Comput Econ (2014) 44:231–251
DOI 10.1007/s10614-013-9381-z

Timescale Analysis with an Entropy-Based Shift-Invariant Discrete Wavelet Transform

Stelios D. Bekiros

Accepted: 2 May 2013 / Published online: 21 May 2013

Abstract This paper presents an invariant discrete wavelet transform that enables point-to-point (aligned) comparison among all scales, contains no phase shifts, relaxes the strict assumption of a dyadic-length time series, deals effectively with boundary effects and is asymptotically efficient. It also introduces a new entropy-based methodology for the determination of the optimal level of the multiresolution decomposition, as opposed to subjective or ad-hoc approaches used hitherto. As an empirical application, the paper relies on wavelet analysis to reveal the complex dynamics across different timescales for one of the most widely traded foreign exchange rates, namely the Great Britain Pound. The examined period covers the global financial crisis and the Eurozone debt crisis. The timescale analysis attempts to explore the micro-dynamics of across-scale heterogeneity in the second moment (volatility) on the basis of market agent behavior with different trading preferences and information flows across scales. New stylized properties emerge in the volatility structure and the implications for the flow of information across scales are inferred.

Keywords Wavelets · Entropy · Exchange rates

JEL Classification C14 · C32 · C51 · F31

S. D. Bekiros (✉)
Department of Economics, European University Institute, Via della Piazzuola 43, 50133 Florence, Italy
e-mail: stelios.bekiros@eui.eu

S. D. Bekiros
Department of Accounting and Finance, Athens University of Economics and Business,
76, Patission str., 10434 Athens, Greece
e-mail: bekiros@aueb.gr

S. D. Bekiros
Rimini Centre for Economic Analysis (RCEA), Via Patara, 3, 47900 Rimini, Italy

1 Introduction

In contrast to the classical Fourier analysis, in which the frequency content of the underlying time series is assumed to be stationary along the time axis, wavelets are defined over a finite domain and unlike the Fourier transform they are localized both in time and in scale. This property, as opposed to the trigonometric or complex exponential functions in Fourier transform, makes wavelets ideal for analysing nonstationary signals especially those incorporating transient phenomena or singularities. Wavelet analysis works with "translates" and "dilates" of a single local function, the so-called *"mother wavelet"*, which is locally defined with compact support or decays sufficiently fast. However, the existence of wavelets is not a trivial analytical issue and the construction of classes of compactly supported wavelets was first addressed by Daubechies (1988). In that work a reliable methodology for obtaining orthogonal wavelet bases by translating and dilating the mother wavelet was provided. As a mother wavelet with compact support is located in a finite interval, the analysis of a singularity is performed by considering only those translates of the mother wavelet that overlap the singularity. Daubechies (1988) contribution was followed by the development of multiresolution analysis by Mallat (1989) and Coifman and Wickerhauser (1992).

Technically, Fourier and wavelet methods involve the projection of a signal onto an orthonormal set of components. Fourier projections are most naturally defined for functions restricted to $L^2(0, 2\pi)$ i.e., the set of square integrable functions in the interval $(0, 2\pi)$, because Fourier series have infinite energy but finite power when extended to being defined over the entire real axis. Based on the complex superposition of individual harmonics, the hypothesis is that over any segment of the time series the exact same frequencies hold at the same amplitudes, namely the signal is *homogeneous over time*. On the contrary, the basis functions in wavelet analysis are defined in $L^2(\mathbb{R})$ and are not necessarily homogeneous over time, meaning that they have narrow compact support so that they rapidly converge to zero as time approaches infinity. Such basis functions are called *wavelets,* in distinction to the trigonometric functions traditionally associated with *waves*.

The flurry of interest in economic applications of wavelets emerged in the mid-90s mostly by Ramsey and his collaborators. Ramsey et al. (1995) pursued a wavelet approach in detecting self-similarity in US stock prices. In addition, Ramsey and Lampart (1998a,b) used a wavelet-based scaling method to investigate the relationship and causality between money, income and expenditure. In other studies Goffe (1994) illustrated the application of wavelets to nonstationary macroeconomic data, in particular for the detection of discontinuities and the occurrence of sharp cusps. In recent works, Almasri and Shukur (2003) address the causal relation between spending and revenue at different timescales, while Gençay et al. (2002) look into dependencies between money, income, expenditure, growth and inflation. Fernandez (2005) deals with the estimation of systematic asset risk. Finally, it is also worth mentioning a stream of papers utilizing wavelet methodology to address theoretical econometric issues and devise new statistical approaches Lee J. and Y. Hong (2001).

1.1 Time Scaling in Economics Revisited

In natural sciences the utilization of sequences of timescales in the analysis of different modes of behavior, or different relationships between variables, is very common. In economics the notion of timescale is related to *time period* segmentation and the examined relationships are described as short-run and long-run, or broadly under the term *scaling laws* (Brock 1999). The scale decomposition often reveals the presence of deterministic regularities or statistical properties of the conditional moments that are seemingly independent of the scale details. However, in wavelet literature the concept of time scaling or "dilation" is quite different from that in economics. The different scales of the wavelet decomposition contain contributions of the signal in different frequencies (although timescale and frequency are not identical concepts as it is further explained below). Based on the selected function space, the time series are analysed into "fine" and "coarse" resolution components extracted from the application of *"father"* and *"mother"* wavelets. The former represent the smooth and low-frequency parts of a signal, whereas the latter incorporate the detailed and high-frequency features (Percival and Walden 2000).

While frequency analysis (Fourier) results in projecting the *entire* signal onto ever lower frequencies, time scaling with wavelets is concerned with projecting a *localized* component of the signal onto an increasingly broader base. Moreover, the basis functions are orthogonal across scales, so that the total variation/energy of the signal at any given point in time is obtained by adding the constituent components extracted at each of the scales. Although at first sight timescale could directly correspond to frequency there is only an indirect connection between these two concepts, as indicated by Priestley (1996). Intuitively, in a naïve interpretation, wide-support wavelets can be associated with low frequencies, while high-frequency analysis requiring high sampling rates can be provided by narrow-support components.

