

Heterogeneity of Price Discovery Processes in Financial Markets

Serge Hayward

Law and Finance Department, Burgundy School of Business - Groupe ESC Dijon Bourgogne, BP 50608 – 21006
Dijon cedex – France – shayward@escdijon.com

Received 12th of June 2007, accepted November 2007.

Abstract

This paper considers heterogeneity-driven asymmetry in the stock market with the low frequency shocks penetrating the entire market, whereas the high frequency shocks are short-lived and often have no impact outside of their boundaries. Testing the long memory versus structural break hypotheses identifies a number of sample periods when structural breaks spuriously induce the long memory effect on a particular frequency, without their presence across all frequencies. Similarly, testing the structural breaks versus phase shifts hypotheses detects periods 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 enhances the understanding of the series' emergent nonstationary behaviour.

A heterogeneous beliefs model with expectations differentiated according to their time dimension is developed. Decomposing a time series into a combination of the underlying series, representing beliefs of major clusters of market participants, a fractal structure of the agents' time preferences is identified. Wavelet transforms are used for analysis of local behaviour of heterogeneous agents.

Key words: max 10 key words

Traders' Heterogeneity

Economic agents' behaviour is determined by their intrinsic opportunity sets. One of the main factors affecting traders' opportunity sets is the time dimension that differentiates market participants according to their time horizons, into short and long term traders. Agents with distinct time scale dimensions respond in a different way, with different reaction time to the news. Low frequency shocks are likely to affect all market participants, though with some time delay. On the other hand high frequency shocks may be ignored, at least for a while, by long term traders. Such information processing 'inefficiency' is heterogeneity-driven asymmetry according to the market efficiency hypothesis. Identified casual relationships are potentially exploitable in trading model development.

Consider a scaling law that relates traders time horizons and frequency of price fluctuations. In these settings rapid price movements that corresponds to high frequency fluctuations are reflected in the frequent trading positions' revaluations by agents with short term horizons, such as (inter-day) speculators. By trading upon the information of high frequency signals short

term traders frequently execute transactions, supplying high frequency information to the market.

On the other hand slow price movements, corresponding normally to larger price shifts, are more apparent in low frequency signals, with less noise interference. Thus traders with long term horizons (e.g. certain institutional investors) tend to trade upon the information of low frequency signals. Long term traders reduce their risk exposure with derivatives, stop-loss limits and the like. They act, to a certain extent, according to the market fundamentals (Dacorogna, Gencay et al. 2001). Long term agents, trading upon low frequency signals, provide low frequency information to the market that is used adaptively by traders with similar time scale.

The interactions between heterogeneous agents in relationship to each other actions rather than to the market news produce endogenous dynamics in the market¹. Such dynamics provide reasonable explanation to some common empirical facts in Finance, as trend persistence or volatility clustering. Differentiating economic agents' expectations according to their time dimension has valuable

¹ Kurz (1994) calls it 'endogenous effects'.

consequences for forecasting. Since the time scale of traders is the key characteristic of the market, the adaptive dynamics of prices reflect beliefs and behaviour of the dominant agents on the market. Establishing a (trading) frequency signal of the dominant agents can be used for tuning into it and exploiting it in profitable predictions.

Time-Scale Domain

Considering that different financial decisions occur at different scales, analysis of separate scales of a complex signal provides a valuable source of information. Stemming out of Fourier approach, the frequency domain's methods analyse the function of interest with some 'basis' functions, producing nonparametric models that are generally more robust to misspecifications.² At the same time traditional spectral analysis' applications to nonstationary financial time series (without strict periodicity) are limited.

Modern flexible tools of time-scale domain like wavelets find their ways into finance, offering adaptive analysis of local behaviour of heterogeneous financial markets. Wavelet transform decomposes a complex time series into separate signals, allowing us to learn the information at separate scales as well as to combine them into meaningfully weighted structures.

Long Memory and Structural Breaks

Consider a time series, $x(t)$ with an order given by a continuous variable t indexed the data. Let a signal, $x(t)$ belong to the vector space of measurable square-integrable $(\int_{-\infty}^{\infty} x^2(t)dt < \infty)$ one-dimensional functions, $L^2(\tilde{N})$:³

$$x(t) = \sum_{n=1}^N \delta x_n, \tag{1}$$

where $N = t/\tau$ and τ stands for a microscopic time scale; $\delta x_n = |\sigma_n| \mathbf{e}_n$ are the elementary increments with σ_n and \mathbf{e}_n representing an amplitude and a sign. The variance of the elementary increments, σ_n^2 for financial prices is claimed to be random, with a (slowly) decaying time correlation function:

² Recall that the alternative time domain's models originate from the classical theory of correlation.

