The Fractal Behaviour of CAC 40 Returns Examined in the Time-Frequency Domain

Serge Hayward

Department of Finance
Ecole Supérieure de Commerce de Dijon, France
shayward@escdijon.com

Abstract

This paper considers a scaling law, a rule relating price movements and time intervals, in the spectral analysis settings. The fractal behaviour of returns, occurrence of similar patterns at different scales, is examined in a diffuse two-dimensional time-frequency domain, rather than in a one-dimensional time domain. The novelty of the approach is to investigate the probability distribution of the data-generating mechanism over frequencies localised in time, as an alternative to sampling intervals.

Keywords: Financial Data Mining; Wavelets.

1. Introduction

Considering that financial decisions occur at different scales, analysis of separate levels of a complex signal provides a valuable source of information. Events of different scales separated in time are interpreted as indicators when their causality is established. In addition, frequency decomposition permits to separate short from long time dynamics, as well as to discover their consequence on the local and global structures. Wavelet transform decomposition of a complex time series into separate scales is a focus of this study. An evolutionary / artificial neural network (E/ANN) is used to learn the information at separate scales and combine it into meaningfully weighted structures.

2. Data Generating Mechanism

Unlike fully revealing equilibrium of homogeneous beliefs, in the environment with heterogeneous beliefs prices are driven by prevailing expectations of market participants. Thus forecasting future prices one must form expectations of others forecasts. Evolution of agents' expectations to a great extent governs the adaptive nature of market prices. Overlapping beliefs of heterogeneous agents largely prevent the effective examination of expectation formation and price forecasting by traditional (time-series) methods.

Traders with similar strategies take comparable actions, forming a particular layer of multi-structured DGM. Thus the search is for an approach capable learning the underlying structures with different time-scales. Frequency decomposition of a time series is used for such structures identification. An inverse operation, synthesizing appropriate structures generates better characterised DGM with further applications for modelling and forecasting.

3. Fractal Behaviour and Long Memory

Let long memory be conventionally defined by autocorrelation function decaying at hyperbolic rate, $f_h(\lambda) = me^{-\lambda}$, rather than at the exponential rate, $f_e(\lambda) = me^{-\lambda/h}$. *m, h* and λ are parameters with λ determining the lag order of the autocorrelation function. Unlike the exponential function, the hyperbolic function simultaneously captures the long and short term persistence in the series.

Consider the fractal behaviour as occurrence of similar patterns at different scales. [1] assert that the distributions of returns must be identical over different time scales, thus satisfying self-affinity (self-similarity) condition:

$$\forall c, k, t_1, ..., t_k \geq 0; \quad \exists S > 0; \qquad (1)$$
$$\{X(ct_1), ..., X(ct_k)\} \overset{d}{=} \{c^S X(t_1), ..., c^S X(t_k)\},$$

where S is self-affinity index or scaling exponent of the process {X(t)}. Among popular self-affine processes in finance are Fractional/Brownian motions and L-stable processes. At the same time empirical evidence demonstrates that financial time series display thinner tails and become less peaked in the bells when the frequency increases, thus rising doubts about their self-affinity.

Unlike the fractal property investigation over different *sampling intervals* and development of multifractal models build upon empirical scaling laws

[2], this research examine the fractal behaviour in the frequency-domain[1] (FD), to identify common empirical regularities over *decomposed frequencies*. The analysis of a signal in FD implies the estimation of some transform coefficients, obtained through this signal decomposition. Reconstruction of the signal corresponds to the computing the linear combination equations, involving those coefficients.

In this respect the long memory hypothesis is tested against structural break hypothesis. Occasional structural breaks might spuriously induce long memory effect with confusing impact on the overall behaviour [3-5]. The effects of genuine long memory must be present across different frequencies in a similar way as over sampling intervals[2]. Examining the implications of long memory in FD, the spectrum is expected to be approximately log-linear and close to the origin with a negative slope at the lowest frequency under the long memory hypothesis. Testing the structural break hypothesis the intention is also to differentiate structural breaks from phase shifts (given as movements of various frequencies into and out of phase with each other).

