Multi-Scaling Properties of Asset Returns:

An Assessment of the Power of the 'Scaling Estimator'

Thomas Lux*

Abstract: We propose a surrogate data test for the significance of multi-scaling of financial data inferred from popular 'scaling estimators'. Applying this test to four long financial time series (the stock price indices of the New York and Frankfurt Stock Exchanges, the U.S.$-Deutsche Mark exchange rate and the price of gold), we can in no case reject the null hypothesis that the apparent curvature of both the scaling function of moments and the Hölder spectrum are spurious results generated by the particular fat-tailed distribution of innovations characterizing these financial data. Given the overwhelming evidence in favor of different degrees of long-term dependence in the powers of returns, we interpret this inability to reject the null hypothesis as a lack of discriminatory power of the standard approach for detection of multi-scaling rather than as a true rejection of multi-scaling in financial data.

Keywords: multi-fractality, long-range dependence, scaling estimator

JEL classification: C20, G12, C15

* Helpful discussions with T. Kaizoji, K. Matia, G. Stanley and K. Yamasaki are gratefully acknowledged. This paper has been written during a sabbatical stay at International Christian University Tokyo whose great hospitality is gratefully acknowledged. The author also gratefully appreciates financial support by the Landeszentralbank Schleswig-Holstein and the Japan Society for the Promotion of Science.

March 2003

Address of author:
Department of Economics, University of Kiel, Olshausenstr. 40, 24118 Kiel, Germany
lux@bwl.uni-kiel.de

1. Multi-Scaling and Multi-Fractality in Asset Returns

Mandelbrot, Calvet and Fisher (1997), Mandelbrot (1999) and Calvet and Fisher (2002) have proposed a compound model consisting of a multi-fractal time transformation θ_t and an incremental Brownian motion B_H as a new model of financial returns (r_t):

(1) $$r_t = B_H\,[\theta(t)]$$

The major new innovation of this model is the use of a multi-fractal cascade as a transformation of chronological time into 'business time'. Put differently, one might also simply interpret the multi-fractal component as a process governing instantaneous volatility. In their original form, multi-fractal cascades are operations performed on probability measures. The 'cascade' starts with assigning uniform probability to the interval [0,1]. In the elementary case of a so-called Binomial cascade, this interval is first split up into two subintervals of equal length, which receive a fraction m_0 and $1 - m_0$, respectively, of the total probability mass. In the next step, each subinterval is again split up into two subintervals, which again receive fractions m_0 and $1 - m_0$ of the probability mass of their 'mother' intervals. In principle, this procedure is, then, repeated *ad infinitum*. While the probabilities m_0 and $1 - m_0$ are constant over the evolution of the cascade in the Binomial model, a straightforward modification is using random numbers rather than the same constant values in each iteration. The Log-normal model is one example along these lines using random numbers from an appropriately scaled Lognormal distribution (Mandelbrot, 1974). One may conveniently define its multipliers, say *M*, via $-\log_2 M \sim N(\lambda, \sigma^2)$. Conservation of average mass in each iteration implies E[*M*] = 0.5, and translates into a restriction on the location and variance of the underlying Normal distribution: $\sigma^2 = 2(\lambda - 1)/\ln(2)$.

The major attraction of this new model is its ability to generate *various degrees* of long memory in different powers of returns – a feature that has been found to characterize virtually all financial prices (e.g., Deng, Engle and Granger, 1993; Lobato and Savin, 1998) but which is absent in traditional volatility models like GARCH and stochastic volatility models as well as in their long-memory variants (e.g., FIGARCH).

