An Introduction to Econophysics: Correlations and Complexity in Finance
by Rosario N. Mantegna and H. Eugene Stanley
Cambridge University Press, 2000
x+148 pp., references, index
USD34.95, ISBN 0-521-62008-2

We all know the argument for the efficient market hypothesis (EMH). Consider an equilibrium financial market, populated by rational agents. The price an agent will pay for a financial instrument is its net present value to him — his estimate of future returns, discounted for time-preference and risk. Since the agents are rational, their estimates of future returns will accurately incorporate everything they know. Hence prices change only when tastes (for risk and for time) change, or when unpredictable information arrives. If a piece of news could have been predicted, it would have been, and already reflected in prices. Hence, given rational expectations and market efficiency, prices are unpredictable, i.e., price changes are independent random variables. A few more hypotheses lead us to expect that changes in the logarithms of prices are independent, identically distributed Gaussians, and so that financial time series should look like (multiplicative) random walks.

This is a strong, elegant and fruitful hypothesis, marred only by being quite wrong. Traders are not perfectly rational, and could not be. Markets are inefficient, both in microscopic ways (e.g., transaction costs), and in global ones, as shown by persistent "anomalies". Even over very short times, log price changes are non-Gaussian — the peaks are too sharp, while the tails are "fat," decaying too slowly, roughly as a power law. Worst of all, financial time series *are* predictable; though correlations in the price changes themselves quickly decay, nonlinear functions of price changes stay correlated over very long times. If markets are efficient, then prices are totally unpredictable; but prices are predictable; therefore markets are not efficient.

There are a couple of ways of moving forward from this point (not counting staying put, out of loyalty to the EMH). Behavioral finance, for instance, aims to replace the rational agents of the EMH — hedonistic sociopaths with supercomputers for brains — with more plausible, if not necessarily more flattering, portraits of market participants. A more pragmatic and modest venture is to find out what, *exactly*, financial time series are like, if they're not as the EMH says they should be.

This is where econophysics comes in. Broadly understood, it's the attempt to understand economic phenomena with mathematical tools from statistical physics. It's one of the things statistical physicists have taken to doing recently, other than statistical physics, partly because they can't stand the prospect of solving Yet Another Spin System. If you look at the on-line archives for econophysics (`http://www.unifr.ch/econophysics/`), you'll see papers about the distribution of wealth and income, the growth of networks and organizations, the "thermodynamics" of teenage pregnancy, models of simple adaptive agents in competition. You'll also see paper after paper about the statistics of financial markets — and almost nothing about any other sort of market. There are two major reasons for this. First, financial markets produce huge volumes of high-quality data. It might be that the econophysicists' tools are even better suited to (say) studying the dynamics of industrial competition, but they'll never find the kind of data on manufacturing that they can get for financial markets. Second, that's where the money is.

The book before us is, the publishers claim, the first English monograph on econophysics. The authors are leading researchers in the field, and were well-regarded statistical physicists before that; Stanley, in particular, was influential in developing and propagating the modern theory of phase transitions and critical phenomena in the 1970s. They have worked in several areas of econophysics, including a fascinating series of studies on the scaling behavior of organizational growth, where they find a sort of quantitative version of Parkinson's Law. Still, in this book, they present econophysics as the phenomenology of financial time series: the study of what the series look like, statistically, with no consideration of what mechanisms make them that way.

After opening with the EMH, our authors move quickly into material on stochastic processes, specifically random walks, and why the central limit theorem says that summing up any collection of independent, identically-distributed price changes, with finite variance, will give you a Gaussian. They even quote some results on the rate at which such sums converge on Gaussians, which surprisingly few books cover. Gaussians are a kind of attractor in the space of random variables — keep averaging independent copies of a variable with finite variance, and you'll approach a Gaussian. But Gaussians are not the only attractors in the space of random variables. The others, with infinite variance, have Lévy distributions, and are reached by averaging IID random variables with infinite variances. These are attractive in modeling because of their extremely fat, power-law tails, resembling the power-law tails of price changes.

Of course, real price changes have only finite variance, so one needs to apply some kind of cut-off to the power-law. This means one recovers the usual central limit theorem in the long run. The trick is making that run long enough. The authors' favorite way to do this is to model a price series by a "truncated Lévy flight" (TLF), a sum of IID variables. Each increment must fall within a certain fixed support, but has a Lévy distribution inside the support. TLFs look like Lévy distributions at short times, but like Gaussians over long times (since the variance of each step is finite). The way in which TLFs cross-over from power-law tails to Gaussian tails is pretty close to the way financial data do, but they fail to capture other aspects of financial time series, which are related to the non-independence of the data.

This brings up the last sort of stochastic process the authors discuss, "autoregressive conditional heteroskedastic" (ARCH) models. As before, we have a sum of increments. Each increment is a random variable, generally assumed to be a Gaussian with zero mean. In a k^{th} order ARCH model, the variance of the increment is not constant but is a weighted sum of the squares of the last k increments. (In a generalized ARCH model, we add the past variances into the sum.) This makes the increments neither independent nor identically distributed. Because correlations can be rigged to extend over long times, ARCH and GARCH processes can mimic the power-law correlation of actual financial data. While ARCH and GARCH can be very good at modeling the distribution of price changes over a fixed time horizon, simply aggregating the model process will not give us the right distribution over a longer horizon. The moral, the authors say, is that there is no completely acceptable model of the statistics of financial time series.

Analyzing individual time series occupies the first three quarters of the book. The last quarter concerns multiple series. After explaining the idea of cross-correlation, it shows how to calculate correlation coefficients. It then describes a simple, but fairly effective, algorithm for the hierarchical clustering of stocks based on their correlations. The last two chapters explain Black-Scholes option pricing theory, and, somewhat sketchily, ways people have modified the theory to accommodate deviations from Black-Scholes assumptions. The authors present no other applications.

You don't *really* need to know any physics to follow this book. The closest Mantegna and Stanley come to using any physics is in a chapter (11) where they examine, and dismiss, other econophysicists' analogies between financial markets and turbulence. (The dismissal is convincing.) Despite the subtitle, "complexity" only shows up on two pages (pp. 11–12), where they discuss the algorithmic information content of financial time series. What they say there is technically accurate, quite misleading, and harmlessly inconsequential[1]. The tools they introduce aren't so much from statistical physics as from stochastic process theory, and thus common to all fields which make serious use of probability. What's "physical" is their general strategy, their way of making idealized models to capture, piece-meal, significant experimental phenomena.

Finance specialists hoping for enlightenment from physics will be disappointed; the average reader of *Quantitative Finance* will find little new. But the book seems aimed the other way, at physicists interested in economics, and for them it would make a good introduction to finance. The writing is clear and friendly, the production-values high (except that all minus signs have been dropped from axes labels in figures!), and the guides to further reading excellent. They will find it well worth their time and money; professionals should save theirs.

[1]They correctly state that, under the EMH, financial time series are IID, and so algorithmically incompressible. They then say that a series's unpredictability is a manifestation of the fact that it "carries a large amount of nonredundant information". True, but only in the sense that the same thing is true of a sequence of coin tosses. It takes a lot of bits to specify an exact sequence of heads and tails, but it takes no bits of memory to generate such sequences statistically. (To model coin-tossing, toss a coin.) So it does not follow that a financial time series contains "valuable and important economic information". By the EMH, the series contains no information about future price changes, and the only information it has about future prices is contained in the current price. The structural or statistical complexity is thus negligible.