³ Standard notation, \tilde{N} and \tilde{U} is invoked for sets of real numbers and integers.

$$\langle \sigma_i^2 \sigma_{i+l}^2 \rangle - \langle \sigma_i^2 \rangle^2 = 2C^2 \langle |l| \rangle \approx_{l \rightarrow \infty} \frac{2\gamma^2}{|l|^{2\nu}}, \tag{2}$$

where l represents lags and $\nu < 0.5$ is the exponent, whose actual value depends on a particular financial market. If $\nu = -1$, the volatility is said to follow a random walk. For any stochastic process X with stationary increments $X(a, \mathbf{D}t) \equiv X(a + \mathbf{D}t) - X(t)$, the dependence in the size of the increments is quantified by the autocovariance in levels:

$$Aucov_X(t, q) = Cov(|X(a, \Delta t)|^q, |X(a+t, \Delta t)|^q), \tag{3}$$

provided that $\Delta |X(a, \Delta t)|^{2q} < \infty$.

Let long memory be conventionally defined by autocorrelation function decaying at hyperbolic rate, $f_h(\lambda) = m e^{-\lambda}$, (rather than at the exponential rate, $f_e(\lambda) = m e^{-\lambda/h}$). In the above 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. For a fixed q in (3) the process is characterised by long memory in the size of increments if the autocovariance in levels is hyperbolic in t when $t/\Delta t \in \mathbb{N}$.

Mandelbrot and van Nesss (1968) present the 'fractional noise' process with a hyperbolically decaying autocorrelation function, which for large number of lags converges to

$$a \approx H(2H - 1)l^{2(H-1)}, \tag{4}$$

where H is the Hurst exponent with the range [0.5; 1.0] for persistent fractional noise. Peters (1991) claims that a drift exponent different from 0.5 is an indication of fractional noise.

Consider the fractal behaviour as occurrence of similar patterns at different scales. Mandelbrot (1963) assert that the distributions of returns must be identical over different time scales, thus satisfying self-affinity (self-similarity) condition:

$$\forall c, k, t_1, \dots, t_k \geq 0; \exists S > 0; \{X(ct_1), \dots, X(ct_k)\} \stackrel{d}{=} \{c^S X(t_1), \dots, c^S X(t_k)\}, \tag{5}$$

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, investigated over different *sampling intervals* with a multifractal model build upon empirical scaling laws (Mandelbrot, Fisher et al. 1997), this research examines the fractal behaviour in the Frequency Domain, to identify common empirical regularities over *decomposed frequencies*. In this respect the long memory hypothesis is tested against structural break hypothesis. Occasional structural breaks spuriously induce long memory effect with confusing impact on the overall behaviour (Diebold 1986; Lamoureux 1990; Granger and Hyung 1999). The effects of the genuine long memory must be present across different frequencies in a similar way as over sampling intervals⁴. Examining the implications of long memory in the Frequency Domain, 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 it is also intended to differentiate structural breaks from phase shifts (given as movements of various frequencies into and out of phase with each other). The null hypothesis is that the series under investigation has an invariant statistical structure. The formal rejection of the null hypothesis assures that the signal reflects genuine structural changes.

Signal Transformation

Wavelet analysis, representing a signal in the *time-scale*⁵ domain provides a fine time resolution, being narrow in time, to zoom in high frequencies with less ambiguity regarding their exact position in time. On the other hand a good frequency resolution is achieved by being short in frequency (long in time) with less ambiguity about the exact value of the frequency (to deal with low frequencies). To accomplish this, a signal is multiplied with the *wavelet* function⁶ and the transform is computed for different segments of the Time Domain signal. Well localized in time, without constant resolution and infinite duration, wavelets adapt the time and frequency support,

⁴ A trend can also induce spurious long memory (Bhattacharya and Gupta, 1983).

⁵ Consider scale as an inverse of frequency.

⁶ 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.

reaching the boundary of Heisenberg Uncertainty Principal⁷

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. The mother wavelet is a source function, from which translated and scaled wavelet functions (with different regions of support) are constructed. Nonorthogonal wavelet functions are used with both Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT), whereas orthogonal⁸ wavelet bases imply exclusively Discrete Wavelet Transform.