The multiresolution features of wavelet decomposition can be useful in econometric analysis. Through wavelet decomposition, the low-frequency content of the data that "captures" the true dynamic relationships can be extracted and the high-frequency fluctuations that might distort the underlying dependencies of the economy can be removed. This paper contributes to the literature by introducing an invariant transform that enables point-to-point comparison among all scales, contains no phase shifts, relaxes the strict assumption of a "dyadic-length" time series, deals effectively with "boundary effects" and is asymptotically efficient. In addition, beyond the existing practice that has utilized subjective judgment in considering the appropriate "depth" of the wavelet analysis, a new entropy-based methodology is introduced to determine the optimal level of decomposition. Finally, an empirical application in financial time series—in particular exchange rates—is pursued. These series are inherently characterized by chaotic patterns, fat tails and long-memory, particularly at high sampling frequencies. The results provide evidence of complex heterogeneous dynamics across and within different scales.

The paper develops as follows: Sect. 2 provides an overview of multiresolution methodology and presents a new shift-invariant discrete wavelet transform. Finally, Sect. 3 provides an empirical application and Sect. 4 the concluding remarks.

2 The Shift-Invariant Discrete Wavelet Transform (SIDWT)

2.1 Preliminaries

Definition 1 Let $\psi(.)$ be a real valued continuous function such that $\int_{-\infty}^{\infty} \psi(t)\, dt = 0$ and $\int_{-\infty}^{\infty} \psi(t)^2\, dt = 1$; then $\psi(.)$ defines a wavelet.

Considering that $h = (h_0, \ldots, h_{M-1})$ is a finite length wavelet filter, the properties of continuous wavelet functions such as integration to zero and unit energy, in discrete time are equivalently given by $\sum_{m=0}^{M-1} h_m = 0$ and $\sum_{m=0}^{M-1} h_m^2 = 1$. The wavelet filter is orthogonal to its even shifts for all nonzero integers$n$

$$\sum_{m=0}^{M-1} h_m h_{m+2n} = \sum_{m=-\infty}^{\infty} h_m h_{m+2n} = 0 \tag{1}$$

Definition 2 If $g = (g_0, \ldots, g_{M-1})$ denotes the complement low-pass (scaling) filter of the wavelet (high-pass) filter then according to Percival and Walden (2000), the scaling filter coefficients are estimated based on the quadrature mirror relationship[1] $g_m = (-1)^{m+1} h_{M-1-m}$ for$m = 0, \ldots, M-1$.

Theorem 1 *The T-length vector of the wavelet coefficients* $\mathbf{w}$ *for a time series* $\mathbf{y} = \{y_t\}_{t=1}^{T}$ *with dyadic length* $\left(T = 2^J\right)$ *is obtained as* $\mathbf{w} = \mathbf{W}\mathbf{y}$. *The* $T \times T$ *orthonormal matrix* $\mathbf{W}$ *defines the Discrete Wavelet Transform (DWT).*

The vector of wavelet coefficients can be decomposed into $J+1$ vectors

$$\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_J, \mathbf{s}_J]^T \tag{2}$$

where $\mathbf{w}_j$ is a $T/2^j$-length vector of wavelet coefficients corresponding to the scale of length $a_j = 2^{j-1}$ and $\mathbf{s}_J$ is a $T/2^J$-length vector of scaling coefficients associated with scale $2a_J$. The $\mathbf{W}$ matrix comprises the wavelet and scaling filter coefficients on a row-by-row representation. Hereby, the vector of zero-padded unit scale wavelet filter coefficients is defined in reverse order by $\mathbf{h}_1 = \left[h_{1,T-1}, h_{1,T-2}, \ldots, h_{1,1}, h_{1,0}\right]^T$, where the coefficients $h_{1,0}, \ldots, h_{1,M-1}$ are derived from an orthonormal wavelet family of length M and all values $M < t < T$ are zero. If $\mathbf{h}_1$ is circularly shifted by factors of two, $\mathbf{h}_1^{(2)} = \left[h_{1,1}, h_{1,0}, h_{1,T-1}, h_{1,T-2}, \ldots, h_{1,3}, h_{1,2}\right]^T$, $\mathbf{h}_1^{(4)} = \left[h_{1,3}, \ldots, h_{1,0}, h_{1,T-1}, h_{1,T-2}, \ldots, h_{1,5}, h_{1,4}\right]^T$ then the $T/2 \times T$ matrix $\mathbf{W}_1$ is defined as the collection of $T/2$ circularly shifted versions of $\mathbf{h}_1$, namely $\mathbf{W}_1 = \left[\mathbf{h}_1^{(2)}, \mathbf{h}_1^{(4)}, \ldots, \mathbf{h}_1^{(T/2-1)}, \mathbf{h}_1\right]^T$. In general, matrices $\mathbf{W}_j$ are defined by circularly shifting the vector $\mathbf{h}_j$ (the vector of zero-padded scale j wavelet filter coefficients) by factors of2^j. Additionally, $\mathbf{S}_J$ is a column vector with all elements equal to $1/\sqrt{T}$ (McCoy and Walden 1996). The $T \times T$ dimensional matrix $\mathbf{W}$ is $\mathbf{W} =$

[1] Quadrature mirror filters are used in engineering for perfect reconstruction of a signal without aliasing effects.

$[\mathbf{W}_1\mathbf{W}_2\ldots\mathbf{W}_J\mathbf{S}_J]^T$. From matrix $\mathbf{W}$ the wavelet filter coefficients for scales $1,\ldots,J$ are computed via the Inverse Discrete Fourier Transform (IDFT). Specifically, given the transfer functions of the wavelet ($H_{j,k}$) and scaling filters ($G_{J,k}$), the wavelet for scale $a_j = 2^{j-1}$ is estimated as the inverse DFT[2]

$$h_{j,l} = \mathcal{F}^{-1}\left\{H_{j,k}\right\} = \mathcal{F}^{-1}\left\{H_{1,2^{j-1}k \bmod T}\prod_{l=0}^{j-2} G_{1,2^m k \bmod T}\right\}, \quad k=0,\ldots,T-1 \quad (3)$$

with length $M_j = \left(2^j - 1\right)(M-1)+1$. In the same way, the scaling filter g_J for scale a_J is derived as the inverse DFT of $G_{J,k}$

$$g_{J,l} = \mathcal{F}^{-1}\left\{G_{J,k}\right\} = \mathcal{F}^{-1}\left\{G_{J,k} = \prod_{l=0}^{J-1} G_{1,2^m k \bmod T}\right\}, \quad k=0,\ldots,T-1 \quad (4)$$

Mallat (1989) introduced the "pyramid algorithm" for the implementation of the DWT. The following proposition proves how the coefficients are produced.