The varying degree of long memory of this process is equivalent to nonlinear scaling of the moments under time aggregation. In fact, as pointed out in Calvet and Fisher (2002), multi-fractal measures are characterized by a non-linear scaling function of moments:

(2) $$E[\theta(\Delta t)^q] == c(q) \cdot \Delta t^{\tau_\theta(q)+1},$$

with c() a "prefactor" depending on the depth of the cascade and $\tau_\theta(q)$ the nonlinear scaling function of moments depending on the particular variant of a multi-fractal process used as the time transformation $\theta(t)$. As shown by Mandelbrot, Calvet and Fisher (1997), the scaling behavior of the multi-fractal time transformation carries over to returns from the compound process (1) which would obey a scaling function $\tau_r(q) = \tau_\theta(q\,H)$. For many popular multi-fractals, their $\tau_\theta(q)$ function can be solved analytically and estimation of the multi-fractal parameters is often done by matching the empirical $\tau(q)$ function to its theoretical counterpart. The traditional approach in the physics literature consists in extracting $\tau_\theta(q)$ or $\tau_r(q)$ from a chain of log-log fits of the behavior of various moments q for a certain selection of time aggregates Δt. One, therefore, uses linear fits to the temporal scaling of powers q:

$$\text{(3)}\quad \ln E\left[\left|r_{t,\Delta t}\right|^{q}\right] = a_0 + a_1 \cdot \ln(\Delta t)$$

and fits the empirical $\tau_r(q)$ curve (for a selection of discrete q) to the hypothesized analytical one. Alternatively and perhaps even more widespread in the literature, one would use the so-called Legendre transformation of the scaling function

$$\text{(4)}\quad f_\theta(\alpha) = \underset{q}{\arg\min}[q\alpha - \tau_\theta(q)].$$

and estimates parameters of the underlying multi-fractal model by matching the empirical and hypothetical spectrum of so-called Hölder exponents $f_\theta(\alpha)$ which for most common multi-fractals can also be derived analytically. Again, the shape of the spectrum carries over from the time transformation to returns in the compound process via a simple relationship: $f_r(\alpha) = f_\theta(\alpha/H)$.

Two particular cases discussed in the above papers are the Binomial and Lognormal multi-fractals whose multipliers are chosen from either a Binomial distribution with probabilities m_o and 1- m_o, or from a Lognormal distribution with scale parameter λ (because of the restriction for the variance σ^2 outlined above, the Lognormal model effectively also has only one parameter). For these simple cases, the pertinent $\tau(q)$ and $f(\alpha)$ functions are obtained as follows:

$$\text{(5)}\quad \begin{aligned} &\tau_B(q) = -\log_2(m_1^q + m_0^q), \\ &f_B(\alpha) = -\frac{\alpha_{max} - \alpha}{\alpha_{max} - \alpha_{min}}\log_2\left(\frac{\alpha_{max} - \alpha}{\alpha_{max} - \alpha_{min}}\right) - \frac{\alpha - \alpha_{min}}{\alpha_{max} - \alpha_{min}}\log_2\left(\frac{\alpha - \alpha_{min}}{\alpha_{max} - \alpha_{min}}\right), \end{aligned}$$

with $\alpha_{min} = -\log_2(m_0)$ and $\alpha_{max} = -\log_2(1 - m_0)$ for the Binomial model and:

$$\text{(6)}\quad \tau_L(q) = q\lambda - q^2(\lambda - 1) - 1,\; f_L(\alpha) = 1 - \frac{(\alpha - \lambda)^2}{4(\lambda - 1)}$$

for the Lognormal model.

The estimation of particular compound multi-fractal models for financial returns via (5) or (6) is a new contribution of Mandelbrot, Calvet and Fisher (1997), and Calvet and Fisher (2002), but interpretation of a non-linear shape of the empirical $\tau(q)$ and $f(\alpha)$ functions alone as evidence for multi-fractality is much more widespread. Papers in this vein include, for example, Ghasghaie, S. *et al*., (1996), Vandewalle and Ausloos (1998a, b), Schmitt, Schertzer and Lovejoy (1999), Pasquini and Serva (1999), Breymann *et al*. (2000) and Muzy *et al.* (2001). Our criticism below, therefore, also extends to these more general, data-analytical approaches for detecting multi-scaling as a time series property of financial data.

