium distribution $\omega_e(x)$:

$$\omega_e(x) = K(1-x^2)^{\varepsilon-1}.$$ (A.9)

The constant K is computed via the normalization condition:

$$K \int_{-1}^{+1} \omega_e(x) \, \mathrm{d}x = 1.$$ (A.10)

Recalling the definition of the Beta function

$$B(\varepsilon_1, \varepsilon_2) := \int_0^1 z^{\varepsilon_1-1}(1-z)^{\varepsilon_2-1} \, \mathrm{d}z = \frac{\Gamma(\varepsilon_1)\Gamma(\varepsilon_2)}{\Gamma(\varepsilon_1 + \varepsilon_2)},$$

we obtain for symmetric coefficients $\varepsilon_1 = \varepsilon_2$ and with the substitution $z = \frac{1}{2}(1+x)$:

$$\int_{-1}^{+1} (1-x^2)^{\varepsilon-1} \, \mathrm{d}x = 2^{2\varepsilon-1} \frac{\Gamma(\varepsilon)^2}{\Gamma(2\varepsilon)}.$$

Solving the integral (A.10), we arrive at $K = \Gamma(2\varepsilon)/(2^{2\varepsilon-1}\Gamma(\varepsilon)^2)$.

Using Eq. (A.9) we can derive a compact formula for the calculation of moments:

$$\mathrm{E}[(1-x^2)^k] = \frac{\Gamma(2\varepsilon)}{2^{2\varepsilon-1}\Gamma(\varepsilon)^2} \int_{-1}^1 (1-x^2)^{\varepsilon-1}(1-x^2)^k \, \mathrm{d}x.$$ (A.11)

Via the definition of the Beta function, Eq. (A.11) becomes

$$\mathrm{E}[(1-x^2)^k] = 2^{2k} \frac{\Gamma(2\varepsilon)}{\Gamma(2\varepsilon+2k)} \left[\frac{\Gamma(\varepsilon+k)}{\Gamma(\varepsilon)} \right]^2$$

which is used in the main text for computing the unconditional moments of returns.

A.3. Autocorrelation functions

Closed-form approximate solutions for the auto-covariances of transformations of the variable x may be obtained via different approaches.

Starting from the Fokker–Planck equation (A.5), we can solve for the conditional probability $\omega(x, t|x_0, 0)$, given the initial distribution $\omega_0(x_0)$. It turns out that $\omega(x, t|x_0, 0)$ can be expanded in Jacobi polynomials, which allows to compute the autocorrelation for any transformation $f(x)$ of the variable x in series of these polynomials (see Alfarano, 2006).

Alternatively, correlation functions can be determined recursively using the Langevin approximation of the Fokker–Planck equation. We proceed here using this second approach. For discrete time steps $t = n\Delta t$, we consider first

$$F_n = E[x_t \cdot x_0], \quad F_0 = E[x_t^2].$$

From the Langevin equation (12), we obtain a recursion for F_n:

$$F_n = (1 - 2a\Delta t)F_{n-1}.$$

Inserting the solution

$$F_n = F_0(1 - 2a\Delta t)^n$$

into the autocorrelation formula of x, we get

$$C_x(t) = (1 - 2a\Delta t)^{t/\Delta t} \simeq \exp(-2at),$$

where the exponential form is exact in the limit $\Delta t \to 0$.

The correlation of $r(t)$ can also be expressed in terms of F_n:

$$E[r_t \cdot r_0] = E[(x_{t+\Delta t} - x_t) \cdot (x_{\Delta t} - x_0)] = 2F_n - F_{n-1} - F_{n+1}.$$

Therefore for $n > 0$, the autocorrelation takes the form:

$$C_r(t) = -\frac{F_{n+1} + 2F_n - F_{n-1}}{E[(x_{\Delta t} - x_0)^2]}. \tag{A.12}$$

Using $E[(x_{\Delta t} - x_0)^2] = 4a\Delta t F_0$ and inserting F_n into (A.12), one proves (20) given in the main text.[17] Eq. (20) can be alternatively derived by means of the following approximation:

$$F_{n+1} - 2F_n + F_{n-1} = \Delta t^2 \frac{\partial^2 F_n}{\partial t^2} = (2a\Delta t)^2 \exp(-2at)F_0. \tag{A.13}$$

Plugging Eq. (A.13) into Eq. (A.12), we arrive at the same result.