Proposition 1 *In the Mallat algorithm the data y_t are filtered using h_1 and g_1, then the outputs are subsampled to half their original lengths and the subsampled filter output from h_1 accounts for the wavelet coefficients. This process is repeated on the subsampled output from g_1 filter.*

The first step of the pyramid algorithm begins by convolving the data with each filter to obtain the following wavelet $w_{1,t} = \sum_{m=0}^{M-1} h_m y_{2t+1-m \bmod T}$ and scaling coefficients $s_{1,t} = \sum_{m=0}^{M-1} g_m y_{2t+1-m \bmod T}$ where $t = 0, 1, \ldots, T/2-1$. This also includes a downsampling operation, in that every other value of the input vector is removed. Consequently, the T-length vector of observations has been high-and low-pass filtered to obtain $T/2$ coefficients. The second step of the algorithm starts by "initializing" the sample now to be the scaling coefficients $\mathbf{s}_1$ and apply the aforementioned filtering procedure to obtain the second level of wavelet and scaling coefficients as $w_{2,t} = \sum_{m=0}^{M-1} h_m s_{1,2t+1-m \bmod T}$ and $s_{2,t} = \sum_{m=0}^{M-1} g_m s_{1,2t+1-m \bmod T}$ respectively with $t = 0, 1, \ldots, T/4-1$. By saving all wavelet coefficients and the final level of scaling coefficients the decomposition becomes $\mathbf{w} = [\mathbf{w}_1\mathbf{w}_2\mathbf{s}_2]^T$. This procedure is repeated up to $J = \log_2(T)$ times and provides the vector of wavelet coefficients in Eq. (2).

The inversion of the DWT is performed by upsampling the final wavelet and scaling coefficients, convolving them with their respective filters and adding the resulting vectors. Upsampling the vectors $\mathbf{w}_J$ and $\mathbf{s}_J$ of the final DWT level produces the new vectors $\mathbf{w}_J^0 = \left[0 w_{J,0}\right]^T$ and $\mathbf{s}_J^0 = \left[0 s_{J,0}\right]^T$. Now the vector of scaling coefficients $\mathbf{s}_{J-1}$ is given by $s_{J-1,t} = \sum_{m=0}^{M-1} h_m w^0_{J,t+m \bmod 2} + \sum_{m=0}^{M-1} g_m s^0_{J,t+m \bmod 2}$ with $t = 0, 1$ and it is twice that of $\mathbf{s}_J$. This is repeated until the first level of all coefficients has been upsampled, in order to produce the original vector of data observations, i.e.,

[2] The modulus operator is required in order to deal with the boundary of a finite length vector of observations.

$$y_t = \sum_{m=0}^{M-1} h_m w^0_{1,t+m \bmod T} + \sum_{m=0}^{M-1} g_m s^0_{1,t+m \bmod T} \quad t = 0, 1, \ldots, T-1.$$

The DWT results in the additive decomposition of the time series. The most important feature of the multiresolution analysis is the reverse process of signal synthesis.

Proposition 2 *Let* $\mathbf{D}_j = \mathbf{W}_j^T \mathbf{w}_j$ *define the wavelet detail corresponding to changes in the series* $\mathbf{y}$ *at scale* a_j *for the level* $j = 1, \ldots, J$. *The coefficients* $\mathbf{w}_j = \mathbf{W}_j \mathbf{y}$ *represent the part of the signal due to wavelet analysis at scale* a_j, *while* $\mathbf{W}_j^T \mathbf{w}_j$ *is the part of the wavelet synthesis attributable to scale* a_j.

The final wavelet detail $\mathbf{D}_{J+1} = \mathbf{S}_J^T \mathbf{S}_J$ is equal to the sample mean of the $T = 2^J$ observations (Gençay et al. 2002). The multiresolution analysis is defined for each observation y_t as the linear combination of wavelet detail coefficients, i.e., $y_t = \sum_{j=1}^{J+1} D_{j,t}$ $t = 0, \ldots, T-1$. Similarly, $\mathbf{A}_j = \sum_{k=j+1}^{J+1} \mathbf{D}_k$ is the cumulative sum of the variations of the details and is defined as the j-th level *wavelet approximation* for $0 \leq j \leq J$ with $\mathbf{A}_{J+1}$ being a vector of zeros. The j-th level *wavelet rough* $\mathbf{R}_j = \sum_{k=1}^{j} \mathbf{D}_k$, $1 \leq j \leq J+1$ incorporates the remaining lower-scale details where $\mathbf{R}_0$ is the zero vector. Overall, the vector of observations may be decomposed for all j through a wavelet approximation and rough as

$$\mathbf{y} = \mathbf{A}_j + \sum_{k=1}^{j} \mathbf{D}_j = \mathbf{A}_j + \mathbf{R}_j \tag{5}$$

Definition 3 Orthonormality of the matrix $\mathbf{W}$ implies, as in case of DFT, that the DWT is an efficient, variance preserving transform[3,4] *i.e.*, $\|\mathbf{w}\|^2 = \sum_{j=1}^{J} \sum_{t=0}^{T/2^j-1} w_{j,t}^2 + s_{J,0}^2 = \sum_{t=0}^{T-1} x_t^2 = \|\mathbf{y}\|^2$.

Consequently, the energy $\|\mathbf{y}\|^2$ is decomposed on a scale-by-scale basis as $\|\mathbf{y}\|^2 = \sum_{j=1}^{J} \|\mathbf{w}_j\|^2 + \|\mathbf{s}_J\|^2$, where $\|\mathbf{w}_j\|^2$ is the energy of $\mathbf{y}$ attributed to variance at scale a_j, and $\|\mathbf{s}_J\|^2$ is the remaining energy in a_J scales or higher. As $\mathbf{D}_j^T \mathbf{D}_j = \mathbf{w}_j^T \mathbf{w}_j$ and $\mathbf{A}_J^T \mathbf{A}_J = \mathbf{s}_J^T \mathbf{s}_J$ apply for $1 \leq j \leq J$ (due to orthonormality of $\mathbf{W}$ and $\mathbf{S}$), an equal decomposition is $\|\mathbf{y}\|^2 = \sum_{j=1}^{J} \|\mathbf{D}_j\|^2 + \|\mathbf{A}_J\|^2$.

2.2 Formal Description of the SIDWT

The classical (decimated) DWT involves subsampling of the filter output to half the original length. This leads to a serious drawback, namely the transform is not invariant in the *real-axis*. Specifically, the DWT of a shifted signal is not the shifted version

[3] The energy of a vector—proportionate to variance—is defined as the sum of its squared coefficients.