2. The Significance of the ‘Scaling Estimator: A Test with Surrogate Data

One interesting (but neglected) aspect of these estimation procedures is that they do not only provide a point estimate, but that they could also be interpreted as a *test* for multi-scaling (non-linear scaling of moments) of the pertinent data set against the alternative hypothesis of uni-scaling (linear behavior). Namely, both the Binomial and Lognormal case have a *uni-fractal* limiting case. In the former case this limit is obtained with $m_o = 0.5$, in the later case it is given by $\lambda = 1$. For the Binomial model, it is easy to see that with a split of mass with probabilities 0.5 no differentiation of the measure (and, hence, no transformation of time) is obtained so that eq. 1 simply boils down to fractional Brownian motion. In the Lognormal case, the same occurs via the dependence of the variance on λ: in the case of $\lambda = 1$, the variance vanishes and we also end up with a flat measure which transforms time only in a trivial way. In terms of the scaling function and Hölder spectrum, we would end up with *linear* scaling $\tau(q) = q\,H - 1$, the well-kown uni-fractal relationship characterizing fractional Brownian motion (and scaling $\tau(q) = q/2 - 1$ for the case of Wiener Brownian motion) which translates into a degenerate Hölder spectrum with $f(\alpha) = 1$ if $\alpha = H$ and 0 elsewhere.

Typically, simulations of uni-fractal processes would neither give rise to *perfectly linear* $\tau(q)$ functions[1] nor *perfectly degenerate* $f(\alpha)$ spectra. Some slight, *spurious* curvature would have to be expected in both of these functions. Although in extant literature, typically *any* estimate of the multi-fractal parameters (or, in the more data-analytical approaches: *any* degree of nonlinearity of the scaling functions) has been accepted as evidence of multi-scaling, one would rather prefer to have a criterion for estimates that are *sufficiently far away* from the borderline cases $m_o = 0.5$ and $\lambda = 1$ to accept the null hypothesis of ‘true’ (as opposed to spurious) multi-fractality.

As no theoretical results are available on the distribution of estimates from the scaling estimators derived from $\tau(q)$ and $f(\alpha)$ functions, we pursue a bootstrap or surrogate data approach in dealing with this issue. In order to do so, we may even allow for a generalization of the model (1) proposed by Mandelbrot *et al.* in that we do not have to postulate any particular distribution of the innovations. Our null hypothesis is absence of multi-fractality, i.e. $m = 0.5$ or $\lambda = 1$, while at the same time we allow ourselves to be totally agnostic about the incremental distribution. Our null hypothesis, thus, implies absence of any temporal structure in the data brought about by the multi-fractal time transformation. To arrive at an assessment of the distribution of the m_o and λ estimates, we, then simply have to repeat our estimation procedure sufficiently often for randomized data and compare the original estimate with the ensemble of estimates from the shuffled series. At significance level x, we would reject the null hypothesis if our original estimate would be further away from $m_o = 0.5$ ($\lambda = 1$) than x percent of the estimates from the synthetic data. Note that this is a one-sided test design as our null-hypothesis holds that the randomized data are characterized by absence of multi-scaling.

In fact, the similarity between the original scaling functions and scaling functions obtained for shuffled data has already been noted by Matia *et al*. (2003). Their visual comparison might serve to raise doubts about the reliability of any inference on ‘true’ underlying multi-scaling to be drawn from such analyses. Nevertheless, they did not question the usual conclusion

[1] Berthelsen *et al*. (1994), Bouchaud *et al*.(1999) and LeBaron (2001) have already pointed out the possibility of obtaining spurious multi-scaling for particular stochastic processes.

drawn from the non-linearity of the $\tau(q)$ and $f(\alpha)$ functions, but conjectured that some part of it may be attributed to a broad probability distribution. What we attempt in the following is a somewhat sharper conclusion in the form of a statistical test for the presence of multi-scaling.

