Estimation of Agent-Based Models: the case of an Asymmetric Herding Model*

Simone Alfarano[a], Friedrich Wagner[b] and Thomas Lux[a]

July 12, 2004

[a] Dept. of Economics, University of Kiel, Germany
[b] Dept. of Physics, University of Kiel, Germany

Abstract

The behavioral origins of the stylized facts of financial returns have been addressed in a growing body of agent-based models of financial markets. While the traditional efficient market viewpoint explains all statistical properties of returns by similar features of the news arrival process, the more recent behavioral finance models explain them as imprints of universal patterns of interaction in these markets. In this paper we contribute to this literature by introducing a very simple agent-based model in which the ubiquitous stylized facts (fat tails, volatility clustering) are emergent properties of the interaction among traders. The simplicity of the model allows us to estimate the underlying parameters, since it is possible to derive a closed form solution for the distribution of returns. We show that the tail shape characterizing the fatness of the unconditional distribution of returns can be directly derived from some structural variables that govern the traders' interactions, namely the herding propensity and the autonomous switching tendency.

Keywords: Herd Behavior; Speculative Dynamics; Fat Tails; Volatility Clustering.
JEL Classification: G12; C61.

Corresponding author:
Dept. of Economics, University of Kiel; Olshausenstr. 40, D-24118 Kiel, Germany.
E-mail: alfarano@bwl.uni-kiel.de

*Earlier versions of this paper have been presented at the 11^{th} Symposium of the Society of Nonlinear Dynamics and Econometrics, Florence March 2003, the 8^{th} Spring Meeting of Young Economists, Leuven April 2003, the 8^{th} Workshop on Economics with Heterogeneous Interacting Agents, Kiel May 2003, the 27^{th} congress of Associazione per la Matematica Applicata alle Scienze Economiche e Sociali, Cagliari September 2003; research seminars at the department of Econometrics, University of Geneva, March 2003, and at the department of Physics, University of Cagliari, May 2003, and have gained considerably from comments by many participants in these events. We are particulary grateful to Manfred Gilli and Tony He for helpful suggestions and intense discussions during the preparation of this paper.

1 Introduction

Research in empirical finance has come up with a number of *stylized facts* of empirical data, involving conditional and unconditional properties of the time series. Starting with the former, the non-Gaussian character of the unconditional distribution of returns is a well-known ubiquitous finding in the literature [26, 23, 31]. In particular, it has been observed that practically all financial time series are *leptokurtic* [26], i.e. they are characterized by a higher concentration of probability mass in the center and in the tail of the distribution than the Gaussian. However, kurtosis is not a fully adequate measure of deviations from normality, since existence of the forth moment itself depends on the tail behavior of the distribution. It has been recently recognized that the extreme part of the distribution is well approximated by a Pareto law [31, 27] (in contrast to the exponential decay of the Gaussian). This indicates that a more appropriate way to measure the dispersion of returns is the *tail index*, defined as the order of the highest finite absolute moment. An analysis of estimated indices across different markets and time resolutions [22, 31] shows a relatively small interval of variability of these estimates, which typically hover between 2.5 and 4. Gopikrishnan *et al.* [13] go even further claiming that with the typical estimates around 3 an *inverse cubic law* holds for financial fluctuations.

Another universal feature that is strongly related to the leptokurtic shape of the distribution of returns is the intermittent behavior of the volatility, or *volatility clustering*, phenomenologically described as *"periods of quiescence and turbulence tending to cluster together"* [31, 35].

While serial correlation is absent in raw returns (which is in harmony with informational efficiency of financial markets), volatility clustering leads to positive autocorrelation for absolute and squared returns over an extended time horizon [31]. The slow decay of the auto-correlation of all measures of volatility[1] is again well represented by a power law, which is the defining property of long memory stochastic processes. It has been shown that the *hyperbolic decline* of the autocorrelation of the volatility is another extremely robust empirical phenomenon [21].

An additional important item to the list of empirical regularities is the so called *unit root property* [31, 35]: one is not able to reject the hypothesis that the prices follow a random walk or a martingale process. The resulting lack of predictability of the future price is in accordance with informational efficiency of the market.

Despite the huge amount of empirical work on the statistical properties of financial data, the origin of the universality of fat tails and volatility clustering in financial markets is still obscure. One could distinguish two competing hypotheses for their origin: one derived from the traditional Efficient Market Hypothesis (EMH) and a recent alternative which we might call the Interacting Agent Hypothesis (IAH). The EMH states that the price fully and instantaneously reflects any new information: the market is, therefore, efficient in aggregating available information with his *invisible hand*. The agents are assumed to be rational and homogeneous with respect to the access and their assessment of information; and, as a consequence, interactions among them can be neglected. The returns are, therefore, a mere reflection of forthcoming information so that, consequently, their empirical regularities would simply mirror those of the

[1]Volatility can be measured, for instance, by the absolute or the squared value of the returns, local variance or standard deviation.

"news arrival process", confining their origin to forces outside the realm of economics.

However, in recent literature, several papers try to explain the stylized facts as the macroscopic outcome of an ensemble of heterogeneous interacting agents [24, 6, 7]. In this view, the market is populated by agents with different characteristics, such as, differences in access to and interpretation of available information, different expectations, or different trading strategies. The traders interact, for example, by exchanging information or they trade imitating the behavior of others. The market possesses, then, an endogenous dynamics, and the strict one-to-one relationship with the news arrival process does not hold any longer (although the market might still be efficient in the sense of a lack of predictability). The universality of the statistical regularities is seen as an emergent property of this internal dynamics, governed by the interactions among agents. One of the main drawbacks of this approach arises from the complexity of the models which typically prevents an analytical solution and leaves only the possibility of a rough calibration of the key parameters (e.g. LeBaron [20]). No wonder, empirical implementations of agent-based models are practically non-existent. While earlier models of chartist-fundamentalist interaction with simple trading rules have been tested (with mixed success) via regime-switching time series models (cf. [30, 34]), for true multi-agent models parameter estimation and evaluation on the base of empirical data is still largely missing in the pertinent literature. As far as we know the only exception is a recent paper by Gilli and Winker [12], who estimate some of the parameters of Kirman's seminal herding model [17, 18]. They introduce a simulated moment approach extracting two key parameters of the model via matching of the empirical kurtosis and the first autocorrelation coefficient of squared returns. The main finding of their application to daily DEM/$ data is that estimated parameters are such that majorities would emerge in the herding process, instead of a balanced distribution of agents on the two groups of chartists and fundamentalists.

The goal of this paper is to add to the direction of empirical research on agent-based models initiated by Gilli and Winker: we set the stage by formulating an extremely simple agent-based model which, however, is still able to reproduce the key stylized facts listed above, in line with previous papers [32, 2]. The main contribution of the present work, then, consists in a *direct* estimation of the underlying parameters of the model using a parametric approach. Our traders are also divided into the groups of fundamentalists and noise traders following the legacy of Beja and Goldman [4], as well as many other authors. Like in [12] the interaction among the traders is based on a variant of the herding mechanism introduced by Kirman, but this herding mechanism is embedded into a simpler framework for the market dynamics which allows to derive a closed-form solution for the distribution of returns. The simplicity of this approach also enables us to consider a generalization of the herding mechanism of our model by allowing for asymmetric transition probabilities.

As it turns out, the resulting structure of our model is analogous to that of stochastic volatility models, which were introduced as a generalization of the Brownian motion in Black and Scholes's option pricing theory [5]. However, although these models can reproduce several empirical regularities, they are nevertheless driven more by mathematical convenience rather than economic intuition. We attempt, on the contrary, to introduce a plausible *behavioral* framework in our model, allowing for an economic interpretation of the resulting dynamics. It turns out that returns are governed by the product of a stochastic volatility term, related to the change in the strategies among traders, and a multiplicative iid noise term linked to the noise traders' misperception of the fundamental price.

The paper is organized as follows: in section 2 we describe the asymmetric herding mechanism that is employed as the main ingredient in our agent-based financial market. The complete framework of our speculative market is introduced in section 3. The estimation procedure for the parameters is detailed in section 4, and some final remarks conclude the paper.

2 The herding mechanism

Kirman [18] has introduced a simple stochastic model of information transmission initially designed to explain the herding behavior in ant colonies, gathering food from two identical sources located in their neighborhood. His formalization is symmetric with respect to switching probabilities generating a symmetric probability distribution of ants' visiting times at both mangers. In a long chain of subsequent papers this and similar approaches have been applied as models of herding and contagion phenomena in financial markets ([17],[19],[25]). Costantini and Garibaldi [8] have derived the equilibrium distribution of the related discrete-time stochastic process within the more general theoretical framework of Polya urn processes. Alfarano *et al.* [2] have derived and solved the associated Fokker-Planck equation for the pertinent continuous symmetric dynamic process. In the following, we generalize the previous approach by considering asymmetric transition probabilities that could describe, for instance, scenarios with non-identical sources of food or biased herding tendencies in a speculative market setting.