Computation of C_{r^2} is slightly more complicated. In order to compute the auto-correlation function of the second moment, we use Eq. (12) to substitute for the auto-covariances of r_t^2:

$$\begin{aligned}
E[r_t^2 \cdot r_0^2] \\
&= E[(x_{t+\Delta t} - x_t)^2 \cdot (x_{\Delta t} - x_0)^2] \\
&= E[(-2a\Delta t x_t + \sqrt{2b\Delta t(1 - x_t^2)}\eta_t)^2 \cdot (-2a\Delta t x_0 + \sqrt{2b\Delta t(1 - x_0^2)}\eta_0)^2] \\
&= 4b^2\Delta t^2 E[1 - x_t^2\eta_0^2 - x_0^2\eta_0^2 + x_t^2 x_0^2\eta_0^2] + o(\Delta t^2). \tag{A.14}
\end{aligned}$$

---

[17]For F_{n+1} and F_{n-1} we use the Taylor expansion $e^x = 1 + x + x^2/2 + o(x^2)$.

It is obvious, that expression (A.14) is identical in leading order Δt^2 to Eq. (21) in the main text. It can be easily derived from Eq. (15) that $E[x_t^2] = 1/(2\varepsilon + 1)$. Then, from Eq. (12), we can work out the following relation:

$$E[x_t^2\eta_0^2] = (1 - 2b\Delta t(2\varepsilon + 1))E[x_{t-\Delta t}^2\eta_0^2] + 2b\Delta t \quad \text{for } t > 1. \tag{A.15}$$

Defining $H_n = E[x_{n\Delta t}^2\eta_0^2]$ and $H_0 = E[x_0^2]$, the recursive approach leads to:

$$H_1 = E[x_{\Delta t}^2\eta_0^2] = (1 - 2b\Delta t(2\varepsilon + 3))E[x_0^2] + 6b\Delta t, \tag{A.16}$$

and

$$H_n = E[x_{n\Delta t}^2\eta_0^2] = (1 - 2b\Delta t(2\varepsilon + 1))H_{n-1} + 2b\Delta t \quad \text{for } n \geqslant 2. \tag{A.17}$$

Solving this recursion yields:

$$H_n = 8a\Delta t E[x_0^2](1 - 2b\Delta t(2\varepsilon + 1))^{n-1} + E[x_0^2] + o(\Delta t). \tag{A.18}$$

In order to obtain a closed-form solution for $E[x_t^2 \cdot x_0^2 \cdot \eta_0^2]$, we use an analogous recursive approach. Defining $G_n = E[x_{n\Delta t}^2 x_0^2\eta_0^2]$, we get

$$G_1 = (1 - 2b\Delta t(2\varepsilon + 3))E[x_0^4] + 6b\Delta t E[x_0^2], \tag{A.19}$$

and

$$G_n = (1 - 2b\Delta t(2\varepsilon + 1))G_{n-1} + 2b\Delta t E[x_0^2] \quad \text{for } n \geqslant 2. \tag{A.20}$$

Solving this difference equation, we arrive at the following expression:

$$G_n = (1 - 2b\Delta t(2\varepsilon + 1))^{n-1}(G_0 - E[x_0^2]^2) + E[x_0^2]^2 + o(\Delta t), \tag{A.21}$$

with $G_0 = E[x_0^4]$. The fourth moment of x_0 can be obtained from Eq. (15):

$$G_0 = E[x_0^4] = \frac{3}{(2\varepsilon + 3)(2\varepsilon + 1)}.$$

Note that the first term on the RHS of Eq. (A.18) is of smaller order than the remaining components of both Eq. (A.18) and Eq. (A.21) and can be relegated to the residual $o(\Delta t^2)$ when inserting the results into Eq. (A.14). In order to solve for $C_{r^2}(t)$, we need some further ingredients:

$$\begin{aligned} E[r_0^2] &= E[(x_{\Delta t} - x_0)^2] = E[(-2a\Delta t x_0 + \sqrt{2b\Delta t(1 - x_0^2)}\eta_0)^2] \\ &= 2b\Delta t E[(1 - x_0^2)\eta_0^2] + o(\Delta t) \simeq 2b\Delta t(1 - E[x_0^2]). \end{aligned}$$