Below we report the results from applying this test design to various financial data sets. Our data are: two stock market indices (the New York Stock Exchange Composite Index and the German DAX), an exchange rate (Deutsche Mark/U.S$), and the daily price of gold from the London Precious Metal Exchange. All time series extend over relatively long time horizons. Time intervals and number of observations are: for the DAX the period 10/59 - 12/98 (n = 9818), for the NYCI: 01/66 - 12/98 (n = 8308), for US$-DM: 01/74 - 12/98 (n = 6140), and for the Gold price: 01/78 - 12/98 (n = 5140). The time series have been demeaned and autocorrelation at the first lag has been filtered out before estimating the multi-fractal parameters.

Tables 1 and 2 about here

Tables 1 and 2 show empirical estimates together with minima and maxima of estimates from 1,000 reshuffled series, and the rank the empirical estimates would occupy within the reshuffled ones.[2] These experiments have been carried out for both the Lognormal and Binomial model with both the $\tau(q)$ and $f(\alpha)$ functions used for estimation of the multi-fractal parameters (cf. Tables 1 and 2, respectively). All results are clearly insignificant at any traditional level of significance, and we, therefore, cannot reject the null hypothesis of absence of multi-scaling (or a multi-fractal time transformation) in any of the four time series under consideration. In none of these eight cases do we find an estimate for the original series that is particularly far in the right tail of the bootstrapped estimates (say, beyond rank 900 of 1,000). Hence, the estimated multi-scaling might well be generated as a spurious result from the particular incremental distribution found in the data. One may note that often the entire distribution of estimates from the shuffled data is placed considerably far from its expected value, e.g. in the interval [0.77,1.00] for the $\tau(q)$ estimates of the Binomial model when applied to NYCI data. Figs. 1 and 2 provide illustrations of original $\tau(q)$ and $f(\alpha)$ functions together with a random selection of those from reshuffled data which shows that there are, in fact, no apparent differences between both the original data and typical randomized data sets.[3]

Figs 1 and 2 about here

---

[2] It is interesting to compare the estimates for λ in the case of the U.S.$-DEM with the one reported by Calvet and Fisher (2002): while our estimate from the $\tau(q)$ function is identical with theirs (which is, however, obtained from a fit of the $f(\alpha)$ function), the one we get for the $f(\alpha)$ function is much lower. However, note that data sources and time periods are different which could, in principle, explain this discrepancy. Furthermore, the scaling estimator is characterized by a large standard error which sometimes generates large discrepancies between estimates obtained for slightly different set-ups (cf., Lux, 2003).

[3] Besides our general caveat against taking curvature of $f(\alpha)$ and $\tau(q)$ plots as convincing evidence of multi-scaling, it might be interesting to point out some perplexing regularities in Tables 1 and 2: First, the ranking of the four markets in terms of deviations from the benchmarks $m_o = 0.5$ and $\lambda = 1$ is the same in all four cases. Since higher m_o (or λ) means more pronounced bursts, this might be viewed as a ranking in terms of volatility persistence. Second, along with the inter-market ranking, the ranking of the original estimates within their shuffled counterparts is also extremely consistent within markets, across the four different sets of experiments.

3. Conclusions

A surrogate data test like the one pursued above can usually be interpreted as testing whether certain features of the data can be explained by a rather simple or uninteresting model. This is exactly what we have found: the apparent nonlinear scaling from the very popular $\tau(q)$ and $f(\alpha)$ approaches can be accounted for by spurious multi-scaling typically obtained with financial data after randomization of their temporal structure. Note that inability of rejection of the null hypothesis (absence of multi-scaling) might either occur because the null hypothesis is indeed true or because the discriminating statistics fails to have sufficient power against the alternatives present in the data. The overall conclusion is, therefore, not necessarily that we have demonstrated the absence of multi-fractality altogether, but could also be that the power of the traditional approaches for detecting this behavior is small at least when confronted with typical financial data. In fact, taking into account the overwhelming evidence in favor of different degrees of long-term dependence in various powers of returns (Deng, Engle and Granger, 1993; Lobato and Savin, 1998), we are rather inclined to follow the latter interpretation of our results. It is also worth pointing out that our results are in harmony with findings of spurious curvature of the $\tau(q)$ and $f(\alpha)$ functions in experiments with some fat-tailed distributions (Nakao, 2000). They also square well with the finding that the scaling estimates suffer from large biases and standard errors (Lux, 2003; Yamasaki, 2003).