We describe the dynamics of the market as a jump Markov process in continuous time. The market is populated by a fixed number of agents N, each of them being either in state 1 or in state 2. The number of agents in the first state is denoted by n, and that in the second state by $N-n$. The conditional probabilities $\bar{\rho}(n', t+\Delta t|n,t)$ of changing from n at time t to n' at time $t+\Delta t$ with $n'-n=\pm 1, 0$ by a single switch are related to the transition probabilities $\pi(n \to n')$ per unit time by $\bar{\rho}(n', t+\Delta t|n,t) = \Delta t \; \pi(n \to n')$. The latter are given by

$$\pi(n \to n+1) = (N-n)(a_1 + bn) \text{ and } \pi(n \to n-1) = n[a_2 + b(N-n)]. \tag{1}$$

The constants a_1 and a_2 describe the idiosyncratic propensity to change the state, while the term b encapsulates the herding tendency. In contrast to earlier contributions, we allow the two constant parameters, a_1 and a_2, to assume different values, generating, then, the required asymmetric behavior. The transition probability $\pi(n \to n)$ to remain in the same state follows from the condition $\sum_{n'} \bar{\rho}(n', t+\Delta t|n,t) = 1$. Note that since transition probabilities should remain restricted to values < 1 per time increment, discrete simulations of the model would only be possible up to an upper limit of Δt which is given, for large N, by

$$\Delta t = \frac{2}{bN^2} \; . \tag{2}$$

The transition probabilities imply the so-called Master equation for the probability[2] $\bar{\omega}_n(t)$ to

[2]We denote probabilities referring to n by $\bar{\omega}$ to distinguish them from the probability densities $\omega(z)$ for the continuous variable $z = n/N$. Both are related by $\omega(n/N) = N\bar{\omega}_n$

find n agents in the state 1 at time t

$$\frac{\Delta\bar{\omega}_n(t)}{\Delta t} = \sum_{n'} \left(\bar{\omega}_{n'}\, \pi(n' \to n) \;-\; \bar{\omega}_n\, \pi(n \to n') \right). \tag{3}$$

For large enough N we can represent the group dynamics by a continuous variable:

$$z = \frac{n}{N}. \tag{4}$$

As shown in appendix A1, the Master equation (3) is equivalent to a Fokker-Planck equation if terms of order $1/N^2$ are neglected

$$\frac{\partial \omega(z,t)}{\partial t} = -\frac{\partial}{\partial z}\left[A(z)\omega(z,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[D(z)\omega(z,t)\right] \tag{5}$$

where $A(z)$ represents the drift term

$$A(z) = a_1 - (a_1 + a_2)z \tag{6}$$

that can be conveniently rewrite as:

$$A(z) = (a_1 + a_2)(\bar{z} - z) \tag{7}$$

where $\bar{z}$ is the mean value of the variable z (see appendix A2). The previous formula underlies the mean-reverting nature of the process (1), with a speed given by the constant $a_1 + a_2$. The diffusion term $D(z)$ is given by:

$$D(z) = 2b(1-z)z \;+\; \frac{1}{N}(a_1(1-z) + a_2 z). \tag{8}$$

It is remarkable that $A(z)$ is independent of the herding mechanism and that the $1/N$ corrections appear only in the diffusion term.

In appendix A2 we compute the equilibrium distribution $\omega_0(z,N)$. Interestingly, it turns out that it depends only on the ratios

$$\varepsilon_1 = \frac{a_1}{b} \quad \text{and} \quad \varepsilon_2 = \frac{a_2}{b}$$

but not on the size of the constants a_i and b:

$$\omega_0(z,N) = K(\varepsilon_1,\varepsilon_2,N)\left(z + \frac{\varepsilon_1}{2N}\right)^{\varepsilon_1 - 1}\left(1 - z + \frac{\varepsilon_2}{2N}\right)^{\varepsilon_2 - 1}. \tag{9}$$

If we neglect the $1/N$ terms in $D(z)$ we obtain

$$\omega_0(z) = \frac{1}{B(\varepsilon_1,\varepsilon_2)}\, z^{\varepsilon_1 - 1}\,(1-z)^{\varepsilon_2 - 1} \tag{10}$$

where

$$B(\varepsilon_1,\varepsilon_2) = \frac{\Gamma(\varepsilon_1)\Gamma(\varepsilon_2)}{\Gamma(\varepsilon_1 + \varepsilon_2)}. \tag{11}$$

We see from (9) and (10) that for $\varepsilon_1 < 1$ or $\varepsilon_2 < 1$ the probability density $\omega_0(z, N)$ does not converge uniformly to $\omega_0(z)$. For $z \sim 1/N$ or $1 - z \sim 1/N$ the $1/N$ term in D has to be kept in order to obtain a finite (however large) value of the probability density.[3] In the following, we ignore the $1/N$ term in D and the correspondent N-dependence in eq. (9). To be able to do so, we restrict our analysis to eq. (10), assuming $\varepsilon_{1,2} > 1$, which avoids the problem of having to estimate the value of N, or imposing an *ad hoc* choice of its order of magnitude. Using $\varepsilon_1 = \varepsilon_2$ in eq. (10), we recover the same functional form of the equilibrium distribution as in Kirman's article [18]. The distribution (10) exhibits a unique mode if both parameters take a value larger than 1 (see panel a in Figure 1), while it shows bi-modality for the case $\varepsilon_{1,2} < 1$ (see panel b). Furthermore, the distribution shows a monotonic behavior if one parameter is larger than 1 and the other smaller than 1 (see panels c and d for an example of a decreasing and increasing monotonic distribution, respectively). The four equilibrium distributions with different choices of ε_1 and ε_2, shown in Figure 1, illustrate this great flexibility of the asymmetric model.[4]

[Insert Figure 1 approximately here]

An alternative way to describe our process (1) is via the Langevin equation:

$$z_{t+\Delta t} = z_t + (a_1 + a_2)(\bar{z} - z)\Delta t + \sqrt{2b\ \Delta t\ (1 - z_t)z_t}\ \lambda_t \tag{12}$$

where λ_t is a iid normally distributed random variable. In appendix A4 we show, that the transition probability

$$\omega(z', t + \Delta t | z, t) = N\bar{\omega}(n', t + \Delta t | n, t) \tag{13}$$

can be written as a Gaussian with mean $z + \Delta t\ A(z)$ and variance $\Delta t\ D(z)$ provided terms of the order $(a_i \Delta t)^2$ and $(b\Delta t)^2$ can be neglected. Eq. (12) is another way to formulate this statement. Usually (see Van Kampen [29] for the relationship between the Fokker-Planck and the Langevin equation) an argument of this type is used to prove that (5) is a consequence of (12) in the limit of $\Delta t \to 0$. In our case both (12) and (5) can be obtained from the Master equation for large N and we do not have to worry about the validity of the Langevin equation at finite Δt (called Euler-Maruyama discretization in mathematical literature) as an approximation to the Fokker-Planck equation in continuous time. Note that the restriction on the size of the time step Δt in the Langevin equation is much less severe then condition (2) when using the transition probabilities (1).

It is interesting to note that De Jong *et al.* [9] and Stegenborg, Larsen and Sørensen [28] have employed the continuous time version of eq. (12) as a stochastic model of floating exchange rates in a target zone. The model is estimated for several European FX rates using GMM and a new estimation technique based on the eigenfunctions of the underlying Fokker-Planck equation [16]. In contrast, our analysis derives eq. (12) as a consequence of an interacting-agent approach, rather than as a starting point of a stochastic model for the description of the data.

[3] It is worth pointing out that, despite this N-dependence of ω_0, the qualitative dynamics of our model is independent of N (cf. Alfarano, Wagner and Lux [2]). The importance of system size N in eq. (9) for $\varepsilon_{1,2} < 1$ stems solely from the relevance of the grid of admissible values of z in scenarios with a bi-modal distribution and concentration of probability mass at the upper and lower boundaries of its support rather than from a "true" qualitative dependence on system size.

[4] Note that the symmetric model, characterized by a single parameter ε, exhibits a uni-modal distribution for $\varepsilon > 1$, a bi-modal distribution $\varepsilon < 1$ and a uniform distribution for $\varepsilon = 1$. However, the monotonic case can not be observed under the symmetric setting.

3 The artificial financial market

3.1 Market implementation

We now implement this two-state dynamics as the main ingredient in a financial market model with interacting heterogenous agents. Our market participants are divided into two groups:

N_F *fundamentalists*, who buy (sell) if the actual market price p is below (above) the fundamental value p_F.

N_C *noise traders*, who are subject to "irrational" fads and moods, following the seminal contribution by De Long *et al.* [10].

The arbitrary states 1 and 2 of the previous section are now identified as noise trader and fundamentalist strategies, respectively. This implies that a_1 and a_2 are the propensities for an agent to switch autonomously to the other strategy, which might include a bias towards one of both alternatives. For example, $a_1 > a_2$ would mean that on average the propensity of autonomous conversion of a former fundamentalist to noise trader behavior is higher than the probability for a switch in the opposite direction. The propensity to change behavior under the influence of herding expressed by b must be the same in both directions, otherwise one would obtain a trivial distribution for z.

In order to arrive at a formalization of the price dynamics, we have to formulate behavioral rules for demand and supply of our two groups. Fundamentalists' excess demand is given by:

$$ED_F = N_F \ln \frac{p_F}{p}. \tag{14}$$

Somewhat deviating from earlier contributions, we assume a dependence on the *log* difference between the fundamental value and the current market price instead of absolute differences. This appears intuitively plausible as fundamentalists' trading should be based on relative rather than absolute under- and overvaluation. Moreover, it facilitates derivation of a simple form of the process governing returns. Since the observed average daily changes are small ($\sim$1%) we expect that results would not be drastically different when using absolute differences instead of relative ones.

Excess demand of the noise traders is formalized as:

$$ED_C = -r_0 N_C \xi \tag{15}$$

where ξ represents the average noise trader's 'mood' (positive or negative). The constant r_0 is a scale factor for their impact on the price formation, and the expression is multiplied by -1 for notational convenience in the following derivations. Within a Walrasian scenario, the equilibrium price is derived by setting $ED_F + ED_C$ to zero. This leads to to the following equation for the equilibrium price:

$$\ln \frac{p}{p_F} = r_0 \frac{N_C}{N_F}\,\xi = r_0 \frac{z}{1-z}\,\xi \tag{16}$$

where z and $1-z$ are the fractions of noise traders and fundamentalists within the population, respectively. Note that our choice of a random development of the noise traders' mood ξ_t (and derived from it, stochastic excess demand of this group) is inspired by the modelling of noise traders in the seminal paper by de Long *et al.* [10]. Our noise traders, therefore, are characterized by random misperceptions rather than some heuristic trading rules such as trend following behavior used, for instance, in Kirman's adaptation of the ant model to foreign exchange trading [17, 19].

We define the continuously compounded return over a time interval Δt by

$$r(t) = \ln \frac{p_{t+\Delta t}}{p_t}. \tag{17}$$

Using (16) we can write it as

$$r(t) = r_0 \left(\frac{z(t+\Delta t)}{1 - z(t+\Delta t)} \, \xi(t+\Delta t) \; - \; \frac{z(t)}{1 - z(t)} \, \xi(t) \right). \tag{18}$$

Empirically r is correlated only for time scales much smaller than a trading day. This property can be build in by assuming that ξ changes much faster than z. Then the return (18) can be approximated by considering only changes in ξ while keeping z constant. Defining $\eta(t) \equiv \xi(t + \Delta t) - \xi(t)$ this approximation yields:

$$r(t) = r_0 \, \frac{z(t)}{1 - z(t)} \, \eta(t). \tag{19}$$

The level of accuracy of this approximation depends on the choice of the underlying parameters, namely b and $\varepsilon_{1,2}$. For given values of $\varepsilon_{1,2}$, the accuracy between eqs. (18) and (19) is a decreasing function of b, which, in fact, governs the diffusion part of the size of the increments Δz (see eq. 12).