4. Wavelet Transforms

Wavelet analysis, representing a signal in the *time-scale*[3] domain provides good time and frequency resolutions. A signal is multiplied with the *wavelet* function[4] and the transform is computed for different segments of TD signal. By applying modified versions of a prototype, the *mother wavelet* to the time series, wavelet transform convolves the data with a series of local waveforms to discover correlated features or patterns.

Being robust against shifting in time wavelets reveal well reappearing features. Defined over a finite length, they are capable effectively to represent complex nonstationary signals. Localising fluctuations power in a time series, a decomposed time-frequency space is used to determine the dominant modes of oscillations and their evolution in time.

5. Significance Testing

Peak-based critical limit significance is used for signal analysis/processing. Two backgrounds are considered: a white noise, H_0: AR(1) = 0 and a red noise (the signal power decreases with increased frequency), H_0: AR(1) > 0. A 95% (99.9%) peak-based critical limit implies that in 1 out of 20 (1000) random noise signals would the largest peak reach this height by a random chance. The Monte Carlo simulation is used to generate the peak-based critical limits. The simulated data is then fitted to bivariate, univariate or trivariate polynomials, depending on the number of factors affecting the significance.

6. Experimental Results

6.1. Data Analysis

The data considered includes prices for CAC 40® share index measuring the evolution of a sample of 40 equities listed on Euronext® regulated markets in Paris. During the trading session 9.00 – 17.30, index levels are calculated in real time using the last trade quoted and are disseminated every 30 seconds. Thus the change in the index is equal to the sum of the change in each of component equity times its weight in the index, where the market price of equity is adjusted for corporate actions taking effect, but not for dividend paid.

The period under investigation runs from 01.03.90 through 07.03.05 with the business time scale, which excludes holidays, Saturdays and Sundays from the physical time scale used in the experiment. The length of the data series is driven by the objective to explain the present behaviour of the index, where the data prior to 1990 is considered to refer to a different from the current phenomena. The original data, consisting of the series with tick frequency of 30 seconds was obtained from Euronext.

After subsampling, the series containing 3779 eight and a half hours prices was extracted. The analysis of such frequency well relates to the objective of examining the behaviour of heterogeneous agents that strategically fulfil certain goals and cluster according to some (nontrivial) time horizons, optimal for their economic type. [7] assert a similar sampling frequency as the most appropriate in designing an ANN by a typical off-floor trader. Furthermore, for practical purposes of construing a model, generating reliable predictions, the issues of the realistic time needed to execute a strategy makes the choice of such a sampling interval justified.

Examining the prices graphically reveals that the price index exhibit an upward, but non-linear trend until around the year 2000, a downward trend until spring 2003 and a mild upward trend afterwards. A possible complete cycle can be visually acknowledged between October 1998 and March 2003. Persistent

[1] The alternative time-domain (TD) originates from the classical theory of correlation.

[2] A trend can also induce spurious long memory [6].

[3] Consider the scale is an inverse of the frequency.

[4] A wavelet refers to a complex function (with a zero integral over the real line) of a small wave, i.e. compactly supported (finite length) oscillatory function.

fluctuations around the trends, which increase in variability with downward moves, can also be seen.

Considering changes in prices, *x(t)* is conventionally expressed in logarithmic terms. Over a fixed time interval, Δt the return, *r(t)* is defined as

$$r(t) = r(\Delta t; t) = x(t) - x(t - \Delta t), \quad (2)$$

where *x(t)* is a price series. It is generally viewed that the return process is closer to stationarity and its distribution is more stable over time than price series. As it can bee seen from Table 1 the distributions of return measures are asymmetric and commonly display a right tail. The skewness coefficients are significant in the modest range, [2.4; 5.8]. Values of kurtosis, significantly greater than zero indicate that the return measures have longer tails than those for a normal distribution. The kurtosis coefficients are significant for all three variables in a sizable range, [32.9; 49.1]. The Kolmogorov-Smirnov tests and their P-values indicate the likely significant variation from a normal distribution for all return measures. This conclusion is supported by the Shapiro-Wilk normality test (SW). The P-values given by this test, P_{SW} indicating how good the fit is, reject normality hypothesis for all three returns.