References:

Berthelsen, C.L., J.A. Glazier and S. Raghavachari, 1994, Effective Multifractal Spectrum of a Random Walk, *Physical Review E* 49, 1860 - 1864

Bouchaud, J.-P., M. Potters and M. Meyer, 2000, Apparent Multifractality in Financial Time Series. *European Physical Journal B* 13, 595 – 599

Breymann, W, S. Ghashgaie and P. Talckner (2000) A Stochastic Cascade Model for FX Dynamics, *International Journal of Theoretical and Applied Finance*, **3**, 357 – 360.

Calvet, L. and A. Fisher (2002a) Multi-Fractality in Asset Returns: Theory and Evidence, *Review of Economics & Statistics* 84, 381 – 406.

Ghasghaie, S. *et al*. (1996) Turbulent Cascades in Foreign Exchange Markets, *Nature* **381**, 767 - 770.

Lobato. I.N. and N.E. Savin (1998) Real and Spurious Long-Memory Properties of Stock Market Data, *Journal of Business and Economics Statistics*, **16**, 261 - 283.

LeBaron, B. (2001) Stochastic Volatility as a Simple Generator of Apparent Financial Power Laws and Long Memory, *Quantitative Finance* 1, 621 – 631.

Lux, T. (2003) *The Multi-Fractal Model of Asset Returns: Its Estimation via GMM and Its Use for Volatility Forecasting*. Manuscript, University of Kiel.

Mandelbrot, B., 1974, Intermittent Turbulence in Self Similar Cascades: Divergence of High Moments and Dimension of the Carrier, *Journal of Fluid Mechanics* 62, 331 – 358

Mandelbrot, B. (1999) A Multifractal Walk down Wall Street, *Scientific American*, February, 50 - 53.

Mandelbrot, B., A. Fisher and L. Calvet (1997) *A Multifractal Model of Asset Returns*. mimeo: Cowles Foundation for Research in Economics.

Matia, K., Y. Ashkenazy and H.E. Stanley (2003) Multifractal Properties of Price Fluctuations of Stock and Commodity Markets, *Europhysics Letters* 61, 422 – 428.

Muzy, J.-F., D. Sornette, J. Delour and A. Arneodo (2001) Multifractal Returns and Hierarchical Portfolio Theory, *Quantitative Finance*, **1**, 131 – 148.

Nakao, H. (2000) Multi-Scaling Properties of Truncated Lévy Flight, *Physical Review Letters* A 266, 282 – 289.

Pasquini, M. and M. Serva, 1999, Multiscale Behavior of Volatility Autocorrelations in a Financial Market, *Economics Letters* 65, 275 – 279.

Schmitt, F., D. Schertzer and S. Lovejoy (1999), Multifractal Analysis of Foreign Exchange Data, *Applied Stochastic Models and Data Analysis*, **15**, 29 – 53.

Vandewalle, N. and M. Ausloos (1998a) Sparseness and Roughness of Foreign Exchange Rates, *International Journal of Modern Physics C*, **9**, 711 - 719.

Vandewalle, N. and M. Ausloos (1998b) Multi-affine Analysis of Typical Currency Exchange Rates, *European Physical Journal B*, **4**, 257 - 261.

Yamasaki, K. (2003) *The Reliability of Mutifractal Indicators*, Manuscript, Tokyo University of Information Sciences.