If we assume the random variable η_t to be iid on the time scale Δt which we might think of as daily data, daily returns in our model will be uncorrelated. Therefore, we model the development of the average noise traders' mood ξ as a random walk with increments η. This choice of the type of stochasticity governing the noise traders' behavior is different from the related literature, in particular the paper of De Long *et al.* [10], where the misperseption of the noise traders itself is modeled as an iid random variable. The random walk assumption, however, avoids the abrupt variations of the market price, which would occur if we would incorporate the De Long formalization.

The behavior of r depends both on $\eta(t)$ and the dynamics of z which is governed by the herding mechanism. In the following we will discuss various choices for the distribution of $\eta(t)$. If sufficiently high moments of η exist, the power law property of the the distribution of $z/(1-z)$ derived in the next paragraph will also carry over to the distribution of returns.

[Insert Figure 2 approximately here]

In Figure (2) we show a typical simulated pattern of the price obtained from (16) using the Langevin equation (12) for the change of z and a uniform distribution for $\Delta\xi = \eta$. In the lower parts we compare the returns calculated from (18) and from the approximation (19). They behave qualitatively similar which can be taken as a justification of the approximation

(19)[5]. One can observe large deviations from the constant fundamental price alternating with small oscillations around it. This feature is related to the volatility clustering phenomenon observed in the time series of returns. Large fluctuations of r happen when a large fraction of traders adopts the noise trader behavior. On the contrary, in the presence of a large fraction of fundamentalists only small variations occur. This intermittent dynamics apparently gives rise to time-series properties capturing some of the stylized facts of empirical data The underlying behavioral mechanisms are similar to those of the more involved models of Kirman [19], Lux and Marchesi [25] and Wagner [32] from which our approach draws inspiration.

3.2 Analytical solutions for the unconditional distribution

The simple structure of eq. (19) permits to derive analytically the unconditional distribution of returns[6]. As a first step we write the return as

$$r = \sigma(t) \cdot \eta(t) \tag{20}$$

with

$$\sigma(t) = r_0 \, \frac{z(t)}{1 - z(t)} \, . \tag{21}$$

The equilibrium probability distribution $p(\sigma)$ can be computed from the general formula

$$p(\sigma) = \frac{dz}{d\sigma} \cdot \omega_0(z). \tag{22}$$

Using the relationship $z = \sigma/(r_0 + \sigma)$ derived from (21) and

$$\frac{dz}{d\sigma} = \frac{r_0}{(r_0 + \sigma)^2}$$

we can express ω_0 in equation (10) as function of σ and obtain the equilibrium distribution of σ

$$p(\sigma) = \frac{1}{r_0} \frac{1}{B(\varepsilon_1, \varepsilon_2)} \left(\frac{\sigma}{r_0}\right)^{\varepsilon_1 - 1} \left(\frac{r_0}{\sigma + r_0}\right)^{\varepsilon_1 + \varepsilon_2} \tag{23}$$

which exhibits a power law decay for $\sigma \gg r_0$

$$p(\sigma) \sim \left(\frac{1}{\sigma}\right)^{\varepsilon_2 + 1} . \tag{24}$$

As is obvious from eq. (24), the exponent of the tail is related to the parameters characterizing the behavior of the fundamentalists, namely the ratio between the tendency a_2 of autonomous switches from fundamentalist to noise trader behavior, and the herding parameter b. The scaling behavior of the tail also carries over to the distribution of r as shown in appendix A3. It is independent of the distribution of η provided the moment $E[|\eta|^{\varepsilon_2}]$ exists[7]. This remarkable property

[5] Given a small value of b, the strict similarity of the two time series of returns in Figure (2) is very robust with respect to the choice of $\varepsilon_{1,2}$.

[6] From the Fokker-Planck equation (5), we can also derive further dynamical properties of returns, such as the autocorrelation of the absolute values. The full analysis of the dynamics is left for future research.

[7] In general, the tail behavior of a product of two independent random variables is governed by the component with the smaller tail index (fatter tail).

of the equilibrium distribution of r is compatible with the empirical finding of power-law tails described in the introduction.

Equation (20) also involves the noise term η. In order to obtain solutions in closed form for the distribution of returns, we assume that η follows either:

- a bimodal distribution with two values ± 1, occurring with equal probability (*spin-noise model*), or
- a uniform distribution over the interval $[-1, +1]$ (*uniform-noise model*)

Since η is not directly observable, we could have chosen any other functional form, for instance a Normal distribution or a Student t-distribution. Since the distribution of the noise is symmetric around zero, we can confine our analysis to the distribution of absolute returns as a measure of the volatility v:

$$v = |r| = \sigma |\eta|. \tag{25}$$

In the case of the spin-noise-model[8], the distribution of the volatility v is given by eq. (23), since the absolute value of returns coincides with σ. The derivation of the distribution of absolute returns in the case of uniformly distributed noise is more involved (see appendix A3), but can also be obtained in closed form:

$$p_u(v) = \frac{1}{r_0} \frac{\varepsilon_2}{\varepsilon_1 - 1} \left[1 - \beta \left(\frac{v}{v + r_0} ; \varepsilon_1 - 1, \varepsilon_2 + 1 \right) \right] \tag{26}$$

where β is the incomplete beta function[9].

[Insert Figure 3 approximately here]

In Figure 3 we compare the probability density functions for the spin-noise-model and the uniform-noise-model. To eliminate the influence of the scale parameter r_0 we show the density function for the normalized volatility $v_n = v/E[v]$. The value of ε_1 has been chosen such that the second moment is the same in both cases. The behavior for $v > E[v]$ is dominated by the asymptotic eq. (24) and therefore very similar, as we expect from the discussion of eq.(23). Both distributions mainly differ for small v, where the uniform noise model avoids convergence to zero resulting from (23). Any continuous distribution of the noise generates a flat region or 'plateau' for small v in agreement with the data. This property will turn out to be the reason for a better descriptive power of the uniform-noise model as compared to the spin noise model.

4 Estimation of the parameters of the models

To apply the model to empirical data we have to estimate the scale parameter r_0, the ratios $\varepsilon_{1,2}$ and the scale of the two constants $a_1 + a_2$. The main difficulty is the occurrence of two sources of randomness in the dynamics of returns in equation (25), the stochastic change of z and the multiplicative noise η, whereas only one measurement (absolute value of the return) is available at each time step. Therefore we divide the estimation into two steps. For a stationary Markov process with exponentially decaying correlations we can use the observed sample

[8] In the following, subscripts s and u indicate the spin-noise and uniform-noise models, respectively.

[9] The incomplete beta function is defined as $\beta(x; a, b) = (1/B(a, b)) \int_0^x t^{a-1}(1-t)^{b-1} dt$.

of returns to estimate the equilibrium parameters r_0 and $\varepsilon_{1,2}$ by maximum likelihood fit to the analytically known unconditional distributions. We should emphasize, however, that we use only an approximation of the "true" likelihood; we pretend, in fact, that the realizations of the Markovian process (19) are independent and identically distributed, according to the unconditional distributions given by equations (23) and (26), in the case of the spin noise or uniform noise respectively. The advantage of using this approximation is the simplicity of its implementation and the reduced computational burden. On the other hand, only parameters from the unconditional distribution can be estimated. The method gives asymptotically consistent estimates as the sample size $T \to \infty$, if the sampling frequency $\Delta = \Delta_T \to 0$ and the increase of T is faster than the decrease of Δ_T, i.e. $T\Delta_T \to \infty$. The estimates are asymptotically normal under the additional condition $T\Delta_T^2 \to 0$. (cf. Genon-Catalot *et al.* [11]). For an application of this approximation to stochastic volatility models see [11]. Once intensities $\varepsilon_{1,2}$ have been estimated, the scale of a_1 and a_2 can be adjusted heuristically from the correlation of $v(t)$. Since the Langevin equation (12) has a drift term which is linear in z, the correlations of any function of z have asymptotically an exponential decay with a characteristic time $\tau = (a_1 + a_2)^{-1}$. Since $\eta(t)$ in equation (25) is uncorrelated, this exponential decay carries over to v, with the same rate. The second step of the estimation, therefore, consists in adjusting τ to the observed decay of the correlation of v at large time lags. This method is not the most efficient, since the correlations are not solely governed by the decay constant τ, but also depend on the herding parameter b via the intensities $\varepsilon_{1,2}$.

Out of the three parameters that enter in the unconditional distribution, namely ε_1, ε_2 and r_0, the later can be expressed in terms of the other two parameters by imposing the following normalization on the empirical data[10]:

$$E[v] = 1. \tag{27}$$

Due to the relation $E[z/(1-z)] = \varepsilon_1/(\varepsilon_2 - 1)$ equation (27) implies the following values of r_0 which define the scale of the fluctuations:

$$r_0 = \frac{\varepsilon_2 - 1}{\varepsilon_1} \cdot \begin{cases} 1 & \text{spin noise} \\ 2 & \text{uniform noise} \end{cases} \tag{28}$$

For our implementation of the the estimation we use the following data sets: The daily returns of the Gold price (5034 entries) over the period 1974-1998 (abbreviated by Gold), the daily quoted stock prices of two large German companies, Deutsche Bank (DB, 6771) and Siemens (6674), in the period 1974-2001 and the daily variations of the index of the German stock exchange (DAX, 9761) from 1959 to 1998.

4.1 Estimation of the spin-noise model

The resulting values of the parameters ε_1 and ε_2 in the case of the spin-noise-model are given in Table 1.

[10]The normalization is based on the implicit assumption of existence of the population mean. Therefore eq. (28) holds under the condition $\varepsilon_2 > 1$, which guarantees the existence of the mean of the process (21) (note that ε_2 governs the asymptotic behavior of the tail and, consequently, the existence of the moments).

