Simple stochastic modeling for fat tails in financial markets

Hans-Georg Matuttis

Department of Mechanical Engineering and Intelligent Systems, The University of Electro-Communications, Tokyo 182-8585, Japan hg@mce.uec.ac.jp

1 Introduction

Stochastic theories for the description of financial markets, e.g. via the Langevin equation, are usually based on terms with uncorrelated noise. The purpose of this paper is to investigate whether correlations in the market can introduce "fat tails" in Random Walk (RW)-like models as well as a narrow center in the distribution as found in the S&P500[1]. Using the discrete version for the RW

$$x_i = \mu + x_{i-1} + \xi_i \tag{1}$$

as a starting point, where μ is the "trend" and ξ_i is Gaussian/ normally distributed uncorrelated noise, we focus on de-trended random walks with $\mu = 0$. Correlated noise for our purpose can easily be produced by "mixing" successive random numbers from the Gaussian random number sequence in a weighted average, e.g.

$$\tilde{\xi}_i = \sum_{j=i-M}^{j} (1-\epsilon)\xi_j + \epsilon\xi_{j-1}. \tag{2}$$

For different ϵ, we get different distributions all narrower than the standard normal distribution, which nevertheless still can all be fitted to a Gaussian with standard deviation $\sigma < 1$. Therefore the narrow center in the S&P500-Data of Mantegna et al.[1] may be explained by short-term/local correlation in the "noise" of the random walk, which corresponds to the correlated information/noise the players receive in the financial world. In the remaining part Gaussian white noise with reduced variance is used instead of explicitly correlated noise, which nevertheless cannot account for the fat tails, because it does narrow, not widen, the distributions. From here on, we use "correlated noise" $\tilde{\xi}_i = a\tilde{\xi}_i$, from a Gaussian distribution ξ_i with the standard deviation $\sigma = a < 1$, so eqn. (1) becomes

$$x_i = x_{i-1} + a\tilde{\xi}_i. \tag{3}$$

2 Technical analysis, correlation and fat tails

A possible reason for much wider market swings than local correlation in market data is the synchronous reaction of many market players to signals from technical/chart analysis: Erratic chart data are averaged or fitted (in a very loose sense of the word), and the market players adapt their expectations and strategies accordingly, which is reflected in the price evolution. As the basis for our "technical analysis", we will use "moving averages", which are computationally easier accessible than chart formations like "double tops", "shoulder-head-shoulder-configurations" or "resistance-lines", on which market analysts often among themselves don't agree.

2.1 The Model: RW with Moving Averages

We set up our model equation for the technical analysis random walk (TARW)

$$x_i = x_{i-1} + a\xi_i + b\eta_i, \tag{4}$$

with Gaussian distributed mini-trends η_i, which react to the crossing of the averages in "bullish" or "bearish" manner as follows: The standard-normal-distributed "mini-trends" η_i with prefactor b don't change as long as the chart x_i does not cross the moving average $\langle x_i \rangle_N = \frac{1}{N} \sum_{j=i-N}^{i} x_j$ from the previous N market transactions:

$$\eta_i = \eta_{i-1} \text{ if } \eta_{i-1} > 0 \text{ and } x_i > \langle x_i \rangle_N \tag{5}$$

$$\eta_i = \eta_{i-1} \text{ if } \eta_{i-1} < 0 \text{ and } x_i < \langle x_i \rangle_N \tag{6}$$

Whenever the chart x_i crosses the moving average $\langle x_i \rangle_N$, a new "mini-trend" η_i with sign opposite the previous one is selected:

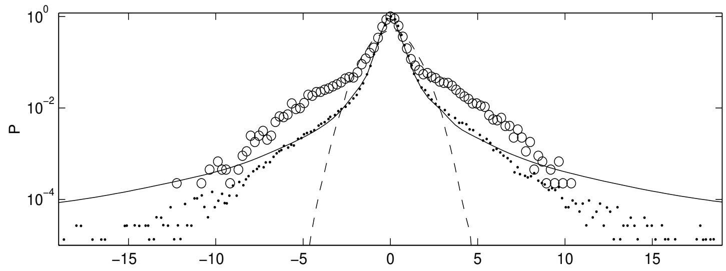
$$\text{new } \eta_i < 0 \text{ if } \eta_{i-1} > 0 \text{ and } x_i < \langle x_i \rangle_N \tag{7}$$

$$\text{new } \eta_i > 0 \text{ if } \eta_{i-1} < 0 \text{ and } x_i > \langle x_i \rangle_N. \tag{8}$$