[4] It can be also proven via matrix operations: $\|\mathbf{y}\|^2 = \mathbf{y}^T \mathbf{y} = (\mathbf{W}\mathbf{w})^T \mathbf{W}\mathbf{w} = \mathbf{w}^T \mathbf{W}^T \mathbf{W}\mathbf{w} = \mathbf{w}^T \mathbf{w} = \|\mathbf{w}\|^2$.

of the DWT of the signal.[5] Alternatively, an undecimated DWT can be implemented without the subsampling technique. According to Coifman and Donoho (1995) undecimated versions of the DWT could handle any sample size T, while the J-th order DWT restricts the sample size to a multiple of 2^J. Moreover, they are invariant to circularly shifting the time series, a property that does not hold for the DWT. Finally, an undecimated wavelet variance estimator is asymptotically more efficient than the DWT estimator (Percival 1995). Hence, in this study a new variation of the undecimated DWT, namely the Shift-Invariant DWT (SIDWT) is proposed via the following theorem.

Theorem 2 *The SIDWT is defined as follows: Let* $\mathbf{y}$ *be an arbitrary* T*- length vector of observations. The* $(J+1)\,T$ *- length vector of SIDWT coefficients* $\tilde{\mathbf{w}}$ *is obtained as* $\tilde{\mathbf{w}} = \tilde{\mathbf{W}}\mathbf{y}$*, where* $\tilde{\mathbf{w}}$ *is a* $(J+1)\,T \times T$ *matrix. The SIDWT coefficient vector, as in DWT, is organized into* $J+1$ *vectors*

$$\tilde{\mathbf{w}} = \left[\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_J, \tilde{\mathbf{s}}_J\right]^T \tag{6}$$

Let $\tilde{\mathbf{w}}_j$ is a $T/2^j$- length vector of wavelet coefficients associated with the scale of length $a_j = 2^{j-1}$ and $\tilde{\mathbf{s}}_J$ is a $T/2^J$- length vector of scaling coefficients corresponding to a length scale of $2^J = 2a_J$. The direct conversion to DWT could be implemented for a dyadic length $\left(T = 2^J\right)$ sample, via subsampling and rescaling of the SIDWT. The converted DWT wavelet coefficients are $w_{j,t} = 2^{j/2}\tilde{w}_{j,2^j(t+1)-1}$ with $t = 0, \ldots, T/2^j - 1$, and the scaling coefficients $s_{J,t} = 2^{J/2}\tilde{s}_{J,2^J(t+1)-1} t = 0, \ldots, T/2^J - 1$. In correspondence to the orthonormal matrix of the DWT, the SIDWT matrix $\tilde{\mathbf{W}}$ comprises $J+1$ submatrices of $T \times T$ dimension expressed as $\tilde{\mathbf{w}} = \left[\tilde{\mathbf{w}}_1 \tilde{\mathbf{w}}_2 \ldots \tilde{\mathbf{w}}_J \tilde{\mathbf{S}}_J\right]^T$. The SIDWT utilizes the rescaled filters from DWT, $\tilde{\mathbf{h}}_j = \mathbf{h}_j/2^j$ and $\tilde{\mathbf{g}}_J = \mathbf{g}_J/2^J$ with $(j = 1, \ldots, J)$. The $T \times T$ submatrix $\tilde{\mathbf{w}}_1$ is constructed by circularly shifting the rescaled wavelet filter vector $\tilde{\mathbf{h}}_1$ by integer units to the right, i.e., $\tilde{\mathbf{w}}_1 = \left[\tilde{\mathbf{h}}_1^{(1)}, \tilde{\mathbf{h}}_1^{(2)}, \tilde{\mathbf{h}}_1^{(3)}, \ldots, \tilde{\mathbf{h}}_1^{(T-2)}, \tilde{\mathbf{h}}_1^{(T-1)}, \tilde{\mathbf{h}}_1\right]^T$ and it can be interpreted as the circularly shifted version of DWT submatrix $\mathbf{W_1}$. The other matrices $\tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_J$ are similarly constructed through replacing $\tilde{\mathbf{h}}_1$ by $\tilde{\mathbf{h}}_j$.

Proposition 3 *The new implementation algorithm starts with the data* y_t*, which is no longer limited to dyadic length, and filters with* $\tilde{h}_1$ *and* $\tilde{g}_1$ *to obtain the* T*-length vectors of wavelet and scaling coefficients* $\tilde{\mathbf{w}}_1$ *and* $\tilde{\mathbf{s}}_1$*, yet without utilizing the downsampling operation.*

In the first step the data is convolved with each filter to obtain the wavelet $\tilde{w}_{1,t} = \sum_{m=0}^{M-1} \tilde{h}_m y_{t-m \bmod T}$ and scaling coefficients $\tilde{s}_{1,t} = \sum_{m=0}^{M-1} \tilde{g}_m y_{t-m \bmod T}$ where $t = 0, 1, \ldots, T-1$. The second step of the SIDWT algorithm uses the "new" data, namely the scaling coefficients $\tilde{\mathbf{s}}_1$ from the previous step, and proceeds with the application of filtering to obtain the second level of wavelet and scaling coefficients i.e., $\tilde{w}_{2,t} =$

[5] Shifting a signal simply means delaying its start in the *real-axis*. In mathematical terms, delaying a function is represented by $f\,(t-d)$.

$\sum_{m=0}^{M-1} \tilde{h}_m \tilde{s}_{1,t-m \bmod T}$ and $\tilde{s}_{2,t} = \sum_{m=0}^{M-1} \tilde{g}_m \tilde{s}_{1,t-m \bmod T}$ with $t = 0, 1, \ldots, T-1$. The resulting T-length decomposition is $\tilde{\mathbf{w}} = \left[\tilde{\mathbf{w}}_1 \tilde{\mathbf{w}}_2 \tilde{\mathbf{s}}_2\right]^T$. The procedure is repeated up to $J = \log_2(T)$ times in order to provide the full vector of SIDWT coefficients in Eq. (6). In the Inverse transform the final-level wavelet and scaling coefficients are convolved with their respective filters and the resulting vectors are added up. Therefore, the vectors $\tilde{\mathbf{w}}_J$ and $\tilde{\mathbf{s}}_J$ of the final level are filtered and combined to produce the vector of scaling coefficients $\tilde{\mathbf{s}}_{J-1}$ in $J-1$ level $\tilde{s}_{J-1,t} = \sum_{m=0}^{M-1} \tilde{h}_m \tilde{w}_{J,t+m \bmod T} + \sum_{m=0}^{M-1} \tilde{g}_m \tilde{s}_{J,t+m \bmod T}$ where $t = 0, 1, \ldots, T-1$. The length of $\tilde{\mathbf{s}}_{J-1}$ is the same as $\tilde{\mathbf{s}}_J$. The algorithm is repeated until the first level of coefficients produce the original vector of observations $y_t = \sum_{m=0}^{M-1} \tilde{h}_m \tilde{w}_{1,t+m \bmod T} + \sum_{m=0}^{M-1} \tilde{g}_m \tilde{s}_{1,t+m \bmod T}$ with $t = 0, 1, \ldots, T-1$.