Data Set	$\hat{\varepsilon}_1$	$\hat{\varepsilon}_2$	$-\ln L_{\varepsilon_1,\varepsilon_2}$	$\hat{\varepsilon}$	$-\ln L_{\varepsilon}$	p-value
Gold	1.23 ± 0.04	4.2 ± 0.4	4926.6	2.04 ± 0.02	5081.9	0.00
DB	1.53 ± 0.05	5.1 ± 0.4	6602.9	2.44 ± 0.03	6746.8	0.00
Siemens	1.69 ± 0.06	4.1 ± 0.1	6457.4	2.49 ± 0.03	6540.0	0.00
DAX	1.37 ± 0.02	14.0 ± 2.0	9569.6	2.50 ± 0.02	10168.0	0.00

Table 1: *Estimated parameters for the spin-noise-model $\varepsilon_{1,2}$ (asymmetric model) with the corresponding Log-Likelihood values at the maximum. The parameter ε refers to the estimation with the constraint $\varepsilon_1 = \varepsilon_2 = \varepsilon$ (symmetric model) with the pertinent Log-Likelihood. The p-values of the Likelihood Ratio test are given in the last column.*

In Figure 4 we show the comparison of the theoretical distribution with the empirical one for Gold and the DAX. For all the considered time series, except for DAX, we obtain a value of $\hat{\varepsilon}_2$ close to the empirical power law (see Table 2). The large entry in the case of DAX is due to the extended plateau at $v = 0$. Since for $\varepsilon_1 > 1$ the theoretical pdf has a zero at $v = 0$, the choice of a large ε_2 is the best compromise the estimation can make. A value $\varepsilon_1 < 1$ would be even worse since it leads to a spike at $v = 0$.

[Insert Figure 4 approximately here]

4.2 Estimation of the uniform-noise model

We also apply the maximum likelihood method to the same data sets to estimate the uniform-noise model. The resulting parameters are listed in Table 2.

Data Set	$\hat{\varepsilon}_1$	$\hat{\varepsilon}_2$	$-\ln L_{\varepsilon_1,\varepsilon_2}$	$\hat{\varepsilon}$	$-\ln L_{\varepsilon}$	p-value	$\hat{\alpha}_H$
Gold	3.2 ± 0.4	3.9 ± 0.4	4942.7	3.6 ± 0.1	4943.2	0.34	2.9 $(2.4, 3.4)$
DB	6.0 ± 1.0	4.4 ± 0.4	6683.0	4.9 ± 0.1	6683.7	0.25	3.4 $(2.9, 4.0)$
Siemens	7.0 ± 2.0	3.9 ± 0.4	6573.9	4.7 ± 0.1	6576.8	0.02	3.7 $(3.2, 4.3)$
DAX	16 ± 5	4.9 ± 0.3	9534.3	6.7 ± 0.2	9542.6	0.00	3.1 $(2.9, 3.6)$

Table 2: *Estimated parameters using the uniform-noise model. Note somewhat different results for the LR test. The last column shows the Hill estimates [14] of the index of the tail for 2.5% tail size, with their 95% asymptotic confidence interval.*

Figure 5 shows the fit for the Gold and the DAX. For the latter, there is a substantial improvement using the uniform-noise model over the spin-noise model, since the plateau is now correctly accounted for. On the contrary, it seems that the goodness of fit is slightly worse for the other cases, as we see from the lower values of the likelihood when compared to the pertinent values of the spin-noise model. In appendix A5 we show that this somewhat disturbing conclusion, namely the superior descriptive power of the spin-noise model over the uniform noise model, might be just a fictitious effect of the discreteness of prices, which creates an artificial concentration of small returns at certain threshold levels. Taking into account this particular feature of the data, we, in fact, observe that the uniform-noise model performs better than the spin-noise framework.

The large values of $\hat{\varepsilon}_1$, especially for DAX, are not unreasonable. They imply that fundamentalists have a relatively high propensity to autonomously switch to a noise trader behavior, while the pertinent herding effect is comparatively smaller.

[Insert Figure 5 approximately here]

Recalling that the parameter ε_2 should coincide with the so-called tail index of the unconditional distribution, we compare the pertinent results to those obtained with the standard conditional ML estimator of the tail index. As can be seen from Table 2, in all cases, the estimated values of $\hat{\varepsilon}_2$ are somewhat below the 95% interval for the semi-parametric Hill estimator for the tail index.[11] The slight tendency of the parametric estimation towards somewhat lower values, when compared with the semi-parametric estimator, may be explained by the influence of the center of the distribution on the estimated values of ε_2. Nevertheless, we observe a relatively small interval of variability of $\hat{\varepsilon}_2$ in harmony with the remarkable homogeneity of the tail index for empirical data.

5 Discussion of the results

Our asymmetric herding model does account for the most pervasive stylized facts. The property of fat tails of the distribution of returns is an essential consequence of our framework. This is also true for volatility clustering, as shown in Figure 6. The clusters arise from equation (19) with values z of the noise traders' concentration getting close to 1. Due to the vanishing diffusion term this situation may persist for some time. Figure 6 shows that the structure of the autocorrelations for simulations of the model is, in fact, similar to that of empirical data in many respects: (i) raw returns appear uncorrelated while (ii) squared and absolute returns have highly significant autocorrelations which apparently fall off only slowly, (iii) autocorrelations of absolute returns are larger throughout than those of squared returns. Absence of correlations in raw returns has been build in by the assumption that the noise variable in equation (19) is iid distributed. A short-coming of our model is the Markov property of herding models of this type which implies that autocorrelations decay exponentially for large time lags. This precludes long-memory *strictu sensu* (i.e., hyperbolic decline of the autocorrelations of some powers of returns). However, an exponential decay with a small decay constant is hard to distinguish from "true" long term-dependence (see Alfarano, Lux and Wagner for analytical details [2]).

[Insert Figure 6 approximately here]

Another important property of our model is that the parameters of the distribution can be related to the behavior of the agents. For example, $\varepsilon_1 > \varepsilon_2$ indicates that, *on average*, the market is dominated by a noise trader attitude (i.e. $E[z] > 0.5$). However, irrespective of whether

[11] One might remark, however, the well known fact that the asymptotic distribution underestimates the finite-sample variability of estimates for processes with volatility clustering (cf [15]).

$E[z]$ is above or below 0.5, the herding component leads to temporal emergence of a majority of noise traders once in a while triggering large fluctuations which are, then, responsible for the volatility clustering observed in Fig. 2. The asymmetry in the transition probabilities, therefore, adds a considerable amount of flexibility to the original framework of the ants model in which $E[z] = 0.5$ is imposed by assumption. A glance at Tables 1 and 2 shows that the spin-noise model always indicates dominance of fundamentalists practices, while the uniform-noise model only finds fundamentalists dominance in the case of Gold.

Note also that although ε_2 exclusively governs the asymptotic tail behavior, the *transition* to the asymptotic power law regime at large v is also affected by ε_1. For $\varepsilon_1 = 2$, one finds an exact Pareto law for $p_u(v)$ (see appendix A.3). With increasing ε_1, the extented tail region to which the asymptotic power law applies decreases exponentially with ε_1. This implies the prediction that markets with a strong herding component in the transition from fundamentalists to noise traders have a pronounced power law behavior.

It seems interesting to compare our results to the estimates of the parameters of the Kirman model by Gilli and Winker [12] who found, in our notation, $\varepsilon_1 = \varepsilon_2 \sim 0.3$. Despite the differences in both the design of the herding mechanism and the market model, we might be able to shed some light on the discrepancy in estimated parameters. In [12], a linear relation between return (or price) and the fraction z of the noise traders is used, whereas we employ the non-linear relation (19). However, to generate volatility clusters in a linear model, z has to stay near the extreme values 0 or 1 most of the time, with a small probability to switch between these modes (a detailed explanation of the underlying mechanism can be found in [1]). In the symmetric herding model analyzed by Gilli and Winker, this can only be achieved with values $\varepsilon = \varepsilon_1 \varepsilon_2 < 1$. In contrast, the nonlinearity and asymmetry of our model allows to obtain volatility clustering without a bimodal distribution.

Finally, comparison of our asymmetric setting with the symmetric counterpart used in previous literature [18, 12], rises the question whether the data fully exploit this enhanced flexibility. In order to shed some light on this issue, we have reported the outcome of likelihood ratio tests in Tables 1 and 2, where the unconstrained model is described by the transition probabilities (1) and the alternative is characterized by the constraint $\varepsilon_1 = \varepsilon_2 = \varepsilon$. As can be seen, the outcome of the test depends to some extent on the type of multiplicative noise considered. In the case of the spin-noise model the symmetric setting is strongly rejected by all data-sets. On the contrary, the uniform-noise version shows less homogeneous behavior with only two cases of rejection of $\varepsilon_1 = \varepsilon_2$ at 5% level of significance. Of course, the inability of rejection of $\varepsilon_1 = \varepsilon_2$ occurs for cases in which we only find a slightly asymmetric distribution (e.g. $E[z] = \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} = 0.46$ in the case of gold) which is hard to distinguish from the symmetric case with a uni-modal distribution (imposed by our assumption $\varepsilon_{1,2} > 1$). It would be interesting, therefore, to allow for the bi-modal case ($\varepsilon_{1,2} < 1$) as an alternative which has a somewhat different mechanism of generating volatility clustering (cf. Alfarano and Lux [1]). To include this scenario, we would, however, have to take into account the N dependency of the limiting distribution of z (cf. eq. (9)), which would translate into N dependency of v as well. Therefore, either an *ad hoc* choice of the number of agents or some additional method of estimating N would be required. We leave this topic for future research.

6 Conclusions

We have presented an extremely simple model of an agent-based artificial financial market, which nevertheless is able to generate the key stylized facts of financial time series. Due to its very simple structure, we were able to obtain a full analytical solution for the probability distribution of returns, which allows a better understanding of the underlying dynamics. Consequently, the parameters of the model can be estimated directly via maximum likelihood, giving more insights into the potential behavioral origins of the statistical regularities of the data. One of the main results of the paper is that we can directly connect some crucial micro variables of the market (such as the traders' tendency to switch between strategies) to a macroscopic observable (the tail index). The estimations are quite encouraging, especially in the quality of fit of the unconditional distribution. It has also been shown that asymmetry of the transitions between the two groups, fundamentalists and noise traders, is not always rejected by the data in favor of a symmetric framework. In conclusion, it appear worthwhile to include this further element of flexibility in agent-based models of interaction in financial markets.

References

[1] S. Alfarano and T. Lux. A minimal noise traders model with realistic time series properties. http://www.bwl.uni-kiel.de/ Ordnung+Wettbewerbspolitik/ewp/ewp.html, 2003. Working paper, University of Kiel.

[2] S. Alfarano, T. Lux, and F. Wagner. Apparent long memory in agent-based models: an explanation, 2003. Work in progress, University of Kiel.

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

[4] A. Beja and M. B. Goldman. On the dynamic behavior of prices in disequilibrium. *Journal of Finance*, 35:235–248, 1980.

[5] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–653, 1973.

[6] W. A. Brock and C. H. Hommes. A rational route to randomness. *Econometrica*, 65:1059–1095, 1997.