To apply a wavelet function to a signal, it is *scaled*⁹ by varying the scale, s and *translated*¹⁰ along the localized time index, n . Note that increments of s and n are continuous in Continuous Wavelet Transform. Consider scaling of the wavelet V with the expansion of its range by a multiplicative factor, $V(1/s)$ and translation of V with shifting its range to the right, $V(n'-n)$:

$$V(n, s) = s^{-1/2} V\left(\frac{n'-n}{s}\right), \quad (6)$$

where $s^{-1/2}$ is energy conservation weighting value, a normalization to keep the total energy of the scaled wavelet constant.

By projecting a time series of interest, $x(t)$ onto the wavelet function $V(n,s)$, decomposition occurs when similar frequency content (correlated features) is identified:

$$CWT(n, s) = \int_{-\infty}^{\infty} x(t)V(n, s) * (t)dt. \quad (7)$$

Providing that a wavelet function satisfy the admissibility condition, given by time-frequency space localization and zero mean, an inverse operation can be performed. Reconstruction (synthesizing) of the original signal from its

⁷ Heisenberg Principal imposes limits on precising an event appearance with arbitrary accuracy in both time and period terms.

⁸ Two vectors are orthogonal if their inner product equals zero. Orthogonal wavelet functions characterised by no overlapping or projection.

⁹ Scaling is a mathematical operation that either dilates or compresses a signal. In wavelet analysis dilated (stretched out) signals are obtained with $s > 1$ and relate to larger scales; whereas compressed signals are delivered with $s < 1$ and correspond to small scales.

¹⁰ Translation indicates the location of the mother wavelet. The widow is shifted through the signal.

wavelet transform is done with integrating over all scales and locations:

$$x(t) = \int_0^{\infty} \int_{-\infty}^{\infty} CWT(n, s) V(n, s)(t) dn \frac{ds}{s^2}, \quad (8)$$

Continuous Wavelet Transform of a discrete time series x_n is defined as the convolution of the sequence x_n with scaled and translated versions of a wavelet function $V_{n,s}$ and in direct form (in the Time Domain) is given by

$$CWT_{n,s}^{dir} = \sum_{n'=0}^{N-1} x_{n'} V_{n,s} * \left[\frac{n' - n\delta t}{s} \right], \quad (9)$$

where n is the localized time index, N is the data series' length, s is the wavelet scale, $(*)$ is the complex conjugate and δt is the sampling interval. Thus a simultaneous presentation of the amplitude versus the scale and fluctuation of the amplitude over time is constructed by varying the wavelet scale and translating along the localized time index.

A function of two continuous parameters (scale and location), Continuous Wavelet Transform is characterised by overlapping (nonorthogonality). A high amount of extra information requires extensive computation and memory, though producing particularly accurate time-frequency spectrum. For computer processing, a discretised Continuous Wavelet Transform (semi-discrete or wavelet series) considers a limited amount of scales with a varying quantity of wavelet coefficients at each scale.

Discrete Wavelet Transform is defined by a square matrix of filter coefficients with the information being sparse in the transform space through compression. A wavelet, characterised by a scaling function that spans the space Λ_2^j , is given by a linear combination of functions from the next subspace Λ_2^{j+1} . Since the subspaces are nested and self-similar a relationship between scaling functions of any two neighbouring subspaces defines the filter coefficients. Thus orthogonal wavelets are characterised by a (finite) set of filter coefficients and scaling functions with different subspaces. Discrete Wavelet Transform uses a pair of filters to isolate high and low frequencies.

Common sampling of Discrete Wavelet Transform coefficients from Continuous Wavelet Transform on a dyadic¹¹ grid ($s_o = 2$ and $n_o = 1$) yields $s = 2^j$ and $n = k2^j$. Following (Mallat 1989), dilating by s

¹¹ Dyadic is a two-fold relationship between resolution scales.