It is emphasized that the SIDWT associates the wavelet details and approximation coefficients with zero-phase filters, thus the details and approximations correspond directly to the original time series in perfect alignment.

Corollary 1 *The multiresolution analysis in SIDWT assumes* $y_t = \sum_{j=1}^{J+1} \tilde{D}_{j,t}, t = 0, \ldots, T-1$ *where* $\tilde{D}_{j,t}$ *is the t- th element of* $\tilde{\mathbf{D}}_j = \tilde{\mathbf{w}}_j^T \tilde{\mathbf{w}}_j$ *for* $j = 1, \ldots, J$. *The SIDWT wavelet approximations and rough are respectively defined as* $\tilde{A}_{J,t} = \sum_{k=j+1}^{J+1} \tilde{D}_{k,t}$ *and* $\tilde{R}_{j,t} = \sum_{k=1}^{j} \tilde{D}_{k,t} t = 0, \ldots, T-1$ *and the original time series are given by*

$$\mathbf{y} = \tilde{\mathbf{A}}_j + \sum_{k=1}^{j} \tilde{\mathbf{D}}_j = \tilde{\mathbf{A}}_j + \tilde{\mathbf{R}}_j \tag{7}$$

Percival and Mofjeld (1997) proved that undecimated, invariant transforms are energy (variance) preserving transforms. Thus, SIDWT is an efficient transform and the total variance of the time series is given by $\|\mathbf{y}\|^2 = \sum_{j=1}^{J} \left\|\tilde{\mathbf{w}}_j\right\|^2 + \|\tilde{\mathbf{s}}_J\|^2$.

2.3 Non-dyadic Length and Time-Invariance

The SIDWT relaxes the assumption of a dyadic length time series which is not always applicable in practice. In case of the DWT a "signal extension" process is usually employed, which involves "padding" the time series with values and increase its length to the next power of two. Ogden (1997) reports various methods such as padding with zeros, repeating the last observation (polynomial of order zero), higher-order polynomials, periodic extension, and numerical integration.

Corollary 2 *SIDWT is time-invariant (as proven in Theorem 2) as opposed to the classical DWT which exhibits some translation in time even after applying signal extension.*

(a) SIDWT is not an orthogonal basis, thus it produces an over-determined (redundant) representation of the series that has advantages in regards to statistical inference.

(b) As the SIDWT entails approximate zero-phase filtering, the details at each timescale and the approximation contain the same number of observations and line up in time with the original series.

This property makes the SIDWT a particularly useful tool in the analysis of time-dependent processes.

2.4 "Periodic Extension" Pattern for Boundary Distortions

Furthermore, the application of the DWT to finite-length time series brings up the crucial issue of "boundary distortions", which concerns the problematic estimation of the remaining wavelet coefficients when the end of the series is encountered in the wavelet transform. To deal with end-of-sample distortions, the border should be treated differently from the other parts of the signal. Although various theoretical methods are available to tackle this problem, they are rather inefficient from a practical viewpoint (Cohen et al. 1993). A common practical technique applied mainly in Fourier analysis involves either taking observations from the initial part of a T-length periodic series to finish computations at the end, or the entire series are duplicated/reflected about the last observation. This may be reasonable for some time series with strong seasonal effects but cannot be applied universally in practice.

Lemma 1 *Based on Theorem 2 (and Proposition 3) the proposed SIDWT employs a specialized "periodic extension" pattern to deal with boundary effects.*

(a) If the series length is odd, the series is first extended by adding an extra-sample equal to the last value on the right. Then a minimal periodic extension is performed on each side. The extension mode used for the inverse SIDWT is the same to ensure a perfect reconstruction.
(b) Using these boundary coefficients, the SIDWT retains its numerically stability (Herley 1995).

2.5 Entropy-Based Determination of the Optimal Decomposition Level

In the literature the depth (level) of the multiscale wavelet decomposition is usually determined arbitrarily or based on some economic rationale with regard to the examined time scales. Alternatively in this study an optimal decomposition is pursued with respect to the *minimization* of an entropy-related criterion. Classical entropy-based criteria describe information-relevant properties for an accurate representation of a given signal (Coifman and Wickerhauser 1992).

Lemma 2 *The depth (level) of the SIDWT is estimated on the basis of the sample length, the selected wavelet class and the boundary-distortion method (based on Theorem 2). The entropy of each level is estimated step-wise and it is compared with the one from the previous level. If it is decreased then the new decomposition "reveals" interesting, non-redundant information and the decomposition continues.*

Corollary 3 *The optimal level is determined at the minimum value of the Shannon entropy-related criterion.*[6] *In the following expressions y is the signal and* c_i *represents the details and the j-th level approximation coefficient of y for scales* $j = 1, \ldots, J$ *in an orthonormal basis. The entropy E must be an additive cost function such that* $E(0) = 0$ *and* $E(y) = \sum_j E(c_j)$*. The entropy for the coefficients in each level is defined as*

$$E_{Shannon}(c_j) = -c_j^2 \cdot \log(c_j^2) \tag{8}$$

and thus for the entire signal is $E_{Shannon}(y) = -\sum_j c_j^2 \cdot \log(c_j^2)$*, with the convention* $0 \cdot \log(0) = 0$.

2.6 Wavelet Filter Class Selection

The selection of a particular wavelet filter class is not trivial in practice and depends upon the complexity of the spectral density function and the underlying features of the data in the time domain. Optimally, in most data sets a balance between frequency localization and time localization should be pursued. According to Gençay et al. (2001) and (2002), a moderate length wavelet filter (e.g., length eight) adequately captures the stylized features of financial data. Given that the wavelet basis functions are used to represent the information contained in the time series, they should "mimic" its underlying features.[7] Usually, smoothness and (a)symmetry are the most crucial factors in selecting suitable wavelet basis functions (Gençay et al. 2002; Ramsey and Lampart 1998a).