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

[8] D. Costantini and U. Garibaldi. A probabilistic foundation of elementary particle statistics. *Stud. Hist. Phil. Phys. 28*, 28:483–506, 1997.

[9] F. De Jong, F. C. Drost, and B. J. M. Werker. A jump-diffusion model for exchange rates in a target zone. *Statistica Neerlandica*, 55:270–300, 2001.

[10] B. J. De Long, A. Shleifer, L. H. Summers, and R. J. Waldmann. Noise trade risk in financial markets. *Journal of Political Economy*, 98:703738, 1990.

[11] V. Genon-Catalot, T. Jeantheau, and C. Laredo. Parameter estimation for discretely observed stochastic volatility models. *Bernoulli*, 5:855–872, 1999.

[12] M. Gilli and P. Winker. A global optimization heuristic for estimating agent based models. *Computational Statistics and Data Analysis*, 42:299–312, 2003.

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

[14] B. M. Hill. A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3:1163–1173, 1975.

[15] P. Kearns and A.R. Pagan. Estimating the density tail index for financial time series. *Review of Economics and Statistics*, 79:171–175, 1997.

[16] M. Kessler and M. Sø rensen. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5:299–314, 1999.

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

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

[19] A. Kirman and G. Teyssiére. Microeconomic models for long memory in the volatility of financial time series. *Studies in Nonlinear Dynamics & Econometrics*, 5:137–156, 2002.

[20] B. LeBaron. Agent based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, 24:679–702, 2000.

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

[22] T. Lux. The stable Paretian hypothesis and the frequency of large returns: An examination of major German stocks. *Applied Financial Economics*, 6:463–475, 1996.

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

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

[25] T. Lux and M. Marchesi. Volatility clustering in financial markets: A micro-simulation of interacting agents. *International Journal of Theoretical and Applied Finance*, 3:67–702, 2000.

[26] B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 35:394–419, 1963.

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

[28] K. Stegenborg Larsen and M. Sø rensen. Diffusion models for exchange rates in a target zone. http://www.math.ku.dk/ michael/target.pdf, 2003.

[29] N. G. van Kampen. *Stochastic processes in Physics and Chemistry.* North Holland, Amsterdam, 1992. Revised edition.

[30] R. Vigfusson. Switching between chartists and fundamentalists: A Markov regime-switching approach. *International Journal of Finance and Economics*, 2:291–305, 1997.

[31] C. G. de Vries. Stylized facts of nominal exchange rate returns. In F. van der Ploeg, editor, *The Handbook of International Macroeconomics*, pages 348–389. Blackwell, Oxford, 1994.

[32] F. Wagner. Volatility cluster and herding. *Physica A*, 322:607–619, 2003.

[33] T. Weiland. On the unique numerical solution of maxwellian eigenvalue problems in three dimensions. *Particle Accelerators*, 17:227242, 1985.

[34] F.H. Westerhoff and S. Reitz. Nonlinearities and cyclical behavior: The role of chartists and fundamentalists. *Studies in Nonlinear Dynamics and Econometrics*, 7, 2003.

[35] M. Youssefmir and A. Huberman. Clustered volatility in multiagent dynamics. *Journal of Economic Behavior and Organization*, 32:101–118, 1997.

[36] J. Zinn-Justin. *Quantum Field Theory and Critical Phenomena.* Clarendon Press, Oxford, 1989.

A Appendix

A.1 Derivation of the Fokker-Planck equation

To derive the Fokker-Plank equation (5) we start by writing the transition probabilities from eq. (1) for the switching process as:

$$\pi(n \to n+1) \quad = \pi_+(n) = \quad (N-n)(a_1 + bn) \tag{29}$$

$$\pi(n \to n-1) \quad = \pi_-(n) = \quad n[a_2 + b(N-n)]. \tag{30}$$

The transition probabilities define a Markovian stochastic process, that belongs to the class of so-called “birth-death processes” [3] or “one-step processes” [29]. The so-called Master equation for the flux of probability from eq. (3) reads[12]

$$\frac{\Delta \bar{\omega}_n(t)}{\Delta t} = \bar{\omega}_{n+1}\pi_-(n+1) + \bar{\omega}_{n-1}\pi_+(n-1) - \bar{\omega}_n\pi_-(n) - \bar{\omega}_n\pi_+(n). \tag{31}$$

Defining the current $\bar{j}_n$ as the probability flow from $n-1$ to n

$$\bar{j}_n = \bar{\omega}_{n-1}\pi_+(n-1) - \bar{\omega}_n\pi_-(n) \tag{32}$$

the Master equation can be written as a *discrete continuity equation*:

$$\frac{\Delta \bar{\omega}_n}{\Delta t} + \bar{j}_{n+1} - \bar{j}_n = 0 \tag{33}$$

One proves from (33) that $\sum_n \bar{\omega}_n = 1$ holds for all t if it holds at $t = 0$, provided the current vanishes at the boundaries ($\bar{j}_0 = \bar{j}_{N+1} = 0$) for any t. This explains the term ’continuity’. For $N \gg 1$ we introduce the continuous variable $z = n/N$. Its probability density $\omega(z)$ is defined as

$$\omega(z) = N\bar{\omega}_n \quad \text{with} \quad z = \frac{n}{N} \tag{34}$$

The normalization factor N ensures that $\int \omega(z)dz = 1$ holds up to terms of order N^{-2}. In the following we will neglect those terms. We also replace (32) for the current by a continuous function $j(z)$:

$$\bar{j}_n = j(z - \frac{1}{2N}). \tag{35}$$

The offset $1/(2N)$ in the argument of $j(.)$ in eq. (35) is a commonly used technique in gauge theories on a discrete lattice [36] or the discretization of Maxwell’s equations for electric phenomena in numerical simulations [33]. It expresses that $\bar{j}_n$ connects the points $n-1$ and n and

[12] In the following we omit the obvious time argument t in $\bar{\omega}_n$ and $\bar{j}_n$.

therefore $j(z)$ should be defined in the middle of $(n-1)/N$ and n/N as is done in equation (35). We plug (35) into the difference $\bar{j}_{n+1} - \bar{j}_n$:

$$\bar{j}_{n+1} - \bar{j}_n = j(z + \frac{1}{2N}) - j(z - \frac{1}{2N}).$$

Using a third order Taylor expansion for $j(z)$ even powers in $1/N$ all vanish, and we get

$$\bar{j}_{n+1} - \bar{j}_n = \frac{1}{N}\left(j'(z) + \frac{1}{24N^2}j'''(z)\right). \tag{36}$$

The continuum version of the Master equation (33) is obtained by inserting the difference (36) and the definition (34) into (33) neglecting terms of order $1/N^2$

$$\frac{\Delta\omega(z)}{\Delta t} + \frac{\partial j(z)}{\partial z} = 0. \tag{37}$$

It is important to define the current (35) with the offset, otherwise we would have got terms of order $1/N$ which are cancelled in a tedious calculation by corresponding corrections in $j(z)$ lateron.
Now we express $j(z)$ in terms of $\omega(z)$. Rewriting equation (32) for $j(z)$ using (35) yields:

$$j(z - \frac{1}{2N}) = \bar{\omega}_{n-1}\pi_+(n-1) - \bar{\omega}_n\pi_-(n)$$

and using the definition (34) we obtain:

$$j(z - \frac{1}{2N}) = \frac{1}{N}\left(\omega(z - \frac{1}{N}) \cdot \pi_+(zN-1) - \omega(z) \cdot \pi_-(zN)\right).$$

In this relation only continuous functions appear and we can shift the variable z by $1/(2N)$ to obtain:

$$j(z) = \frac{1}{N}\left(\omega(z - \frac{1}{2N}) \cdot \pi_+(zN - \frac{1}{2}) - \omega(z + \frac{1}{2N}) \cdot \pi_-(zN + \frac{1}{2})\right). \tag{38}$$

Up to terms of order $1/N^2$ we can use the approximation $\omega(z \pm 1/(2N)) = \omega(z) \pm 1/(2N)\omega'(z)$ and get

$$j(z) = \omega(z)\,\frac{\pi_+(zN - \frac{1}{2}) - \pi_-(zN + \frac{1}{2})}{N} - \omega'(z)\,\frac{\pi_+(zN - \frac{1}{2}) + \pi_-(zN + \frac{1}{2})}{2N^2}. \tag{39}$$

$\pi_\pm(n)$ from equation (32) can be written as functions of z and N. Introducing the function $D(z)$

$$D(z) = 2bz(1-z) + \frac{1}{N}\left(a_1 - z(a_1 - a_2)\right) \tag{40}$$

the coefficient in the second term in eq. (39) can be written as

$$\frac{\pi_+(zN - \frac{1}{2}) + \pi_-(zN + \frac{1}{2})}{N^2} = D(z) + \frac{a_1 + a_2 - b}{N^2}. \tag{41}$$

With the definition

$$A(z) = a_1 - z(a_1 + a_2) \tag{42}$$

or equivalently (see appendix A2)

$$A(z) = (a_1 + a_2)(\bar{z} - z) \tag{43}$$

the first coefficient in (39) can be expressed in terms of $A(z)$ and $D(z)$ as:

$$\frac{\pi_+(zN - \frac{1}{2}) - \pi_-(zN + \frac{1}{2})}{N} = A(z) - \frac{1}{2}D'(z). \tag{44}$$

Inserting the expressions (44) and (41) without the $1/N^2$ term into (39) we finally get

$$j(z) = A(z)\omega(z) - \frac{1}{2}\frac{\partial}{\partial z}(D(z)\omega(z)). \tag{45}$$

Since Δt is also of order $1/N^2$ the difference in time can be replaced by a partial derivative which, together with the current (45), leads to the Fokker-Planck equation (5). The structure of the Fokker-Planck equation is correct up to terms of order $1/N^2$ which involve higher derivatives of $\omega(z)$. It is remarkable that no contribution of the herding mechanism to the drift term A(z) appears and the $1/N$ correction shows up only in the diffusion term $D(z)$.