Test/Stat	R	|R|	R^2
SK	0.191	0.231	0.096
SE_{SK}	0.040	0.040	0.040
KU	2.662	3.926	2.631
SE_{KU}	0.080	0.080	0.080
K-S/Pr	0,043/< 0,001	0,138/< 0,001	0,321/< 0,001
SW	0.972	0.933	0.973
P_{SW}	0.000	0.000	0.000
H-σ_H- r²	0.57-0.002-0.99	0.79-0.002-0.99	0.78-0.002-0.99

Table 1**.** Return Measures

Return series also displays statistically significant positive correlations with volatility. This fact indicates that heterogeneous traders have different beliefs about price dynamics, executing their transactions in different situations. Increased market activity results in increased volatility with prices not readily converging to their rational expectation values.

6.2. Long Memory and Fractal Behaviour

Descriptive statistics for the price series identify strong positive autocorrelation up to very high lags, indicating a linear dependence between the current and past values. Using the Growth in Cumulative Range algorithm, the Hurst exponent, H is estimated to measure the fractal dimension of the data series. H_P = 1.014, being significantly greater than 0.5, indicates the long-term memory effect in the price series[5]. Long-memory hypothesis was also confirmed by spectrum analysis. The spectrum was estimated to be approximately log-linear, close to the origin with a negative slope. The statistical analysis confirms the Lagged Adjustment Model hypothesis [8]. As stocks, constituting an index react with different speed to the aggregate information and the autocovarience of a diversified portfolio is the average of cross-covariance of its constituents, positive autocorrelation in the index is present despite its absence in stocks.

Considering the memory, the autocorrelation function of the return does not display hyperbolic decay; it quickly dies out and stays within the confidence intervals at higher lags. Though significant autocorrelation for small lags implies that the volatility clusters or patters might be present. The autocorrelations of absolute and squared returns, on the other hand, are significant even at high lags, as it is confirmed by the Hurst exponents, given in the table above, and can also be seen on the Figure 1. Autocorrelations of square returns display a few peaks[6]. In magnified scale one can detect that those peaks are fairly regular, occurring every 500 days, around 02.94; 02.96; 02.98; 02.00 and 02.02.

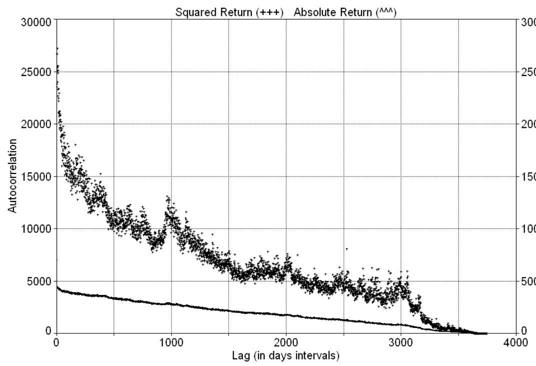

Fig. 1: Autocorrelation Functions of Returns

Profoundly significant autocorrelation of squared returns displays the initial rapid decay followed by a slow dissipation. This behaviour is commonly attributed to slowly mean-reverting fractionally integrated processes. Note that all the returns considered follow fat-tailed and skewed distributions, while the distribution of pure fractional noise process is expected to be normal.

[5] The coefficient standard error derived from the fit, σ_H = 0.001. The goodness of fit index, based on a sum of squares criterion, r^2 = 0.999.

[6] Note that theoretical autocorrelation of squared returns is meaningful only if the kurtosis of the returns is finite.

The analysis performed supports the long memory hypothesis for the underlying data-generating process of returns with the results similar to the studies of the US stock indexes [9]. Although, the long memory behaviour was not found to be the most apparent for the absolute returns, as claimed previously [9, 10]. Spectrum analysis was able to identify the long memory with a stable slope across data frequencies of up to 8 hours, similar to the findings of [11].