Assumption(s) 1 (a) The SIDWT coefficients are calculated from the Daubechies family of compactly supported wavelet filters, which are well localized in time (Daubechies 1992). The wavelet and scaling coefficients of the Daubechies class are $w_{1,t} = \sum_{m=0}^{M-1} h_m y_{2t-m}$ and $s_{1,t} = \sum_{m=0}^{M-1} g_m y_{2t-m}$ respectively with $t = M/2, 1 + M/2, \ldots, T/2$.

(b) The Daubechies wavelet filter of length eight, (db8) is selected in order to balance smoothness, length and symmetry (Jensen and Whitcher 2000; Gençay et al. 2001).

(c) It achieves an "ideal compromise" between competing requirements in that it has reasonably narrow, compact support, is fairly smooth, has vanishing

[6] The Shannon entropy criterion shows a downward trend until a minimum value—corresponding to a "threshold" scale level—is reached and then it begins to rise revealing that further signal decomposition "contains" redundant information. The maximum level of decomposition tried in this study is ten, based on the appropriate "translation" of the wavelet scales into economic time horizons.

[7] For instance, if the data appear to be constructed of piecewise linear functions, then the Haar wavelet may be the most appropriate choice, while if the data is fairly smooth, then a longer filter such as the Daubechies asymmetric wavelet filter may be desired.

moments, is twice differential, nearly symmetric and has a moderate degree of flexibility.[8]

3 Empirical Application

3.1 Data Description and Econometric Analysis

The time series used for the empirical exercise is the Great Britain Pound (GBP) daily closing rates (5 days) denoted relative to United States dollar (USD) i.e., the GBP/USD ratio. The currency returns are defined as $r_t = \log(P_t) - \log(P_{t-1})$, where P_t is the closing level on day t, and the volatility series as the absolute value of the returns $u_t = |r_t|$ as in Jensen and Whitcher (2000) and Gençay et al. (2002). The data sample span the period from January 5, 1999 to May 10, 2010, namely from the introduction of the Euro until the ECB and the IMF agreed on a program of bond purchases and an defence package of 750,000€ in order to deal with the Eurozone sovereign-debt crisis. The multiresolution analysis is performed in sub-periods based on the application of stability tests for structural breakpoints as well as on economic foundation. In this study *March 10, 2000* is used as the first breakpoint, when the technology NASDAQ Composite index peaked at 5,048.62 (intra-day peak 5,132.52), more than double its value just a year before, corresponding to the date when the dot-com bubble "burst" Greenspan (2007). This date also coincides with the after-Euro era. Then, the global financial crisis of 2008 is examined that was triggered by a liquidity shortfall in the United States banking system and resulted in the collapse of large financial institutions, turbulence and downturns in stock markets around the world Krugman (2009). The crisis began to affect the financial sector in *February 22, 2007*, when HSBC, the world's largest bank of 2008, wrote down its holdings of subprime-related mortgage-backed-securities by $10.5 billion. This particular date is used as the second breakpoint. Finally, the sovereign debt crisis in early 2010 concerning Eurozone countries such as Greece, Spain, Ireland, and Portugal is also investigated. It led to the widening of bond yield spreads on credit default swaps between these countries and other Eurozone members, especially regarding Germany. The date *December 8, 2009* is set as the third breakpoint, corresponding to the first Greek rating cut by Fitch. Further to the economic justification the breakpoint selection is statistically tested via the application of Chow's test (Chow 1960) for known (imposed) breaks and the cumulative sum (CUSUM) test (Brown et al. 1975) for unknown points. These tests are applied both on the GBP return and volatility series to investigate also for volatility breaks (McConnell and Perez-Quiros 2000; Dijk et al. 2005). All possible calendar combinations are examined, i.e. one imposed breakpoint of March 10, 2000, February 22, 2007 or December 8, 2009 separately, then two points (3 cases) and finally all three points/dates of structural change. The Chow test in this paper uses the methodology of McConnell and Perez-Quiros (2000) who estimate an AR(1) model with a constant for each sub-sample separately, to see whether there are significant

[8] In the empirical part, alternative choices of wavelet classes were also applied, but the results were very robust to such changes and the current selection appeared to be the most balanced.

differences in the estimated equations. A significant difference indicates a structural change in the relationship. Two statistics for the Chow test are used, namely the log-likelihood ratio χ^2 and the F-statistic, which are both based on the comparison of the restricted and unrestricted sum of squared residuals. For the GBP, the null hypothesis of no structural change for the one break of February 22, 2007 as well as for the specific case of the two breaks of February 22, 2007 and December 8, 2009, is rejected at 10 % significance level. The CUSUM test is based on the cumulative sum of the recursive residuals. It detects parameter instability—though marginally—around the region of the February 22, 2007 break (e.g., 1,950–2,150 observations). Regarding the volatility series, the Chow test rejects the null of no structural change at 1 % level for the February 22, 2007 breakpoint. Moreover, it does not reject the null hypothesis for the one break of December 8, 2009. In addition, it rejects the null for all date combinations not including December 8, 2009. Finally, the CUSUM test strongly detects parameter instability around the February 22, 2007 breakpoint. The selected breakpoints have also been verified with the Bai and Perron (2003) and Zivot and Andrews (1992) tests. Thus, the breakpoint of *February 22, 2007* is finally selected for the return and volatility series, offsetting statistical and economic motivation. Overall, the examined sub-periods are the following: P_I: January 5, 1999 to February 21, 2007 (2,122 observations), P_{II}: February 22, 2007 to May 10, 2010 (838 observations). In addition, the entire sample period P_T: January 5, 1999 to May 10, 2010 (2,960 observations) is comparatively investigated in an extensive robustness analysis.