$= 2^j, V_{2^j}(t) = 2^j V(2^j t)|_{\forall j \in \mathbb{Z}}$ and translating by $n = k2^j$ define an orthonormal¹² wavelet basis:

$$\forall j, k \in \mathbb{Z}, \quad V_{j,k} = 2^{-j/2} V_{2^j}(n - k2^{-j}). \quad (10)$$

Discrete Wavelet Transform analyses a signal at different frequency bands with different resolutions by successive decomposing it into a (coarse) approximation and detail information. Consider a^j as a coarser approximation at the resolution J of x_n and d^j as the detail information that gets lost while going from the approximation a^j to the coarser approximation a^{j+1} . The scaling and wavelet functions, associated with low and high pass filters, are used for filtering the Time Domain signal. *Filtering* changes the resolution through convolution of the signal with the impulse response of the filter. After passing a (finite length) signal through half bands high/low pass filters the amount of detail information in the signal is halved. Following such filtering according to the Nyquist rule, a half of the samples can be discarded since the signal has a highest frequency of $\mathbf{p}/2$ radians (rather than \mathbf{p}). *Subsampling* the signal by two, halves the number of points and double the scale (reducing the sample rate). If y_k^H and y_k^L denote the output of high, h_n and low pass, l_n filters the decomposition of the signal, x_n is thus

$$\begin{aligned} y_k^H &= \sum_n x_n h(2k - n); \\ y_k^L &= \sum_n x_n l(2k - n). \end{aligned} \quad (11)$$

The length of the signal determines the maximum number of decomposition levels. The full transform demands the data series length to be of a power 2 and the partial transform requires it to be an integer of 2^j . Since financial data is not of dyadic length or even an integer multiple of it, it is necessary to truncate the series to the closest integer multiple of 2^j or 'pad' it with zeros. Pollock and Lo Cascio (2005) consider collections of wavelets and filters related to each other by other dilation factors than the factor of 2.

Particularly for large series the Discrete Wavelet Transform provides an efficient data reduction scheme, where the unimportant frequency bands, characterised by low amplitude are eliminated without significant loss of information. For an orthogonal wavelet basis, the number of convolution at each scale is proportional to its width at that scale. As a result a wavelet spectrum

¹² A complete set with its individual components being orthogonal and with a unit length is orthonormal.

is confined by discrete blocks of wavelet power. Since this number of blocks at each scale is a function of wavelet width, the compact representation of the signal is achieved. Both orthonormal and biorthogonal¹³ wavelet basis result in non-redundant transforms.

The reconstruction of the original signal from its Discrete Wavelet Transform is done through reversing the order of the decomposition procedure. Though the perfect reconstruction is only attainable for ideal halfband filters. Otherwise, Daubechies filters (Daubechies 1992) might provide such reconstruction under certain conditions.

Simulation Procedures¹⁴

Filters

The time series under consideration determines the choice of the wavelet function. For a smoothly fluctuating signal, a smooth function (e.g. a damped cosine) is appropriate. On the other hand, for a signal with sudden jolts a boxcar function (e.g. the Haar) is better suited. See Farge (1992) for comprehensive overview of the criteria determining the choice of the wavelet function.

Consider the period of a wavelet, p as the number of complete cycles or the rate of change of a signal. Rapid appearance/disappearance entail rapid rate of change of a signal, implying a short period, whereas a slow rate of change implies a long period. Varying the scale of the wavelet by changing its width, the wavelet is tuned to identify certain periods and their positions in time. Further consider the width of the filter, μ as the number of sample points on each side of its centre over which the filter gathers more than negligibly small information. Outside this limit the filter is modified to force its response to zero. A wide wavelet covers more time series and is capable of better identifying a particular period but is worse in detecting its exact position in time. On the other hand a narrow wavelet has broad response to features that are moderately close to the specified period with a good localisation in time.

The maximum length of the series, considered by a filter, is usually limited to twice the width plus one.

¹³ Separate bases orthogonal to each other but without individually forming orthogonal sets are biorthogonal.

¹⁴ It was pointed out by a referee that the difference between an 'experiment' and a 'simulation' is important. An experiment allows one to produce data in the lab or in the field, whereas a 'simulation' occurs with data already available. This difference in terminology appears to be crucial for 'financial experimentalists'.

For the current application, periodic sensitivity of a filter is given by the following relationship, $p \leq \mu \leq 2p$. The minimal periodic sensitivity that is practically useful is $p \leq \mu$. A narrow sensitivity filter is characterised by $\mu > 2p$, typically seeking a particular period. For pattern detection in financial time series, $\mu = p$ appears to be appropriate.

When a filter is applied to a part of a time series in which only the information about its past is available, the lag need to be determined explicitly. The specified lag of the filter determines the peak of its response. From theoretical perspective, to have samples on both sides of the centre, the filter should be lagged at its width length to avoid (unknown) future interferences. Although for practical applications such extreme might be unnecessary.