We, therefore, get for the numerator of $C_{r^2}(t)$ in leading order of Δt:

$$E[r_t^2 \cdot r_0^2] - E[r_0^2]^2 = 4b^2\Delta t^2(1 - 2b\Delta t(2\varepsilon + 1))^{n-1}(E[x_0^4] - E[x_0^2]^2),$$

with $n = t/\Delta t$. Finally, in order to pin down the denominator, consider

$$
\begin{aligned}
\mathrm{E}[r_0^4] &= \mathrm{E}[(-2a\Delta t x_0 + \sqrt{2b\Delta t(1 - x_0^2)}\eta_0)^4] \\
&= \mathrm{E}[(2b\Delta t)^2(1 - x_0^2)^2\eta_0^4] + o(\Delta t^2) \\
&\simeq (2b\Delta t)^2 \mathrm{E}[(1 - 2x_0^2 + x_0^4)\eta_0^4] \\
&= 3(2b\Delta t)^2 \frac{4\varepsilon(\varepsilon + 1)}{(2\varepsilon + 1)(2\varepsilon + 3)}.
\end{aligned}
$$

It, then, turns out that

$$
\mathrm{E}[r_0^4] - \mathrm{E}[r_0^2]^2 = (2b\Delta t)^2 \frac{(4\varepsilon^2 + 6\varepsilon + 3)4\varepsilon}{(2\varepsilon + 1)^2(2\varepsilon + 3)}.
$$

One solves for

$$
C_{r^2}(t) = \frac{\mathrm{E}[r_{t+\Delta t}^2 \cdot r_{\Delta t}^2] - \mathrm{E}[r_t^2]^2}{\mathrm{E}[r_t^4] - \mathrm{E}[r_t^2]^2},
$$

and obtains Eq. (22).

A.4. Mean first passage time

To compute the MFPT $T(x_0 \to x_2)$ one integrates the Fokker–Planck equation with a reflecting boundary condition at $x_0 = -1$ and an absorbing boundary condition at $x_2 = +1$. Evoking the following general formula (see Gardiner, 2003) for MFPTs, we can try to find a closed-form solution for this quantity as well

$$
T(x_0 \to x_2) = \int_{x_0}^{x_2} \mathrm{d}x (\psi(x))^{-1} \int_x^{x_1} \frac{2\psi(y)}{D(y)} \, \mathrm{d}y,
$$

with the auxiliary $\psi(x)$ function being defined as

$$
\ln \psi(x) = \int_0^x \frac{2A(y)}{D(y)} \, \mathrm{d}y.
$$

Note first that with drift and diffusion terms $A(x) = -2ax$ and $D(x) = 2b(1 - x^2)$, according to our previous computations $\psi(x) = (1 - x^2)^\varepsilon$. Taking stock of this result, we obtain

$$
T(-1 \to 1) = \int_{-1}^1 (1 - x^2)^{-\varepsilon} \left(\int_{-1}^x \frac{2(1 - y^2)^\varepsilon}{2b(1 - y^2)} \, \mathrm{d}y \right) \mathrm{d}x.
$$

Because of the symmetry of $(1 - y^2)^{\varepsilon-1}$ we can solve the double integral in a few steps:

$$
bT(-1 \to 1) = \int_{-1}^1 (1 - x^2)^{-\varepsilon} \left(\int_{-1}^0 (1 - y^2)^{\varepsilon-1} \, \mathrm{d}y - \int_0^{-x} (1 - y^2)^{\varepsilon-1} \, \mathrm{d}y \right) \mathrm{d}x,
$$

where the second integral is identically zero since the integrand is an odd function integrated in a symmetric interval. We arrive, then, at

$$
\begin{aligned}
bT(-1 \to 1) &= \frac{1}{2} \int_{-1}^{1} (1 - x^2)^{-\varepsilon} \, dx \int_{-1}^{1} (1 - y^2)^{\varepsilon-1} \, dy \\
&= \frac{1}{2} \int_{-1}^{1} 2^{2\varepsilon-1} \frac{\Gamma(\varepsilon)^2}{\Gamma(2\varepsilon)} (1 - x^2)^{-\varepsilon} \, dx \\
&= \frac{1}{2} \cdot \frac{\Gamma(\varepsilon)^2 \cdot \Gamma(1 - \varepsilon)^2}{\Gamma(2\varepsilon) \cdot \Gamma(2 - 2\varepsilon)}.
\end{aligned}
$$

The identity $\Gamma(1 - x) \cdot \Gamma(x) = \pi / \sin(\pi x)$ finally leads to

$$
T(-1 \to 1) = \frac{1}{b} \frac{\pi}{(1 - 2\varepsilon)} \frac{\cos(\pi\varepsilon)}{\sin(\pi\varepsilon)}.
$$