Table 1: Reshuffle Test for Significance of Log-Normal and Binomial Multi-Fractal Estimates from Scaling Estimator – τ(q) fit

Data	Empirical data	Reshuffled data (1,000 runs)		
	$\hat{\lambda}$ from $\tau(q)$	Min $\hat{\lambda}$	Max $\hat{\lambda}$	Rank of empirical $\hat{\lambda}$
NYCI	1.226	1.086	1.300	705
DAX	1.180	1.150	1.248	844
US$-DEM	1.109	1.050	1.233	166
Gold	1.140	1.059	1.260	322
	$\hat{m}_0$ from $f(\alpha)$	Min $\hat{m}_0$	Max $\hat{m}_0$	Rank of empirical $\hat{m}_0$
NYCI	0.961	0.774	1.000	731
DAX	0.893	0.659	0.977	848
US$-DEM	0.762	0.667	0.977	159
Gold	0.818	0.657	1.000	308

Note: The scaling estimator is implemented in the following way: 25 time increments Δt ranging from $\Delta t = 5$ to $\Delta t = T/5$ (T the length of the time series) have been used which are equally spaced in logs (i.e. the next Δt is computed as $\Delta t' = \exp(\ln(\Delta t) + \ln(T/5)/25)$, only positive moments are used, q = 0.1, 0.2…(0.1)…3, 3.5,. (0.5),… 10, and the estimates of λ and m_0 are found by minimizing the squared deviation between the theoretical and empirical scaling function at the above q values.

The time intervals and number of observations are:
DAX: 10/59 - 12/98 (n = 9818),
NYCI: 01/66 - 12/98 (n = 8308),
US$-DM: 01/74 - 12/98 (n = 6140),
Gold price: 01/78 - 12/98 (n = 5140).

Table 2: Reshuffle Test for Significance of Log-Normal and Binomial Multi-Fractal Estimates from Scaling Estimator – f(α) fit

Data	Empirical data	Reshuffled data (1,000 runs)		
	$\hat{\lambda}$ from f(α)	Min $\hat{\lambda}$	Max $\hat{\lambda}$	Rank of empirical $\hat{\lambda}$
NYCI	1.240	1.033	1.549	751
DAX	1.140	1.011	2.162	816
US$-DEM	1.031	1.007	2.644	130
Gold	1.076	1.009	2.582	401
	$\hat{m}_0$ from f(α)	Min $\hat{m}_0$	Max $\hat{m}_0$	Rank of empirical $\hat{m}_0$
NYCI	0.811	0.607	0.870	738
DAX	0.741	0.554	0.990	844
US$-DEM	0.616	0.501	0.992	153
Gold	0.675	0.554	0.992	341

Note: Same time intervals, number of observations, and implementation of the scaling estimator as in Table 1, but here the estimates of λ and m_0 are found by minimizing the squared deviation between the theoretical and empirical spectrum at the α coordinates of the empirical spectrum.