The descriptive statistics for the all series are presented in Table 1. The Jarque-Bera multiplier in all periods is statistically significant, thereby implying that the return distributions are not normal. In general, kurtosis for returns in all periods is larger than normal which indicates the presence of fat tails, extreme observations and possibly volatility clustering. Kurtosis is also significantly higher than normal for the distribution of the absolute returns. As indicated by skewness, GBP returns have a longer left tail. Based on the Ljung-Box Q-statistic, the hypothesis that all correlation coefficients of the returns up to 12 are jointly zero is rejected in the majority of cases. Therefore, it can be inferred that the return series present some linear dependence. On the contrary, the statistically significant serial correlations in the volatility series imply nonlinear dependence due possibly to clustering effects or conditional heteroscedasticity. The differences between the two periods P_I and P_{II} are quite evident in Table 1 where a significant increase in variance can be observed in P_{II} as well as increased fat-tailedness of the return and volatility distributions reflected in the higher kurtosis. Additionally, P_{II} witnessed many occasional negative spikes as it can be inferred from the skewness of GBP returns. Volatility series also present more spikes in P_{II}. The results from testing nonstationarity are also presented in Table 1. Specifically, the Augmented Dickey–Fuller (ADF) and Phillips-Perron (PP) unit root tests are applied to the log-levels, returns and volatility series. The lag lengths were selected using the Schwartz Bayesian Information Criterion (SIC), while for the PP test the bandwidth was automatically selected using Newey and West (1994) method with Bartlett kernel. The GBP variable appears to be nonstationary in log-levels and stationary in log-returns based on the reported p-values. Specifically, the ADF and PP tests indicate that the null of a unit root cannot be rejected at 1 % for the log-levels

Table 1 Descriptive statistics and unit root tests

Statistic	GBP returns		
	P_T	P_I	P_{II}
Mean	0.000	0.000	0.000
Std. dev.	0.006	0.005	0.008
Skewness	−0.379	−0.008	−0.553
Kurtosis	5.844	3.640	5.547
JB test	1068.48*	36.28*	269.17*
Q(12)	11.38	9.24	16.52
Statistic	GBP volatility		
	P_T	P_I	P_{II}
Mean	0.004	0.004	0.006
Std. dev.	0.004	0.003	0.005
Skewness	2.042	1.250	1.978
Kurtosis	10.191	4.720	8.316
JB test	8433.69*	814.37*	1532.96*
Q(12)	832.81*	42.96*	465.12*

Unit root testing		P_T				P_I				P_{II}			
Periods		ADF		PP		ADF		PP		ADF		PP	
Variables		ADF_c	ADF_τ	PP_c	PP_τ	ADF_c	ADF_τ	PP_c	PP_τ	ADF_c	ADF_τ	PP_c	PP_τ
GBP series	P_t	0.65	0.94	0.65	0.93	0.87	0.43	0.88	0.42	0.87	0.77	0.87	0.76
	r_t	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*
	u_t	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*

Price variables are in logarithms and reported numbers for the augmented Dickey–Fuller (ADF) and Phillips-Perron (PP) test are *p*-values (both are one-sided tests of the null hypothesis that the variable has a unit root). The index *c* indicates that the test allows for a constant, while τ for a constant and a linear trend. The number of lags for the ADF was selected using the Schwarz information criterion. The lag truncation for the PP test was selected using Newey and West (1994) automatic selection with Bartlett kernel. (*) denotes significance at 1 % confidence level. The periods are P_I: 01/05/1999-02/21/2007, P_{II}: 02/22/2007-05/10/2010 and P_T: 01/05/1999-05/10/2010

in all periods, regardless of whether a constant and linear trend or only a constant is included in the deterministic component. Furthermore, both tests show that the log-returns and volatility series are stationary as the null can be soundly rejected for all periods.[9] The combined results from the unit root tests suggest that the investigated log-levels appear to be $I(1)$ processes.

3.2 Multiscale Analysis based on the SIDWT

As opposed to Fourier transforms and to classical or other undecimated DWTs where the depth (level) of the multiscale wavelet decomposition is determined arbitrarily or based on trial and error, for SIDWT it is estimated on the basis of the boundary-distortion method and the minimum-entropy decomposition criterion. The step-wise entropy of each level is calculated and it is compared to the one from the previous level. If it is decreased then the new decomposition "reveals" non-redundant information and the decomposition continues. The results of the optimal entropy-based decomposition level for the GBP returns and volatility are presented in Table 2. For the GBP FX returns the optimal decomposition level is the seventh for P_T and P_I and the sixth in P_{II}, while for the volatility series the minimum value of the Shannon entropy criterion is calculated at the fourth scale. Moreover, the appropriate "depth" of the wavelet analysis corresponds to a specific "translation" of the wavelet scales into time horizons. Table 3 provides insight on the relation between SIDWT levels and time scales for the time series. Each scale of the wavelet multiscale analysis corresponds to a frequency interval, or equally an interval of periods, therefore each scale is associated with a range of time horizons that span from several days to one year. For instance, the wavelet detail $\mathbf{D}_2$ is associated with a frequency range of 4–8 days or 0.8–1.6 weeks, while $\mathbf{D}_4$ (optimal level for the volatility series) is associated with approximately one month.[10] At scale level $j = 7$, the frequency range corresponds to a cycle length between a period of 2.1–4.3 quarters, namely between a semester and a yearly variation. Hence, the GBP returns series are decomposed at scale level $j = 7$ therefore "containing" up to yearly frequencies, while the volatility series are analyzed up to the $j = 4$ scale which is associated with a frequency range of 0.8–1.6 weeks (1 month). The economic interpretability is also substantiated. In particular, due to the nature of the series it is reasonable to investigate the returns from daily to yearly frequencies, whereas up to monthly variations for the volatility.[11]

[9] Due to the nature of volatility, it is assumed that there is no time trend in the series in the long run (Nikkinen et al. 2006). However, the unit root tests were also performed with a time trend and the results remain unchanged. Moreover, the test results are generally not sensitive to the number of lags used.

[10] Thereafter the notation $\mathbf{D}_j$ (and not the $\tilde{\mathbf{D}}_j$ used in Sect. 3) corresponds to the SIDWT details, to enhance readability.

[11] In empirical applications, quarterly, semi-annual or yearly volatility is not interesting for the economic analysis of high-frequency (daily) FX series, nor "traded" in currency markets, as opposed to daily or weekly volatility. However, the analysis of the returns up to yearly variations can be very useful in detecting FX market linkages with macroeconomic fundamentals (e.g., GDP, CPI, Interest rates) or in producing multi-step ahead price forecasts.