Significance Testing

Peak-based critical limit significance is used for signal analysis/processing. Two backgrounds are considered in this research: 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 the largest peak would reach this height by a random chance. The Monte Carlo simulation is used to generate the peak-based critical limits¹⁵. The simulated data is then fitted to bivariate, univariate or trivariate polynomials, depending on the number of factors affecting the significance.

Simulation Results

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. Created in June 1988, the CAC 40 is designed to represent the French stock exchange listed companies and serve as a support for derivatives markets. During the trading session 9.00 – 17.30, index levels are calculated in real time using the last trade quoted¹⁶ 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,

¹⁵ The Monte Carlo method is used to determine significance levels for a process with unknown statistical distribution.

¹⁶ On the exceptional basis in the event that at least 35% of the market capitalization of the index has not been traded, a price indicator - the forerunner is substituted for 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 simulation. 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 frequency¹⁷ of 30 seconds was obtained from Euronext.

The series containing 3779 eight and a half hours prices was extracted from the original data after subsampling. The study of such data frequency relates well 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. A similar sampling frequency was asserted as the most appropriate in designing an Artificial Neural Network by a typical off-floor trader in Kaastra and Boyd (1996). Lastly, for practical purposes of constructing a model, that generates reliable predictions, the issue of the realistic time needed to execute a strategy provides further support for the choice of such sampling interval.

Examining the price series reveals that the index exhibit an upward, but non-linear trend until around the year 2000, a downward trend until the spring 2003 and a mild upward trend afterwards. Persistent fluctuations around the trends, which increase in variability with downward moves, can also be seen. A possible complete cycle can be visually acknowledged between October 1998 and March 2003.

To test the data for normality, the Kolmogorov-Smirnov test (with Lilliefors' correction¹⁸) is used. Considering cumulative distribution functions, the test indicates that the data varies significantly from the pattern expected if the data was drawn from a population with a normal distribution. The maximum cumulative distance between the Gaussian distribution curve and the histogram for the data considered, K-S = 0.176. The P-value, $P_{K-S} < 0.001$ indicates that the difference in the means of the Gaussian and the analysed distributions is

¹⁷ A 'tick' is one logical unit of information.

¹⁸ The test assesses the normality of the standardized variables.

significantly greater than would be expected by chance.¹⁹

A measure of the symmetry of a distribution about its mean, skewness, $SK = 0.884$ being significantly nonzero, indicates that the distribution is asymmetric. A positive value also demonstrates a right tail. A skewness coefficient is considered to be significant as the absolute value of skewness over its standard error, $SK/SE_{SK} = 0.894/0.04$ is greater than 2. Interestingly, the values of kurtosis, $KU = -0.182$, being negative indicates that the distribution of the data considered is flatter than a normal distribution. A kurtosis coefficient $KU/SE_{KU} = 0.182/0.08$ although significant is just above 2.

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

$$r(t) = r(\Delta t; t) = x(t) - x(t - \Delta t), \quad (12)$$

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 a distribution of the price series. The distributions of the return measures are asymmetric and commonly display a right tail as it can be seen from Table 1. 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 a normal distribution expects. 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)²⁰. The P-values, PSW, given by this test, reject normality hypothesis for all three returns.²¹

¹⁹ The P-value determines the probability of being incorrect in concluding that the data is not normally distributed.

²⁰ Common for small samples, the test is given by

$$SW = \left[\sum_{i=1}^n a_i x_i \right]^2 / \sum_{i=1}^n (x_i - \bar{x})^2, \quad \text{where } a = (a_1,$$

$a_2, \dots, a_n) = m'V^{-1}[(m'V^{-1})(V^{-1}m)]^{-0.5}$, where m is a vector of expected values of standard normal order statistics, and V is the corresponding covariance matrix.

²¹ The P-value indicates how good the fit is.

Table 1: Return Measures

Tests Statistics	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-S _H - r ²	0.57-0.002-0.99	0.79-0.002-0.99	0.78-0.002-0.99

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

Long Memory

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. Applying the Growth in Cumulative Range algorithm, the Hurst exponent, H is estimated to measure the fractal dimension of the data series²². Its value ($H_p = 1.014$) indicates the long-term memory effect in the price series, being significantly greater than 0.5.²³

The long-memory hypothesis was also supported by the 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 (Ahn, Boudoukh et al. 2000). As stocks,

constituting an index react with different speed to the aggregate information, and the autocovariance of a diversified portfolio is the average of cross-covariance of its constituents, positive autocorrelation in the index is present despite its absence in the stocks.