A.2 Unconditional distribution of z

To compute the equilibrium distribution $\omega_0(z)$ in eq. (9), we use the standard formula:

$$\omega_0(z, N) = \frac{K}{D(z)} \exp\left(\int^z \frac{2A(y)}{D(y)} dy\right) \tag{46}$$

(see for instance Van Kampen [29]), derived from the condition $\partial j(z,t)/\partial z = 0$. The drift term D can be written up to order $1/N^2$ as a product

$$D(y) = 2b\left(y + \frac{\varepsilon_1}{2N}\right)\left(1 - y + \frac{\varepsilon_2}{2N}\right).$$

Writing $A(y)$ as

$$A(y) = b\left[\varepsilon_1\left(1 - y + \frac{\varepsilon_2}{2N}\right) - \varepsilon_2\left(y + \frac{\varepsilon_1}{2N}\right)\right],$$

the integral in the argument of the exponential is elementary with the result

$$\int^z \frac{2A(y)}{D(y)} dy = \varepsilon_1 \ln\left(z + \frac{\varepsilon_1}{2N}\right) + \varepsilon_2 \ln\left(1 - z + \frac{\varepsilon_2}{2N}\right). \tag{47}$$

Inserting this integral into (46) we arrive at the following formula for $\omega_0(z, N)$:

$$\omega_0(z, N) = K(\varepsilon_1, \varepsilon_2, N)\left(z + \frac{\varepsilon_1}{2N}\right)^{\varepsilon_1 - 1} \cdot \left(1 - z + \frac{\varepsilon_2}{2N}\right)^{\varepsilon_2 - 1} \tag{48}$$

The normalization constant K follows from $\int \omega_0(z, N)\, dz = 1$. If we neglect the $1/N$ term in D we can express K as the beta function:

$$B(\varepsilon_1, \varepsilon_2) = \int_0^1 z^{\varepsilon_1 - 1}(1 - z)^{\varepsilon_2 - 1} = \frac{\Gamma(\varepsilon_1)\Gamma(\varepsilon_2)}{\Gamma(\varepsilon_1 + \varepsilon_2)} \tag{49}$$

which leads to the distribution given in (10). Given (49) and (10), it is straightforward to compute the mean value of the distribution of z:

$$E[z] = \frac{a_2}{a_1 + a_2} = \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2}. \tag{50}$$

Note that in the text we used the convenient notation $E[z] \equiv \bar{z}$ (see eq. (7))

A.3 Unconditional distribution of uniform-noise model

The volatility v is a product of the noise variable $|\eta|$ distributed according to $\rho(\eta)$ and σ distributed according to $p(\sigma)$ from equation (23). For the distribution of v under the condition of a given value of σ we have

$$p(v|\sigma) = 2\frac{d|\eta|}{dv}\,\rho(\frac{v}{\sigma}) = \frac{2}{\sigma}\,\rho(\frac{v}{\sigma}). \tag{51}$$

The unconditional distribution of v is obtained by integrating $p(v|\sigma)$ over σ with the probability distribution $p(\sigma)$

$$\begin{aligned} p_u(v) &= 2\int p(v|\sigma)\,p(\sigma)\,d\sigma & (52)\\ &= 2\int \frac{1}{\sigma}\,\rho(\frac{v}{\sigma})\,p(\sigma)\,d\sigma. & (53)\end{aligned}$$

For a uniform distribution $\rho(\eta)$ we get

$$p_u(v) = \int_v^\infty \frac{1}{\sigma}\,p(\sigma)\,d\sigma.$$

Inserting $p(\sigma)$ we find

$$p_u(v) = \frac{1}{r_0^2 B(\varepsilon_1,\varepsilon_2)}\int_v^\infty \left(\frac{u}{r_0}\right)^{\varepsilon_1-2}\left(\frac{r_0}{u+r_0}\right)^{\varepsilon_1+\varepsilon_2} du. \tag{54}$$

If we we make the substitution $u = r_0 t/(1-t)$ the integral can be reduced to the incomplete beta function β defined by:

$$\beta(x;a,b) = \frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt \tag{55}$$

with B given in (49). Using the recursion formula $bB(a,b) = (a-1)B(a-1,b+1)$ this leads to eq. (26):

$$p_u(v) = \frac{1}{r_0}\frac{\varepsilon_2}{\varepsilon_1-1}\left[1-\beta\left(\frac{v}{v+r_0};\varepsilon_1-1,\varepsilon_2+1\right)\right]. \tag{56}$$

If either ε_1 or ε_2 is an integer the integral in (54) can be solved explicitly. In the case $\varepsilon_1 = 2$ we obtain a pure Pareto law

$$p_u(v) = \frac{\varepsilon_2}{r_0}\left(\frac{r_0}{v+r_0}\right)^{\varepsilon_2+1}. \tag{57}$$

The case $\varepsilon_1 = 3$ leads to a distribution with the same asymptotic behavior but modified at small and intermediate v:

$$p_u(v) = \frac{\varepsilon_2}{r_0}\left(\frac{r_0}{v+r_0}\right)^{\varepsilon_2+1}\frac{(\varepsilon_2+2)v+r_0}{2(r_0+v)}. \tag{58}$$

To prove the power law behavior of $p_u(v)$ for any $\rho(\eta)$ we rewrite equation (52) using the symmetry $\rho(\eta) = \rho(-\eta)$

$$p_u(v)v^{\varepsilon_2+1} = 2\int_0^\infty \frac{d\eta}{\eta}\rho(\eta)\,p(\frac{v}{\eta})\,v^{\varepsilon_2+1}. \tag{59}$$

The integral is well defined for any v. If the limit $v \to \infty$ exists we can interchange the limit $v \to \infty$ with the integration. Since for large v we have

$$p(\frac{v}{\eta})v^{\varepsilon_2+1} \simeq \frac{1}{r_0 B(\varepsilon_1, \varepsilon_2)}\,(r_0\eta)^{\varepsilon_2+1} \tag{60}$$

we find the limiting behavior

$$\lim_{v\to\infty} p_u(v)v^{\varepsilon_2+1} = \frac{2}{B(\varepsilon_1, \varepsilon_2)} \int_0^\infty d\eta\,(r_0\eta)^{\varepsilon_2}\,\rho(\eta). \tag{61}$$

This limit exists if the ε_2-th moment of $\rho(\eta)$ exists. If the moment does not exist, the asymptotic behavior is dominated by $\rho(\eta)$ for large $|\eta|$.

A.4 Proof of the Langevin equation

To prove equation (12) we begin with the following moments of the transition probability $\bar{\omega}(n', t+\Delta t|n, t)$:

$$M_k(n) = \sum_{n'} \left(n' - n\right)^k \bar{\rho}(n', t+\Delta t|n, t). \tag{62}$$

Inserting the probabilities (29) and (30) we get

$$\begin{aligned} M_0(n) &= 1 \\ M_k(n) &= \Delta t \begin{cases} \pi_+(n) - \pi_-(n) & \text{for} \quad k > 0 \quad \text{odd} \\ \pi_+(n) + \pi_-(n) & \text{for} \quad k > 0 \quad \text{even.} \end{cases} \end{aligned} \tag{63}$$

Introducing the continuum transition probability

$$\rho(z'|z) = N\bar{\rho}(n', t+\Delta t|n, t)$$

its moments are given by

$$\int dz'(z'-z)^k \rho(z'|z) = N^{-k} M_k(zN).$$

The characteristic function $C(q) = E[\exp(iq(z'-z))|z]$ can be computed from the moments

$$C(q) = \sum_{k=0}^{\infty} \frac{(iq)^k}{k!} M_k(zN) \tag{64}$$

$$= 1 + \Delta t\,(\pi_+(zN) - \pi_-(zN)) \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \left(\frac{iq}{N}\right)^{2k+1} \tag{65}$$

$$+\Delta t\,(\pi_+(zN) + \pi_-(zN)) \sum_{k=1}^{\infty} \frac{1}{(2k)!} \left(\frac{iq}{N}\right)^{2k}. \tag{66}$$

Using the identities

$$\pi_+(zN) - \pi_-(zN) = NA(z) \quad \text{and} \quad \pi_+(zN) + \pi_-(zN) = N^2 D(z)$$

and neglecting terms of order $\Delta t/N^k$ with $k \geq 2$ we find

$$C(q) = 1 + \Delta t \left(iqA(z) - \frac{q^2}{2} D(z) \right) . \tag{67}$$

If we further assume $a_i \Delta t \ll 1$ and $b\Delta t \ll 1$ this characteristic function C corresponds to a Gaussian for $\omega(z'|z)$. This implies that z' for given z is Gaussian distributed with mean $z + \Delta t\ A(z)$ and variance $\Delta t\ D(z)$, which is the content of (12). Note that the above inequalities are less restrictive than (2).

A.5 Discreteness of price records

In the estimation procedure, an important issue arises from the discrete nature of the records of prices. Close inspection of our empirical time series shows a variability of the precision of the entries that changes over time. The DAX, for instance, exhibits a decimal precision for the first 7064 data points, which is increased to the second digit afterwards; the price of Gold is recorded, for a long period, with a precision of a quarter of dollar, but changes to smaller increments later on. The discreteness of prices is reflected in an artificial large variability of small returns ($|r| < E[|r|]$), where certain values are more frequent than others, due to some sort of threshold effect. Consequently, the empirical proxy for volatility $|r|$ is poorly approximated by a continuous variable in the region close to zero, while the key variable in our model is continuous over its entire range (see eq. 25). In order to induce continuity, one could add a small white Gaussian noise with mean zero to the time series of prices, to avoid the discreteness of small returns. If one takes a standard deviation of the noise of 0.05, this should be sufficiently small to just marginally affect the original price level, but large enough to avoid the discreteness in the last decimal[13]. To evaluate the impact of the noise in the estimation of ε_1, ε_2 and the values of the likelihood, we have performed 100 Monte Carlo simulations for different realizations of the noise entries for all the time series, and for the case of both the uniform and spin noise models. It turns out that the estimates are not very different across replications of this Monte Carlo setting. The resulting standard deviation of the parameters ε_1 and ε_2 is, in fact, much smaller than the associated error of a single estimate due to the Maximum Likelihood method. For example, in the case of Gold and spin-noise model, the average value of ε_1 over 100 simulations is 4.7 and the standard deviation is 0.1, while the error associated to the ML estimate is 0.4. (further details upon request). Therefore, estimates from a particular realization of the modified data sets should be quite representative. Table 3 shows representative numbers from one such particular realization.

A comparison of Tables 1 and 2 with Table 3 shows that when adding noise to the prices, the uniform-noise model is preferred in all cases over the spin-noise model on the basis of its higher likelihood value. Therefore, the superior performance of the simpler multiplicative spin-noise in describing the raw data seems, in fact, to be due to the discreteness of the price records. The additive noise, in fact, destroys the artificial abrupt variations of small returns, generating a flat region around the origin, which, then, is responsible for the large entries of ε_2 estimates in the spin-noise model, cf. Table 3 (except for Gold case, which does not exhibit such an extended

[13] In the case of Gold price the standard deviation is 0.1, since the minimum increment is 1/4.