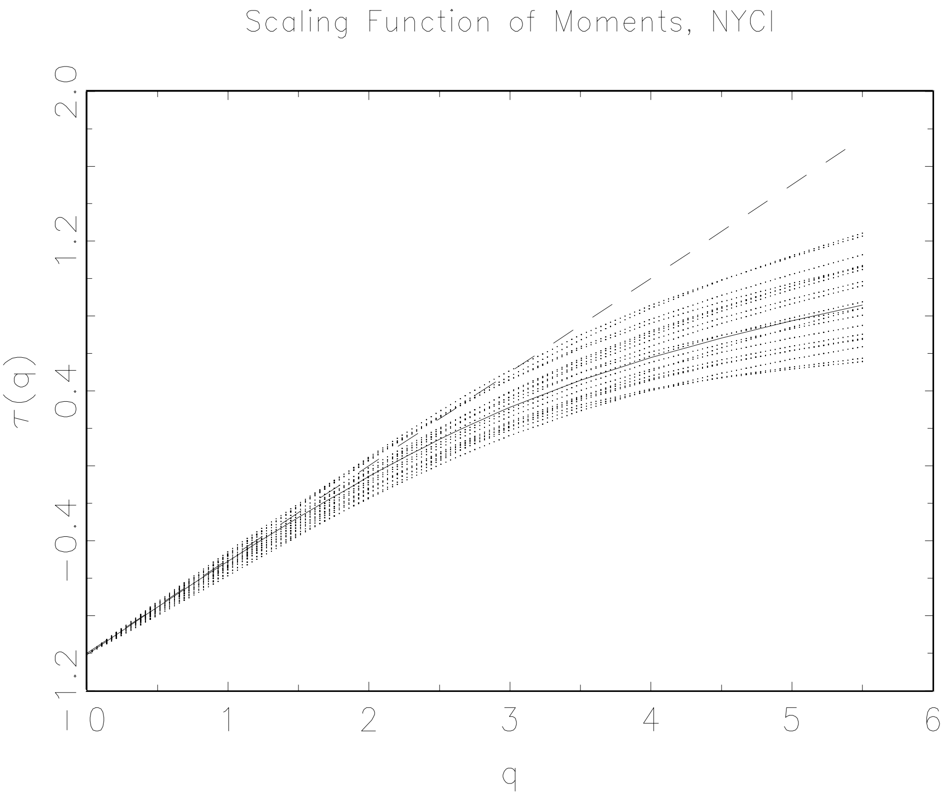

Fig 1: τ(q) function for New York Stock Exchange Composite Index returns (solid line) and twenty reshuffled series (dotted lines). The broken line gives the scaling behavior τ(q) = q/2 – 1 of Gaussian random variables.

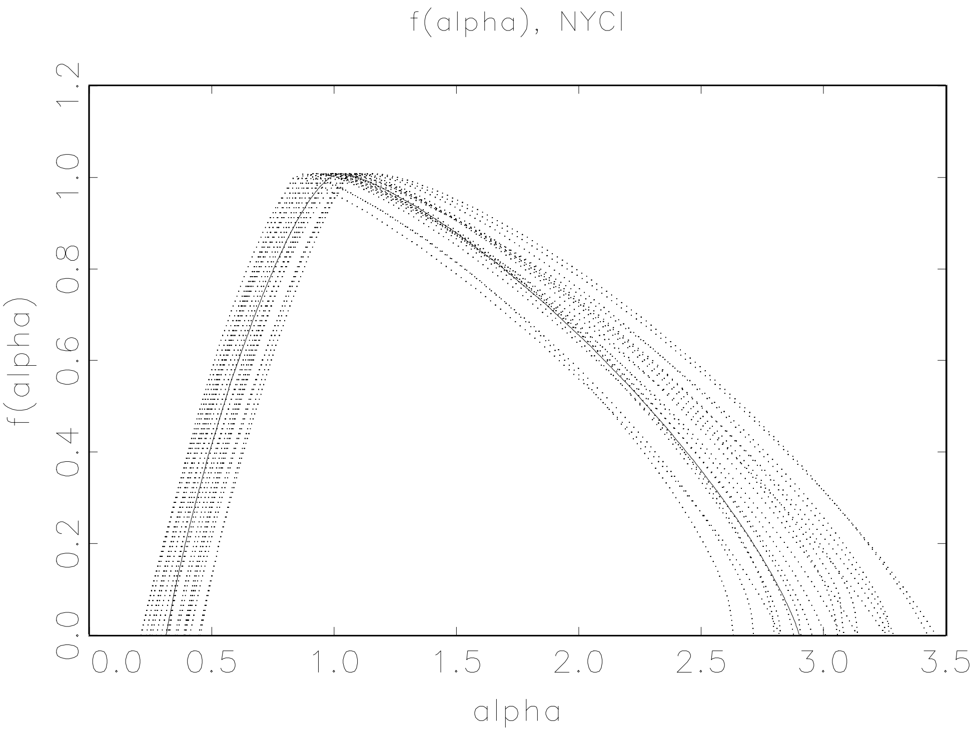

Fig. 2: f(α) spectrum for New York Stock Exchange Composite Index returns (solid line) and twenty reshuffled series (dotted lines).