The long memory hypothesis was examined for three return series. The autocorrelation function of the return (R) 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 patterns might be present. The autocorrelations of absolute ($|R|$) and squared returns (R^2), 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 Figure 1. The autocorrelations of square returns display a few peaks²⁴. 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.

²⁴ Note that theoretical autocorrelation of squared returns is meaningful only if the kurtosis of the returns is finite.

²² To compute the growth in range of a time series, a time scale index is used by partitioning number of observations, N elements and averaging over these groups. A normalized (rescaled) range value, R/S is the range divided by the standard deviation. The Hurst equation is thus given by $R/S = k \cdot N^H$, where H is the Hurst exponent.

²³ The coefficient standard error derived from the fit, $S_H = 0.001$. The goodness of fit index, based on a sum of squares criterion, $r^2 = 0.999$.

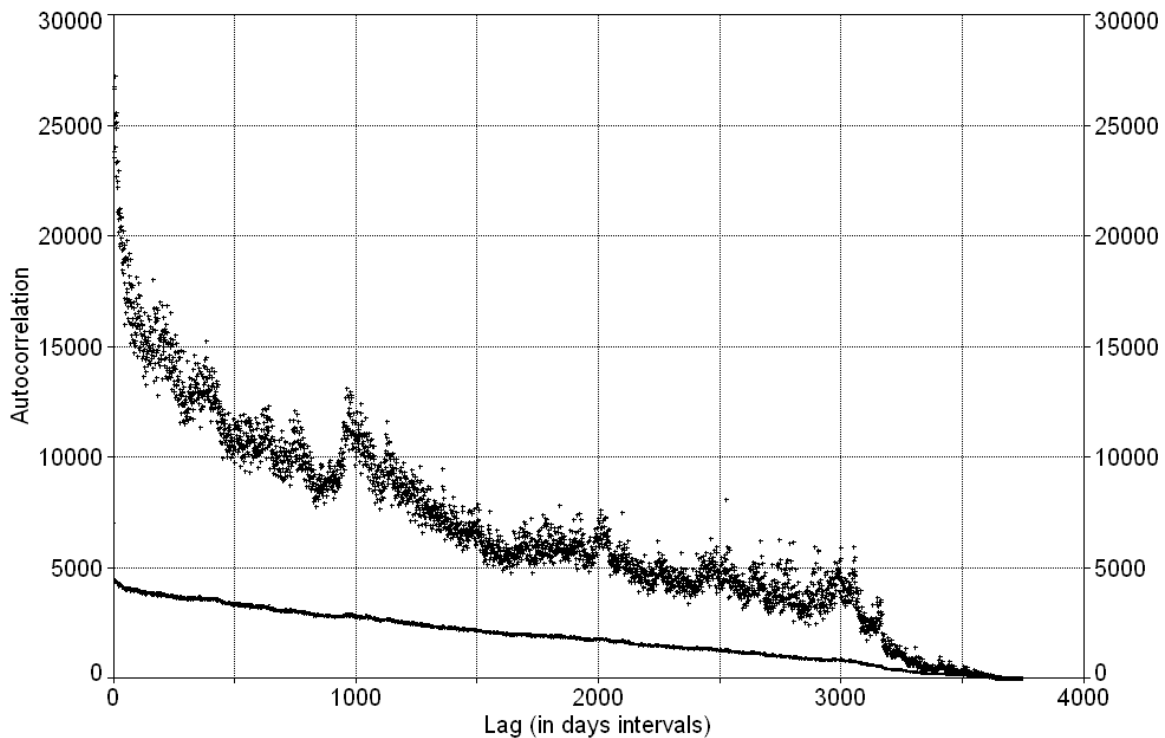


Figure 1. Autocorrelation Functions of Returns
 Squared Return (+++); Absolute Return (^ ^)

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

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 (Ding, Granger et al. 1993). At the same time, the long memory behaviour was not found to be the most apparent for the absolute returns, as claimed previously (Ding, Granger et al. 1993; Lobato and Savin 1998). 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 Andersen and Bollerslev (1977).

On the whole, the long memory was found to be an intrinsic feature of the system, present in different scales. Since the conventional statistical inference are limited in the case of the signals with long memory, the wavelet analysis that approximately decorelate the signal with long memory appears to

be particular appropriate. In addition it provides a platform for modelling nonstationary features without the exact knowledge of the correlation structure.