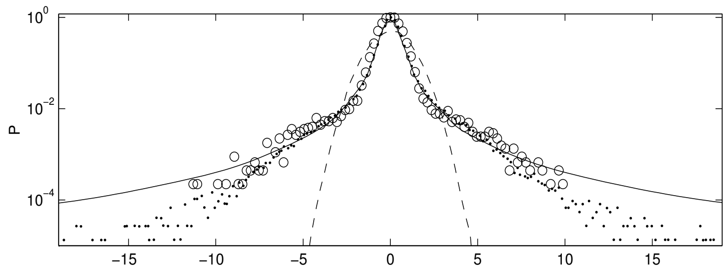
The result for the distribution[1] of such a TARW is shown in Fig. 1(a) with fit parameters $a = 0.4$, $b = 4$. The inclusion of technical-analysis-like decision-making leads to fat tails and also a curvature change similar to the one seen in the empirical S&P500-data. Nevertheless, the fat tails in Fig. 1(a) branch

[1] This and all the following distributions have been plotted with the maximum probability normalized to one to allow the simple comparison with the empirical data, with 300000 time steps, 8000 equilibration steps before the first measurement, and moving averages of length $N = 5000$. Distributions did not change significantly if the moving averages were computed for $N = 5000$, $N = 50000$, or $N = 500$. Our "technical analysis" is "time-scale-free" in the sense that the sum of Gaussian random numbers produces a Gaussian again. The structure of the time-series itself varied considerably with the length of the moving averages.

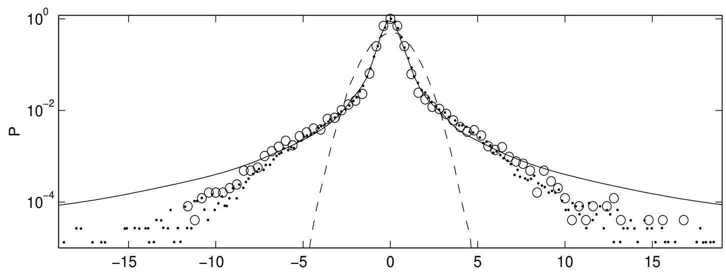
out at too high probabilities compared to the data from Mantegna[1]. The TARW-mechanism can be interpreted as the superposition of two Gaussian distributions, one narrow, one wide, where the wider distribution is selected very rarely, due to the relative motion of chart and moving average. Though we can conclude that the "technical analysis" (i.e. a quantifiable herd-like behavior produced by the herd itself) is a very efficient way to obtain "fat tails" when introduced in random walk models, the TARW-mechanism is not the final answer: The simulated curve deviates significantly from the empirical data, the "onset" of the fat tails is much too high.



(a) After eqn. (4)-(8) with $a = 0.4$, and $b = 3$,



(b) After eqn. (9) -(8) with $a = 0.4$, $b = 1.4$ and $D = 25$.



(c) After eqn. (9)-(8), (10)-(15), $a = 0.32$, $b = 1$ and $D = 15$.

Fig. 1. Comparison for TARW (a), DTARW (b) and DTARWB (c) (moving averages of length $N = 5000$, circles) with the empirical data for the S&P500 (full dots), the fitted Levy-distribution (solid line) and Gaussian (broken line) after Ref. [1]

2.2 Introducing Delay

In our TARW-Model exchange, each transactions relates to the previous one as reference level. In reality, in the trader-based New York Stock, transactions for the S&P500 do not take place instantaneously, but there is a certain time lag between the transaction decision, the completion of the transaction and the actual display of the new price. Data for the duration of this delay are hard to come by, Wall Street seems to be quite reluctant to talk it. To mimic the time lag, we introduce the delay time D so that between display-timestep j and the next display-timestep $j+D$ all players base their transaction decision on the data at the display-timestep j

$$x_{j+i} = x_j + a\xi_{j+i-1} + b\eta_{j+i-1}, \quad D \geq i \geq 1, \tag{9}$$

all other terms are defined as in the TARW-model. For this "delayed technical analysis random walk model"[2] (DTARW), the delay parameter of D means that the average number of transactions between a transaction and the reference price is $D/2$. The "best fit" for the empirical S&P500-data obtained with the parameters $a = 0.4$, $b = 1.4$ and delay $D = 25$ is shown in Fig. 9. For the DTARW, the variance b for the "mini-trend" has decreased to 1.4, from 4 for the TARW, which is an improvement for the sake of plausible simulation parameters. Nevertheless, the simulation data are higher than the empirical distribution for the probability interval between 10^{-2} and $10^{-4.5}$. Another setback of the model is that we had to introduce an additional simulation parameter, the delay D (again the distribution did not change with the length of the moving average), so the DTARW needs three fit parameters, which should account for any curve symmetric to the y-axis with two different curvatures, not very satisfying from the point of data modeling.