A.5. Convergence to the Gaussian in the extensive case

We have shown that in the case of local interactions the drift and diffusion terms are given by Eq. (11), where $a = 0.5/N_0$. Given the N-dependence of b, the $1/N$ correction to the diffusion cannot be neglected any longer, and both terms have to be taken into account in deriving the equilibrium distribution. For the number of agents N much higher than the benchmark value N_0, the equilibrium distribution converges to a Gaussian. Using Eq. (A.8), we obtain:

$$
\omega_e(x) \propto \left(1 - x^2 + \frac{2a}{\bar{b}} \right)^{\varepsilon_N - 1},
$$

where $\varepsilon_N = N/2\bar{b}N_0$. The previous expression can be rewritten as:

$$
\omega_e(x) \propto \exp\left((\varepsilon_N - 1) \ln\left(1 - x^2 + \frac{2a}{\bar{b}} \right) \right).
$$

Recalling Eq. (14), the equilibrium distribution is well approximated by a Gaussian with mean zero and variance:

$$
\sigma^2 = \frac{1}{2(N/2\bar{b}N_0 - 1)} \approx \frac{\bar{b}N_0}{N}.
$$

A.6. Langevin equation

The basic idea of the Langevin approach is to find an appropriate time interval, which one might call a *macro-interval*, for which the *conditional* distribution of the discrete variable x is well approximated by a Gaussian. In the following we sketch a derivation of the Langevin equation as an approximation to the discrete process (3)

based on some heuristic arguments. Let us start with the identity:

$$x_{t+\Delta t} = x_t + \sum_i^M \eta_{t+i\Delta t_0},\tag{A.22}$$

where $\Delta t = M\Delta t_0$ and $\eta(t + i\Delta t_0)$ might take the values in the finite set $S \equiv \{0, 2/N, -2/N\}$; Δt_0 is the elementary time interval for a possible change of strategy by one agent. The previous decomposition is based on a separation of the dynamics of x into two different time scales: a micro time scale Δt_0 and a macro time scale Δt. During the micro interval Δt_0, the variation of x is constrained to the numerable set S. If we concentrate on the aggregate variation Δx during the macro-time interval Δt, we sum up over many of these elementary increments, loosing the information on the fine structure of the dynamics of x (i.e. the single movements of the distinguishing individuals). Our aim is exactly to give up this too detailed perspective in favor of a more general macroscopic description of the system. In order to do so, we have to find a proper time scale, long enough to aggregate the tiny details, and not too long to lose relevant information. This perspective is sometimes called a mesoscopic approach, as it is somehow between the micro and macroscopic approach (Ramsey, 1996). We have, then, to find the proper value for the aggregation level M. The random variable $\eta(t + (i + 1)\Delta t_0)$ can take just three values with the following probabilities:

$$\eta_{i+1} = \begin{cases} +\dfrac{2}{N} & \text{with prob } p_i\Delta t_0 = (1 - x_i)\left(\dfrac{2\varepsilon}{N} + 1 + x_i\right)bN^2\Delta t_0, \\[2mm] 0 & \text{with prob } 1 - (p_i + q_i)\Delta t_0, \\[2mm] -\dfrac{2}{N} & \text{with prob } q_i\Delta t_0 = (1 + x_i)\left(\dfrac{2\varepsilon}{N} + 1 - x_i\right)bN^2\Delta t_0, \end{cases}\tag{A.23}$$

where we explicitly take into account the dependence on the previous value of x (here $x_i = x(t + i\Delta t_0)$ and $\eta_{i+1} = \eta(t + (i + 1)\Delta t_0)$). For notational convenience, we label the time step by i. The mean μ_{i+1} and variance σ_{i+1} of η_{i+1} are given by, respectively:

$$\mu_{i+1} = \frac{\Delta t_0}{N}(p_i - q_i),\tag{A.24}$$

$$\sigma_{i+1}^2 = \mathrm{E}[\eta_{i+1}^2] - \mu_{i+1}^2 = \frac{\Delta t_0}{N^2}(p_i + q_i) + o\left(\frac{\Delta t_0}{N^2}\right).\tag{A.25}$$

Following the Central Limit Theorem (CLT), a sum of M independent random variables, drawn from a common distribution, with mean μ and finite variance σ^2, converges to a normal with mean $M\mu$ and variance $M\sigma^2$. We can, therefore, approximate the sum in (A.22) by a normally distributed random variable. The requirement of a finite variance is certainly fulfilled by the distribution of the variable η. However, the iid-ness assumption does not strictly hold. The probabilities $p\Delta t_0$ and $q\Delta t_0$ depend on x, which dynamically changes with η itself. We have to impose a further condition on x: it should not vary 'too much' during the time interval Δt, in such a way that it can be treated as a constant. Finally, M and, consequentially, Δt

should be large enough to assure the convergence towards the Gaussian, but not too large to prevent the approximation of constant x.