Low and High Frequencies

Assuming that lower level decomposition captures the long range dependencies whereas the higher levels identifies short term relationships, the signal is decomposed into low and high frequencies. To explore the features of interest with a minimum delay to be of value for financial applications, Continuous Wavelet Transform with Gabor wavelet function generated the best performance. This result is attributed to Gabor's ability of efficient filtering with minimal lag.

Considering a lowpass filter data smoothing, the best performance was obtain by the Gabor wavelet ($\mu = 350$; $p = 350$; $l = 350$) for both price and return series. Balancing spurious frequency responses against lag reduction benefits, (real) mirroring technique was found to be the most effective for the current application. Note that the issue of balancing costs and benefits of the mirroring is application-dependent and should be considered for the problem in hand.

In the high frequency filtering a highpass real Gabor ($\mu = 3$; $p = 5$; $l = 2$) generated the most accurate result for both price and return series. Applying a very narrow in Time Domain filter, Morlet shape modification is used to avoid contamination from slow trends. Such modification centres the filter and eliminate/reduce the problem.

The memory analysis of two series obtained indicates that the low frequency signal now carries most of the memory (the Hurst exponent, $H = 0.098$). Low persistency for the high frequency signal is confirmed by the value of H close to 0.5. These results demonstrate that the original data series is successfully decomposed into the signals, carrying separately individual characteristics of the original data. The analysis and modelling of a

particular signal is expected to be with limited undesirable interferences from its counterpart.

Now applying the decomposition technique as above, the individual outputs of the lowpass and highpass filters are examined. The simulation detects that low and high frequencies move into and out of phase with each other. Figure 2 presents low and high frequencies on the same plot²⁵. Around 03-09.00 [2501-2648] and 03-04.03 [3262-3302] one can see apparent phase shifts, rather than structural breaks, as it was claimed previously (Hayward 2005).

²⁵ Note that the graphs are shifted so as to cause the features to align in a visually obvious manner.

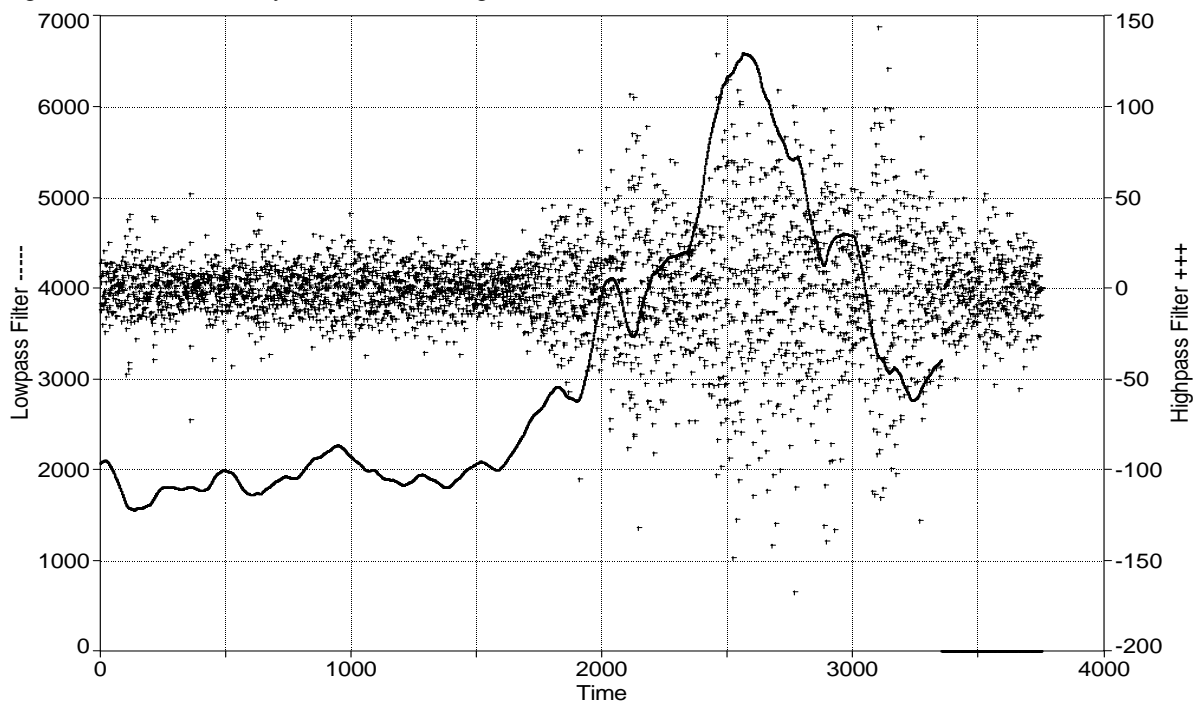


Figure 2: Low and High Frequencies.