2.3 Bollinger Bands

Our models up to here[2] allowed only a change from "down" to "up" in the mini-trends, and vice versa. A convenient way to implement a steepening or flattening of a trend in the same direction is via Bollinger Bands, so that

$$\eta_i > \eta_{i-1} \ \text{ if } \eta_{i-1} > 0 \text{ and } \ x_i > \langle x_i \rangle_N + 2\text{STD}(x_i) \tag{10}$$

$$\eta_i = \eta_{i-1} \ \text{ if } \eta_{i-1} > 0 \text{ and } \ \langle x_i \rangle_N + 2\text{STD}(x_i) > x_i > \langle x_i \rangle_N \tag{11}$$

$$\text{new } \eta_i < 0 \ \text{ if } \eta_{i-1} > 0 \text{ and } \ x_i < \langle x_i \rangle_N \tag{12}$$

$$\eta_i < \eta_{i-1} \ \text{ if } \eta_{i-1} < 0 \text{ and } \ x_i < \langle x_i \rangle_N - 2\text{STD}(x_i) \tag{13}$$

$$\eta_i = \eta_{i-1} \ \text{ if } \eta_{i-1} < 0 \text{ and } \ \langle x_i \rangle_N - 2\text{STD}(x_i) < x_i < \langle x_i \rangle_N \tag{14}$$

$$\text{new } \eta_i > 0 \ \text{ if } \eta_{i-1} < 0 \text{ and } \ x_i > \langle x_i \rangle_N, \tag{15}$$

[2] This model is not a delay differential equation in the conventional sense where each x_i would be computed from the previous x_{i-D}.

where the "mini-trends" η_i are chosen with absolute value larger than the previous one, but the same sign, if the market breaks out beyond the Bollinger bands (twice the standard deviation above and below the moving average[3]). The result for this "Delayed technical analysis random walk with Bollinger Bands" (DTARWB) can be seen in Fig. 1(c) for $a = 0.32$, $b = 1$ and delay interval $D = 15$. The tails have been significantly straightened in comparison to the DTARW-case. Not only are the simulated data very close to the empirical data, the scattering in the tails and the convex part are quantitatively well reproduced. The delay-parameter D to model the empirical S&P500-data has been reduced to 15 in the DTARWB from 25 for the DTARW, so between a transaction decision and the display of the price on average 7 to 8 transactions have occurred. More important: The previous fit-parameter b has been reduced to unity, which means that though the returns $x_i - x_{i-1}$ are not Gaussian distributed, the mini-trends η_i are. Therefore, our DTARWB has only two free fit-parameters, the local correlation a and the delay D. Algebraically, a curve with two different curvatures needs at least three fit parameters, so our DTARWB-theory seems to supply some additional information.

3 Summary and Conclusions

We have shown that the return distributions observed in the S&P500 can be obtained for a random-walk which reacts to moving averages in the technical analysis sense. Characteristic ingredients are mini-trends in accordance with moving averages, which lead to fat tails, delay in trading, which shifts the tails lower in the distributions and a reaction to break-outs of the market (in our case, Bollinger bands) which straighten out the curvature of the tails. Though the chart values of the S&P500 are not Gaussian distributed, it is the mini-trends which follow a random walk/ Gaussian distribution with unit variance. This leaves considerable doubts about the actual "efficiency" of the market. It will be interesting to analyze in market data whether the local correlation a is universal, the mini-trends η_i are always standard-normal-distributed and whether the delay-parameter D is shorter in markets with electronic trading.

References

1. R. N. Mantegna, H. E. Stanley, Scaling Behavior in the Dynamics of an Economic Index, Nature 376, p. 46-49 (1995).
2. H.-G. Matuttis, Simple stochastic modeling for financial markets, Meeting Abstracts of the Physical Society of Japan, Vol. 59, Issue 1, part 2, p. 346, March 2004.

[3] Changes in the definition of the Bollinger bands from a prefactor two to three gave only moderate changes in the distributions. A publication with a more detailed discussion of the scattering is in preparation.