	Data Set	$\hat{\varepsilon}_1$	$\hat{\varepsilon}_2$	$-\ln L_{\varepsilon_1,\varepsilon_2}$	$\hat{\varepsilon}$	$-\ln L_\varepsilon$	p-value
Spin	Gold	1.14 ± 0.03	4.8 ± 0.4	4944.9	1.99 ± 0.02	5181.2	0.00
	DB	1.20 ± 0.02	13.0 ± 2.0	6726.2	2.24 ± 0.02	7228.5	0.00
	Siemens	1.22 ± 0.03	16.0 ± 3.0	6621.6	2.30 ± 0.03	7117.0	0.00
	DAX	1.34 ± 0.02	15.0 ± 2.0	9592.0	2.46 ± 0.02	10261.1	0.00
Uniform	Gold	3.3 ± 0.4	3.9 ± 0.4	4943.2	3.6 ± 0.1	4943.5	0.43
	DB	7.2 ± 1.6	4.7 ± 0.5	6691.5	5.4 ± 0.2	6292.6	0.13
	Siemens	7.3 ± 1.5	5.0 ± 0.5	6590.9	5.7 ± 0.2	6591.8	0.18
	DAX	15 ± 5	4.9 ± 0.3	9533.9	6.7 ± 0.2	9542.8	0.00

Table 3: *Estimated parameters for the spin-noise-model and the uniform-noise model with the additive noise in the time series of prices.*

plateau). A close inspection of the estimates shows that the uniform-noise model is more robust against additive noise in the prices than the spin-noise model, as its estimates are only marginally affected by the modification.

Finally, this estimation exercise reveals that discreteness of the price levels, even for daily data, can have a sizeable influence on statistical estimates. The effect is especially noticeable if one works with parametric models capturing the entire distribution, since in such models small returns constitute a large fraction of the sample size.

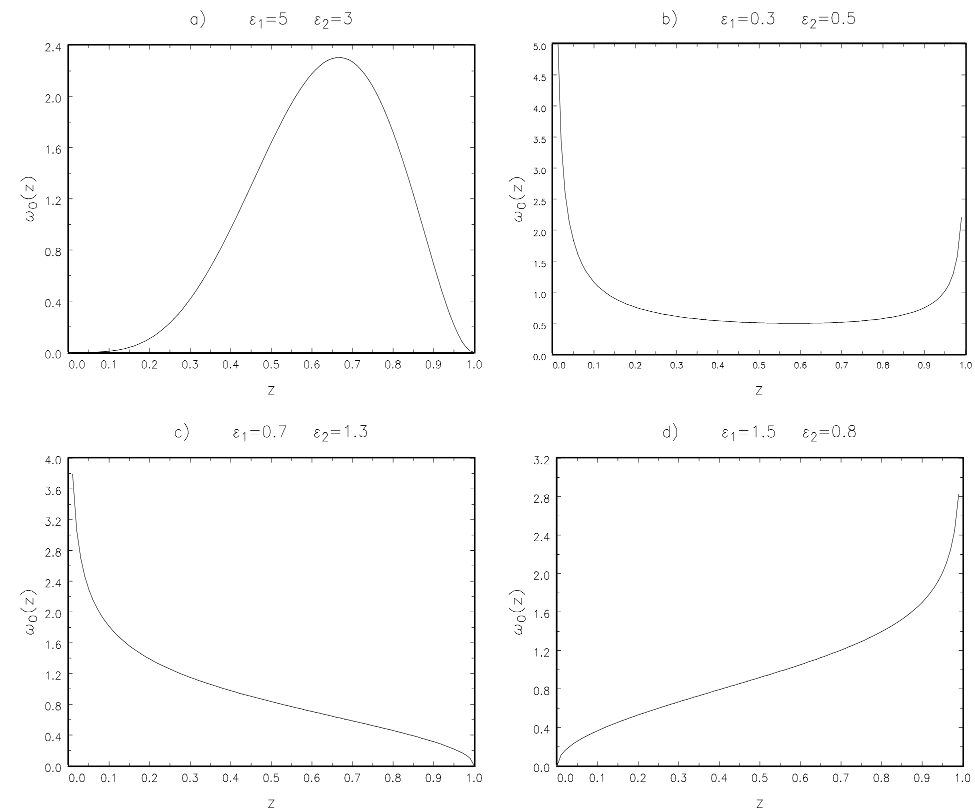


Figure 1: *The four panels show different probability densities derived from eq. (10) for several choices of the parameters ε_1 and ε_2. In the panel (a) a uni-modal distribution, in (b) a bi-modal distribution and in panels (c) and (d) two cases of asymmetric monotonic distributions are shown.*

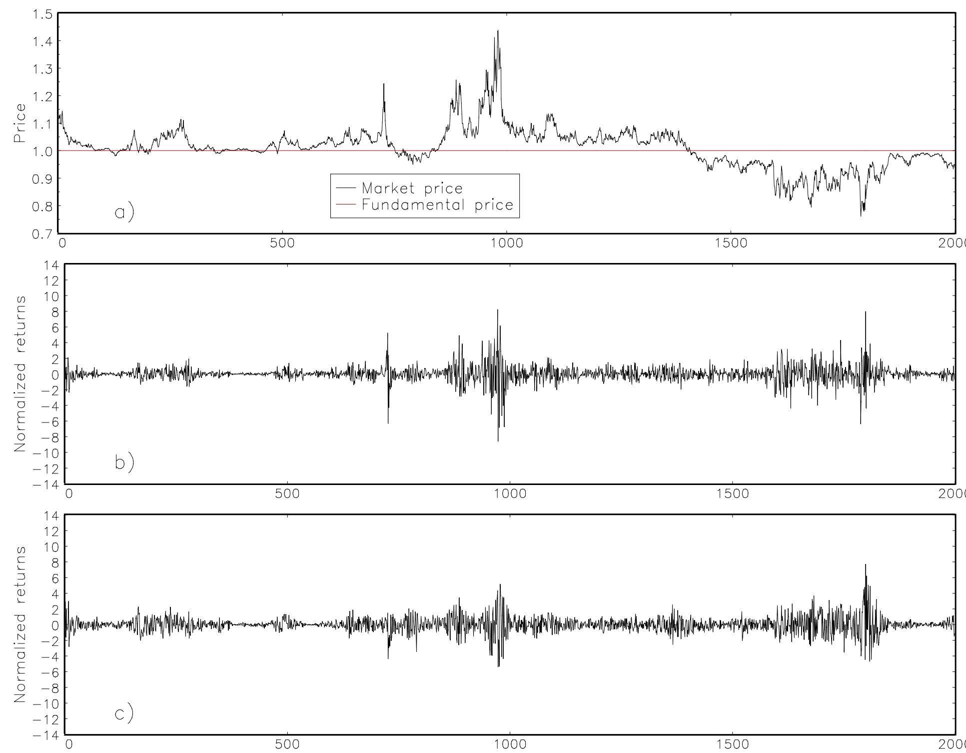


Figure 2: *The upper panel shows a simulation of the price derived from equation (16) using a uniform distribution of* η *over* $[-0.1, +0.1]$. *In the middle panel the returns corresponding to (18) are shown. The lower panel shows the returns obtained by using eq. (19). As parameters we have chosen* $p_f = 1$, $r_0 = 0.1$ *and* $\Delta t = 1$. *The herding parameters* $\varepsilon_1 = 3$, $\varepsilon_2 = 4$ *and* $b = 0.003$ *are representative for the case of the price of Gold.*

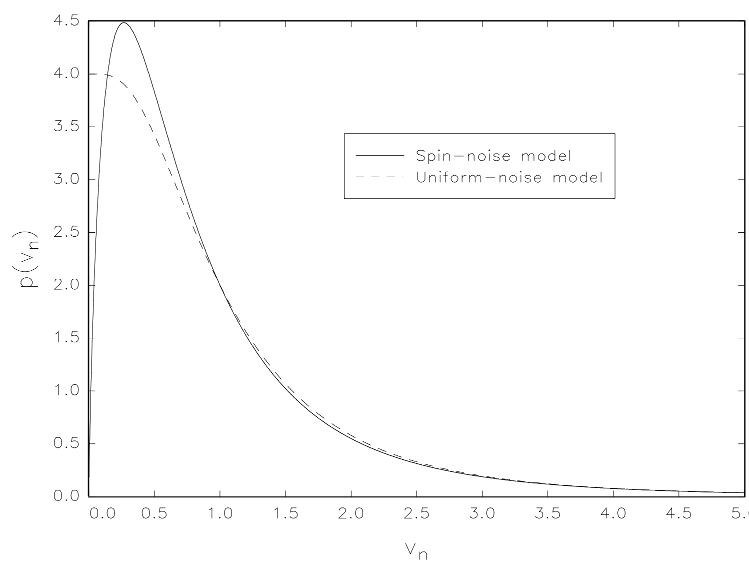
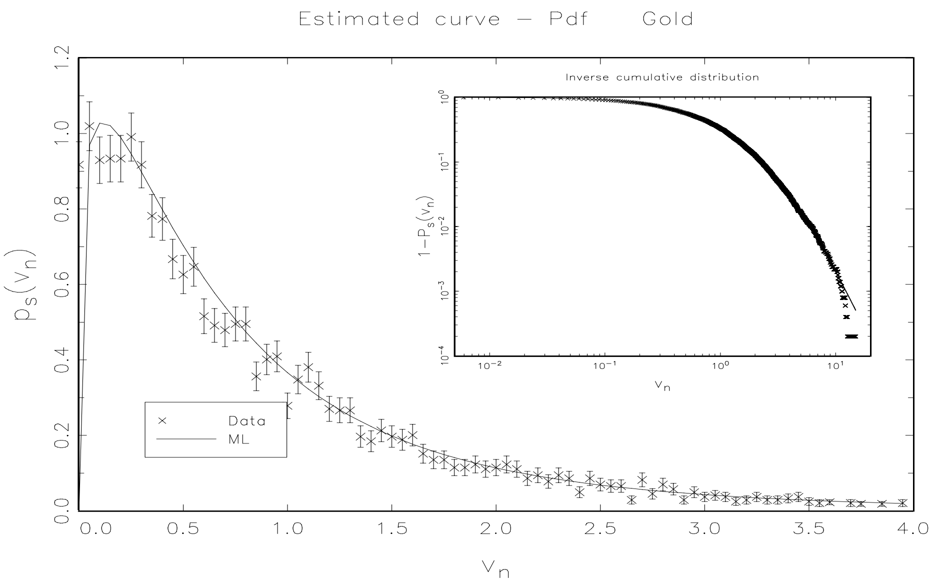
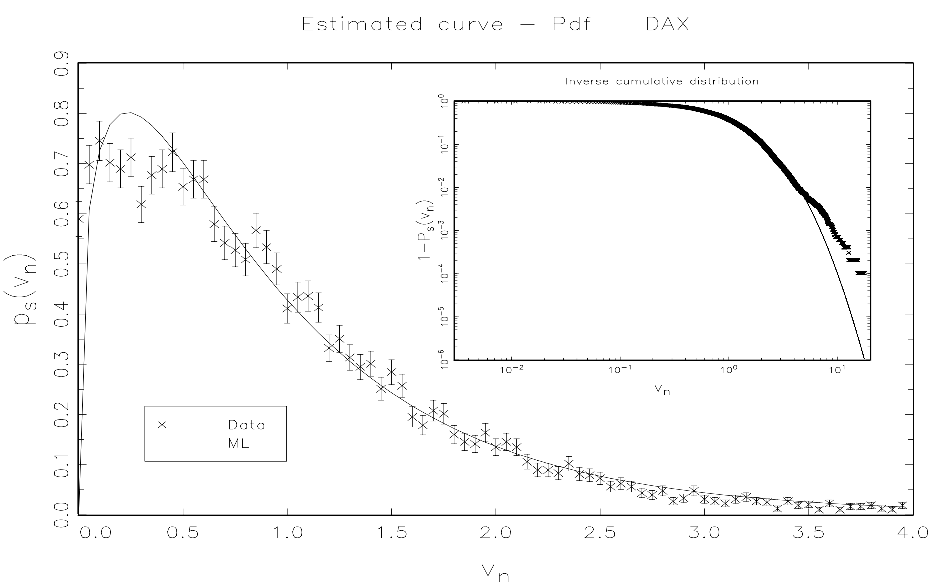