The long memory was established to be an intrinsic feature of the system, present in different scales. Since the conventional statistical inference appears to be limited in the case of long memory signals, the wavelet analysis that approximately decorelate the signal with long memory is particular appropriate. Furthermore, it provides a platform for modelling nonstationary features without the exact knowledge of the correlation structure.

A hypothesis that different scaling regimes characterize intradaily and lower frequency signals [12] was tested. The difference in dynamics of two scales is confirmed by the current experiment. It is primarily found in significantly different levels of volatility in high frequency signal. Such large volatility is well explained by the behaviour of intraday speculators.

Multiscale analysis establishes the distributional nonlinearities and the presence of the fractal property in data frequencies from 2 hours up to a few months. The distributions of returns display a progressive convergence to the Gaussian process beginning at about monthly frequency. The spreading of the unconditional distribution of returns was detected: increasing the time scale results in thinner tails. It would not be appropriate to explain the scaling law outcomes as a consequence of a (stable) random process. Since return distributions are unstable the results presented are attributed to the phenomenological law.

7. Conclusion

Testing the long memory versus structural brake hypotheses, identifies a number of sample periods when structural breaks spuriously induce long memory effect on a particular frequency, without their presence across all frequencies. Similarly, testing the structural brakes versus phase shifts hypotheses detect times when the low and high frequencies move into and out of phase with each other, resulting in phase shifts rather than structural breaks, claimed by other studies. Distinguishing long memory, structural breaks and phase shifts enhance the understanding of the series' unstationary behaviour.

Multiscale analysis establishes the presence of the fractal property in data frequencies from 2 hours up to a few months. A progressive convergence to the Gaussian process begins at about monthly frequency. The results presented are attributed to the phenomenological law.

8. References

[1] Mandelbrot, B.B., *The Variation of Certain Speculative Prices.* Journal of Business, 1963. 36(4): p. 394-419.

[2] Mandelbrot, B.B., I.J. Fisher, and L.E. Calvet, *A Multifractal Model of Asset Returns*. 1997, Cowles Foundation Discussion Paper.

[3] Diebold, F., *Modeling the Persistance of the Conditional Variances: A Comment.* Econometric Reviews, 1986. 5: p. 51-56.

[4] Lamoureux, C.L., W., *Persistance in Variance, Structural Change, and the GARCH Model.* Journal of Business and Economic Statistics, 1990. 8: p. 225-234.

[5] Granger, C.W.J. and N. Hyung, *Occasional Structural Breaks and Long Memory.* University of California San Diego Discussion Paper, 1999. 99-14.

[6] Bhattacharya, R.N., V.K. Gupra, and E. Waymire, *The Hurst Effect under Trends.* Journal of Applied Probability, 1983. 20: p. 649-662.

[7] Kaastra, I. and M. Boyd, *Designing a Neural Network for Forecasting Financial and Economic Time Series.* Neurocomputing, 1996. 10: p. 215-236.

[8] Ahn, D.-H., et al., *Partial Adjustment or Stale Prices? Implications from Stock Index and Futures Return Autocorrelations*. 2000, Stern Business School Working Paper.

[9] Ding, Z., C.W.J. Granger, and R.F. Engle, *A Long Memory Property of Stock Market Returns and a New Model.* Journal of Empirical Finance, 1993. 1: p. 83-106.

[10] Lobato, L.N. and N.E. Savin, *Real and Spurious Long-Memory Properties of Stock-Market Data.* Journal of Business and Economic Statistics, 1998. 16: p. 261-267.

[11] Andersen, T.G. and T. Bollerslev, *Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns.* Journal of Finance, 1977. 52: p. 975-1005.

[12] Gencay, R., F. Selçuk, and B. Whitcher, *Scaling Properties of Foreign Exchange Volatility.* Physica A, 2001. 289: p. 89-106.