We approximate the sum in Eq. (A.22) by a Gaussian distribution with mean $M\mu$ and standard deviation $\sqrt{M}\sigma$. Therefore, we end up with

$$x_{t+\Delta t} = x_t - 2ax_t M\Delta t_0 + \sqrt{2(1 - x_t^2)Mb\Delta t_0} \cdot \eta_t, \tag{A.26}$$

where η_t is a normally distributed random variable. The inequality $Mb\Delta t_0 \ll 1$ should be satisfied in order to guarantee small variations of the variable x and, therefore, to preserve the approximation of a constant x during the interval Δt. Recalling the inequality (24), we come up with the following scaling relation between the number of micro steps M and the number of agents N

$$M = k\frac{N^2}{2}, \tag{A.27}$$

with an arbitrary small number k. Finally, the Langevin approximation for the stochastic process Eq. (A.22) is given by

$$\Delta x = -2a\Delta t x_t + \sqrt{2b\Delta t(1 - x_t^2)} \cdot \eta_t, \tag{A.28}$$

where $\Delta t = M\Delta t_0$. The limit of such an approximation is related to the assumption of the CLT and, therefore, to the goodness of the Gaussian approximation for the sum in Eq. (A.22). Near the edges, $x = -1$ or 1, the variable x cannot be considered to be constant, and the assumption of identically distributed variables does not hold any longer. Additionally, Eq. (A.28) does not incorporate neither the conditions of boundedness of x to the compact interval $[-1, 1]$, nor the natural boundary conditions of Eqs. (1) and (3). The boundary conditions, then, should be put in equation (A.28) 'by hand'. They are given by

$$\text{if } x(t) > 1 \quad \text{then} \quad \frac{x(t + \Delta t) + x(t)}{2} = 1, \tag{A.29}$$

$$\text{if } x(t) < -1 \quad \text{then} \quad \frac{x(t + \Delta t) + x(t)}{2} = -1. \tag{A.30}$$

A glance at Eqs. (A.29) and (A.30) shows that they are equivalent to a reflection around the edges of the domain of x, $x = 1$ and -1, respectively.[18]

References

Alfarano, S., 2006. An agent-based stochastic volatility model. Ph.D. Thesis, Department of Economics, University of Kiel.

---

[18]Alternative boundary conditions are also possible. For example, the boundary conditions can also be implemented as follows:

$$\text{if } x(t) > 1 \quad \text{then } x(t + \Delta t) = 1,$$
$$\text{if } x(t) < -1 \quad \text{then } x(t + \Delta t) = -1.$$

Alfarano, S., Lux, T., 2007. A noise trader model as a generator of apparent financial power laws and long memory. Macroeconomics Dynamics, in press.

Alfarano, S., Lux, T., Wagner, F., 2005. Estimation of agent-based models: the case of an asymmetric herding model. Computational Economics 26, 19–49.

Aoki, M., 1996. New Approaches to Macroeconomic Modeling: Evolutionary Stochastic Dynamics, Multiple Equilibria, and Externalities as Field Effects. University Press, Cambridge.

Aoki, M., 2002. Modeling Aggregate Behavior and Fluctuations in Economics. University Press, Cambridge.

Arifovic, J., 1996. The behaviour of the exchange rate in the genetic algorithm and experimental economies. Journal of Political Economy 104, 510–541.

Arthur, W.B., Holland, J.H., LeBaron, B., Palmer, R., Tayler, P., 1997. Asset pricing under endogenous expectations in an artificial stock market. Economic Notes 26, 297–330.

Barucci, E., 2003. Financial Markets Theory: Equilibrium, Efficiency and Information. Springer, Berlin.

Beja, A., Goldman, M.B., 1980. On the dynamic behavior of prices in disequilibrium. Journal of Finance 35, 235–248.