Low frequency - solid curve; - high frequency - dotted curve.

Conclusion

The interactions between heterogeneous agents in relationship to each other actions rather than to the market news result in endogenous dynamics in the market. Differentiating economic agents' expectations according to their time dimension has valuable consequences for forecasting. Since the time scales of the market participants are the key characteristics of the market, the adaptive dynamics of such market prices reflect beliefs and behaviour of the dominant agents on the market.

Assuming that lower level decomposition capture the long range dependencies, whereas the higher levels identify short term relationships, the signal is

decomposed into low and high frequencies. Considering the individual outputs of the lowpass and highpass filters, the simulation detects that low and high frequencies move into and out of phase with each other. Two apparent phase shifts (rather than structural breaks, as it was claimed previously) were identified.

Testing the structural break hypothesis the simulation successfully differentiated structural breaks from phase shifts. Continuous wavelet transform spectra analysis of return series with the Morlet wavelet function identified that high frequencies follow low frequencies in time, indicating a possible casual relationship. Low frequency shocks affect all market participants,

though with some time delay. On the other hand high frequency shocks are ignored, at least for a while, by long term traders. Such information processing 'inefficiency' is a heterogeneity-driven asymmetry in information processing. Identified casual relationships are potentially exploitable in trading model development.

References

- Ahn, D.-H., J. Boudoukh, et al. (2000). Partial Adjustment or Stale Prices? Implications from Stock Index and Futures Return Autocorrelations, Stern Business School Working Paper.
- Andersen, T. G. and T. Bollerslev (1977). "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns." *Journal of Finance* 52: 975-1005.
- Bhattacharya, R. N., V. K. Gupta, et al. (1983). "The Hurst Effect under Trends." *Journal of Applied Probability* 20: 649-662.
- Dacorogna, M., R. Gencay, et al. (2001). *An Introduction to High-Frequency Finance*, Academic Press.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics.
- Diebold, F. (1986). "Modeling the Persistence of the Conditional Variances: A Comment." *Econometric Reviews* 5: 51-56.
- Ding, Z., C. W. J. Granger, et al. (1993). "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance* 1: 83-106.
- Farge, M. (1992). "Wavelet Transformations and their Applications to Turbulence." *Ann. Rev. Fluid Mech.*(24): 397-457.
- Granger, C. W. J. and N. Hyung (1999). "Occasional Structural Breaks and Long Memory." University of California San Diego Discussion Paper 99-14.
- Hayward, S. (2005). "The Role of Heterogeneous Agents' Past and Forward Time Horizons in Formulating Computational Models." *Computational Economics* 25(1-2): 25-40.
- Kaastra, I. and M. Boyd (1996). "Designing a Neural Network for Forecasting Financial and Economic Time Series." *Neurocomputing* 10: 215-236.
- Kurz, M. (1994). *Endogenous Economic Fluctuations, Studies in the Theory of Rational Beliefs*. Berlin, Springer.
- Lamoureux, C. L., W. (1990). "Persistence in Variance, Structural Change, and the GARCH Model." *Journal of Business and Economic Statistics* 8: 225-234.
- Lobato, L. N. and N. E. Savin (1998). "Real and Spurious Long-Memory Properties of Stock-Market Data." *Journal of Business and Economic Statistics* 16: 261-267.
- Mallat, S. G. (1989). "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7): 674-693.
- Mandelbrot, B. B. (1963). "The Variation of Certain Speculative Prices." *Journal of Business* 36(4): 394-419.
- Mandelbrot, B. B., I. J. Fisher, et al. (1997). *A Multifractal Model of Asset Returns*, Cowles Foundation Discussion Paper.
- Mandelbrot, B. B. and J. W. van Ness (1968). "Fractional Brownian Motion, Fractional Noises and Application." *SIAM Review* 10(4): 422-437.
- Peters, E. E. (1991). *Chaos and Order in Capital Markets*. New York, John Wiley & Sons.
- Pollock, D. S. G. and I. Lo Cascio (2005). *Orthogonality Conditions for Non-Dyadic Wavelet Analysis*. Queen Mary, University of London, Department of Economics, Working Papers.