Figure 3: *The figure compares the probability density functions of the spin-noise-model and the uniform-noise-model as functions of the normalized volatility* $v_n = v/E[v]$. *The parameters for the spin-noise-model are* $\varepsilon_1 = 1.8$ *and* $\varepsilon_2 = 4$, *for the uniform-noise-model* $\varepsilon_1 = 6$ *and* $\varepsilon_2 = 4$. *For this choice of* ε_1 *both densities have the same mean and second moment.*

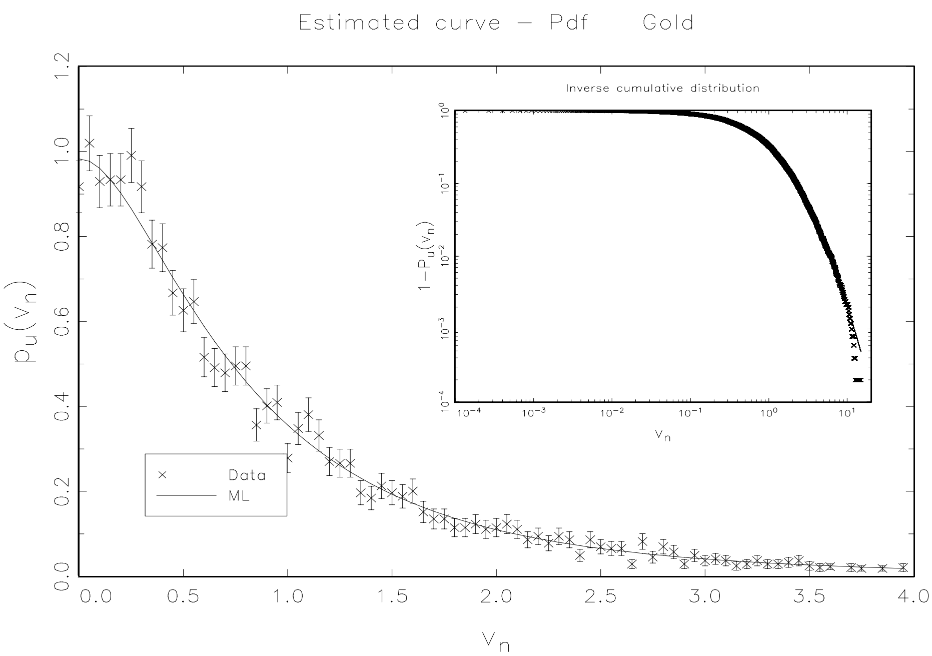


(a) Gold

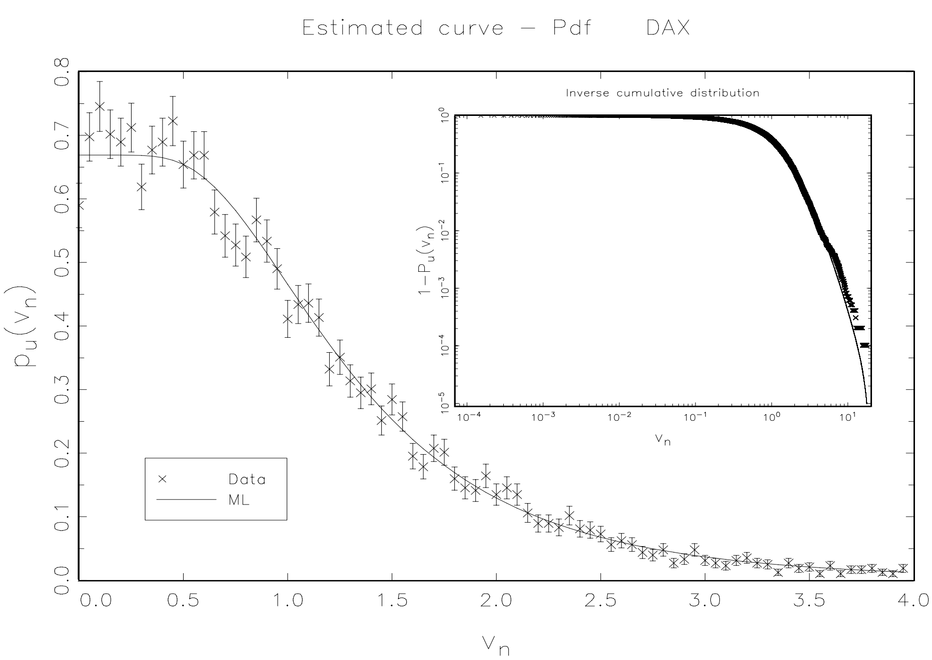


(b) DAX

Figure 4: *Spin-noise model: In panel (a) the probability density of the time series of Gold is shown as a function of the normalized volatility v_n together with the distribution (23) with estimated parameters given in Table 1. The inlet shows for both cases the cumulative distribution $P(|r| > x)$ in a log-log plot. Panel (b) shows the same comparisons for the time series of DAX. Note the deviation of the theoretical curve in the outer part of the cumulative distribution and the poor fit in the central part of the pdf in the case of the DAX. The graph also shows intervals of $\pm 1x$ standard deviation, which are computed assuming a normal distribution for the entries in every bin of the histogram. The same procedure is applied also in Figure 5.*



(a) Gold



(b) DAX

Figure 5: *Uniform-noise model: In panel (a) the probability density of the time series of Gold is shown as a function of the normalized volatility* v_n *together with the distribution (26) with estimated parameters given in Table 2. The inlet shows the cumulative distribution* $P(|r| > x)$ *in a log-log plot. Panel (b) shows the fit for the time series of DAX. In contrast to the spin-noise model the uniform-noise model gives a good fit for all* v_n*.*

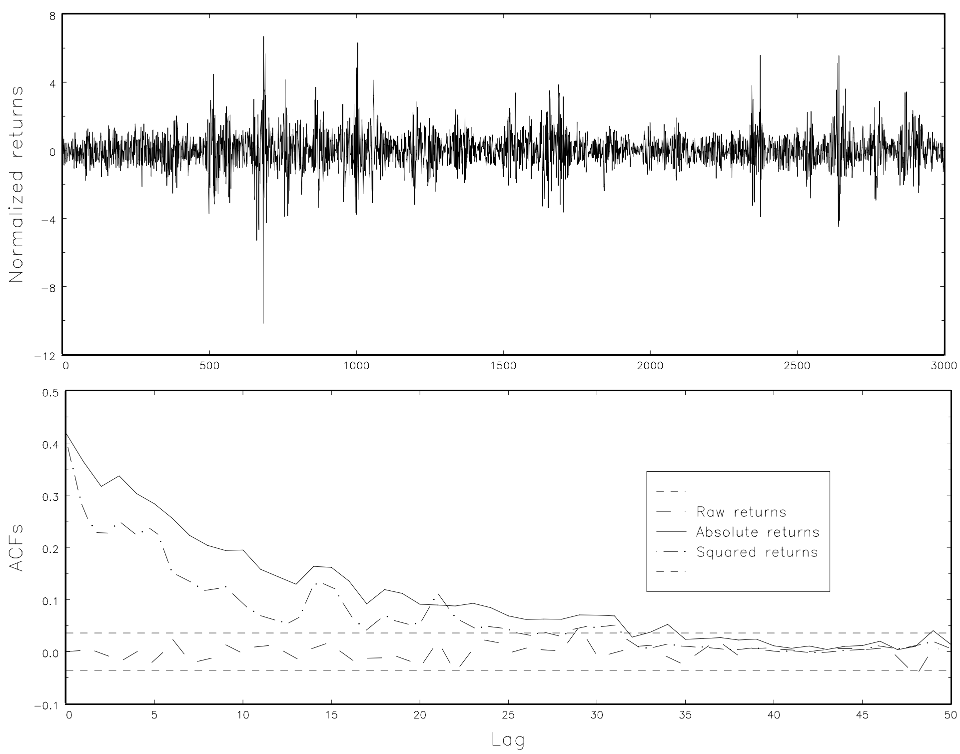


Figure 6: *The upper panel shows simulated returns; autocorrelation function of raw, absolute and squared returns are exhibited in the lower panel. Underlying parameters are:* $b = 0.0025$, $\varepsilon_1 = 16$, $\varepsilon_2 = 4.9$, $r_0 = 1$ *and* $p_F = 1$. *The time series is derived from (19) using the uniform-noise model and* $z(t)$ *simulated via the Langevin equation (12).* ε_1 *and* ε_2 *are representative for the case of DAX.*