Bornholdt, S., 2001. Expectation bubbles in a spin model of markets: intermittency from frustation across scales. International Journal of Modern Physics C 12, 667–674.

Bouchaud, J.-P., Potters, M., 2000. Theory of Financial Risks. From Statistical Physics to Risk Management. University Press, Cambridge.

Challet, D., Marsili, M., 2004. From minority game to the real markets. Physical Review E 68, 168–176.

Chen, S.H., Yeh, C.H., 2002. On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis. Journal of Economic Behavior and Organization 49, 217–239.

Chiarella, C., Iori, G., 2002. A simulation analysis of the microstructure of double auction markets. Quantitative Finance 2, 346–353.

Cont, R., Bouchaud, J.P., 2000. Herd behaviour and aggregate fluctuations in financial markets. Macroeconomic Dynamics 4, 170–196.

Day, R.H., Huang, W., 1990. Bulls, bears, and market sheep. Journal of Economic Behavior and Organization 14, 299–329.

DeGrauwe, P., Dewachter, H., Embrechts, M.J., 1993. Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets. Blackwell, Oxford.

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

Föllmer, H., Horst, U., Kirman, A., 2005. Equilibria in financial markets with heterogeneous agents: a probabilistic perspective. Journal of Mathematical Economics 41, 123–155.

Gardiner, C.W., 2003. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, third ed. Springer, Berlin.

Gaunersdorfer, A., Hommes, C., 2007. A nonlinear structural model for volatility clustering. In: Teyssière, G., Kirman, A. (Eds.), Long Memory in Economics. Springer, Berlin.

Giardina, I., Bouchaud, J.-P., 2003. Bubbles, crashes and intermittency in agent based market models. Europhysics Journal B 31, 421–437.

Grannan, E., Swindle, G., 1994. Contrarians and volatility clustering. Complex Systems 8, 75–89.

Horst, U., 2004. Financial price fluctuations in a stock market model with many interacting agents. Economic Theory 25 (4), 917–932.

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 Behaviour and Organisation 49, 269–285.

Kirman, A., 1991. Epidemics of opinion and speculative bubbles in financial markets. In: Taylor, M.P. (Ed.), Money and Financial Markets. Blackwell, Cambridge, pp. 354–368.

Kirman, A., 1993. Ants, rationality, and recruitment. Quarterly Journal of Economics 108, 137–156.

Kirman, A., Teyssière, G., 2002. Microeconomic models for long memory in the volatility of financial time series. Studies in Nonlinear Dynamics and Econometrics 5, 137–156.

LeBaron, B., Arthur, W.B., Palmer, R., 1999. The time series properties of an artificial stock market. Journal of Economic Dynamics and Control 23, 1487–1516.

Levy, M., Levy, H., Solomon, S., 1994. A microscopic model of the stock market: cycles, booms, and crashes. Economics Letters 45, 103–111.

Lo, A., MacKinlay, C., 1999. A Non-Random Walk Down Wall Street. University Press, Princeton.

Lux, T., 1995. Herd behaviour, bubbles and crashes. Economic Journal 105, 881–896.

Lux, T., 1997. Time variation of second moments from a noise trader/infection model. Journal of Economic Dynamics and Control 22, 1–38.

Lux, T., 1998. The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions. Journal of Economic Behavior and Organization 33, 143–165.

Lux, T., Marchesi, M., 1999. Scaling and criticality in a stochastic multi-agent model of a financial market. Nature 397, 498–500.

Lux, T., Marchesi, M., 2000. Volatility clustering in financial markets: a microsimulation of interacting agents. International Journal of Theoretical and Applied Finance 3, 67–702.

Lux, T., Schornstein, S., 2005. Genetic learning as an explanation of stylized facts of foreign exchange markets. Journal of Mathematical Economics 41, 169–196.

O'Hara, M., 1995. Market Microstructure Theory. Blackwell, Cambridge.

Raberto, M., Cincotti, S., Focardi, S., Marchesi, M., 2003. Traders' long-run wealth in an artificial financial market. Computational Economics 22, 255–272.

Ramsey, J.B., 1996. On the existence of macro variables and of macro relationships. Journal of Economic Behavior and Organisation 30, 275–299.

Stauffer, D., Sornette, D., 1999. Self-organized percolation model for stock market fluctuation. Physica A 271, 496–506.

Van Kampen, N.G., 1992. Stochastic Processes in Physics and Chemistry, revised ed. North-Holland, Amsterdam.