Table 2 Minimum-entropy decomposition level

Wavelet level	GBP/USD returns			GBP/USD volatility		
	P_T	P_I	P_{II}	P_T	P_I	P_{II}
Raw	0.966	0.538	0.428	0.966	0.538	0.428
1	0.998	0.789	0.446	0.429	0.328	0.199
2	1.033	0.768	0.503	0.453	0.357	0.205
3	0.988	0.698	0.530	0.448	0.364	0.204
4	1.062	0.777	0.523	**0.392**	**0.289**	**0.193**
5	1.000	0.863	0.435	0.424	0.295	0.227
6	0.942	0.912	**0.367**	0.573	0.494	0.287
7	**0.715**	**0.456**	0.407	1.377	0.643	1.203
8	1.623	0.533	1.224	3.009	0.512	2.500
9	1.244	0.788	0.803	7.482	1.320	6.648
10	1.629	0.939	1.296	15.405	2.029	9.764

Bold numbers report the corresponding optimal level of decomposition for each time series. It indicates the minimum value of the Shannon entropy criterion for the wavelet details and j-th level approximation

Table 3 Translation of wavelet scales into time horizons

Wavelet scale	Time horizons				
	Days	Weeks	Months	Quarters	Years
a_1	2–4				
a_2	4–8	0.8–1.6			
a_3	8–16	1.6–3.2			
a_4	16–32	3.2–6.4	0.8–1.6		
a_5	32–64	6.4–12.8	1.6–3.2	0.5–1.1	
a_6	64–128	12.8–25.6	3.2–6.4	1.1–2.1	
a_7	128–256	25.6–51.2	6.4–12.8	2.1–4.3	0.5–1.1

Each scale of the SIDWT corresponds to a frequency interval, or conversely an interval of periods, and thus each scale is associated with a range of time horizons. The time horizons are expressed in base units (daily frequency) as follows: week = 5 trading days, month = 20 trading days, quarter = 60 trading days, year = 240 trading days

The wavelet *approximation* and *details* of the returns and volatility series, as well as the original time series in all periods are depicted in Figs. 1 and 2. As it can be demonstrated from the graphs, a major advantage of the new SIDWT representation—as opposed to the classical and other undecimated DWTs—is that it contains no phase shifts in the wavelet *components,* which enables aligned temporal comparison among all scales. Furthermore, the employment of Daubechies (db8) class in the implementation of the SIDWT decomposition reveals that the wavelet details display a complicated structure that cannot be attributable to an oscillation at a single frequency, that for example a Short-time Fourier transform with an arbitrary window function could capture. Moreover, the Haar class usually utilized in DWT application would fail to "mimic" the underlying features of the time series as they are not constructed of piecewise linear functions. The use of the Daubechies (db8) in the SIDWT balances smoothness, length and symmetry, in that it has reasonably narrow and compact support, it has vanishing moments and it is twice differential and nearly symmetric.

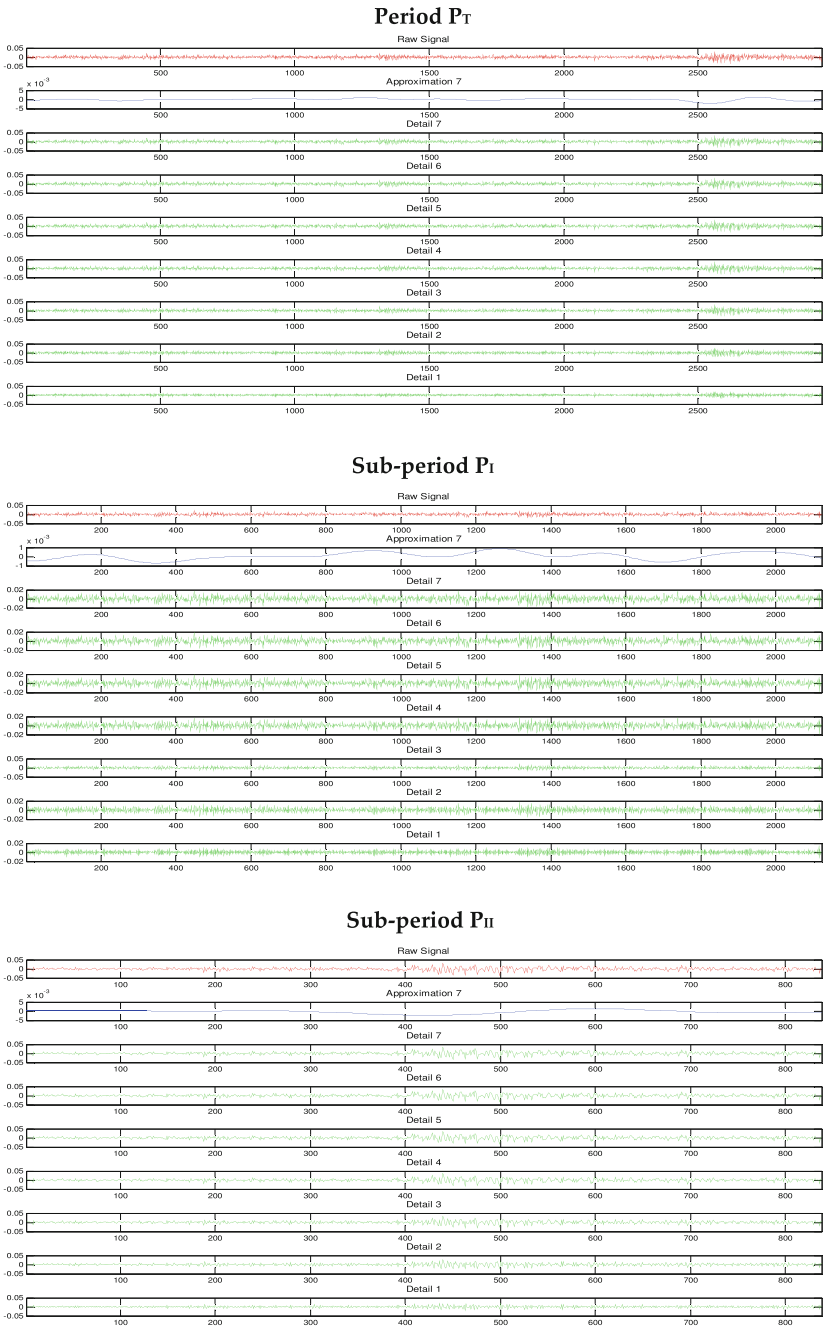

Fig. 1 SIDWT Decomposition of GBP returns. The results of SIDWT (db8) multiresolution wavelet analysis include the D_1–D_7 wavelet details and the 7-th level approximation A_7. The raw signal is also displayed

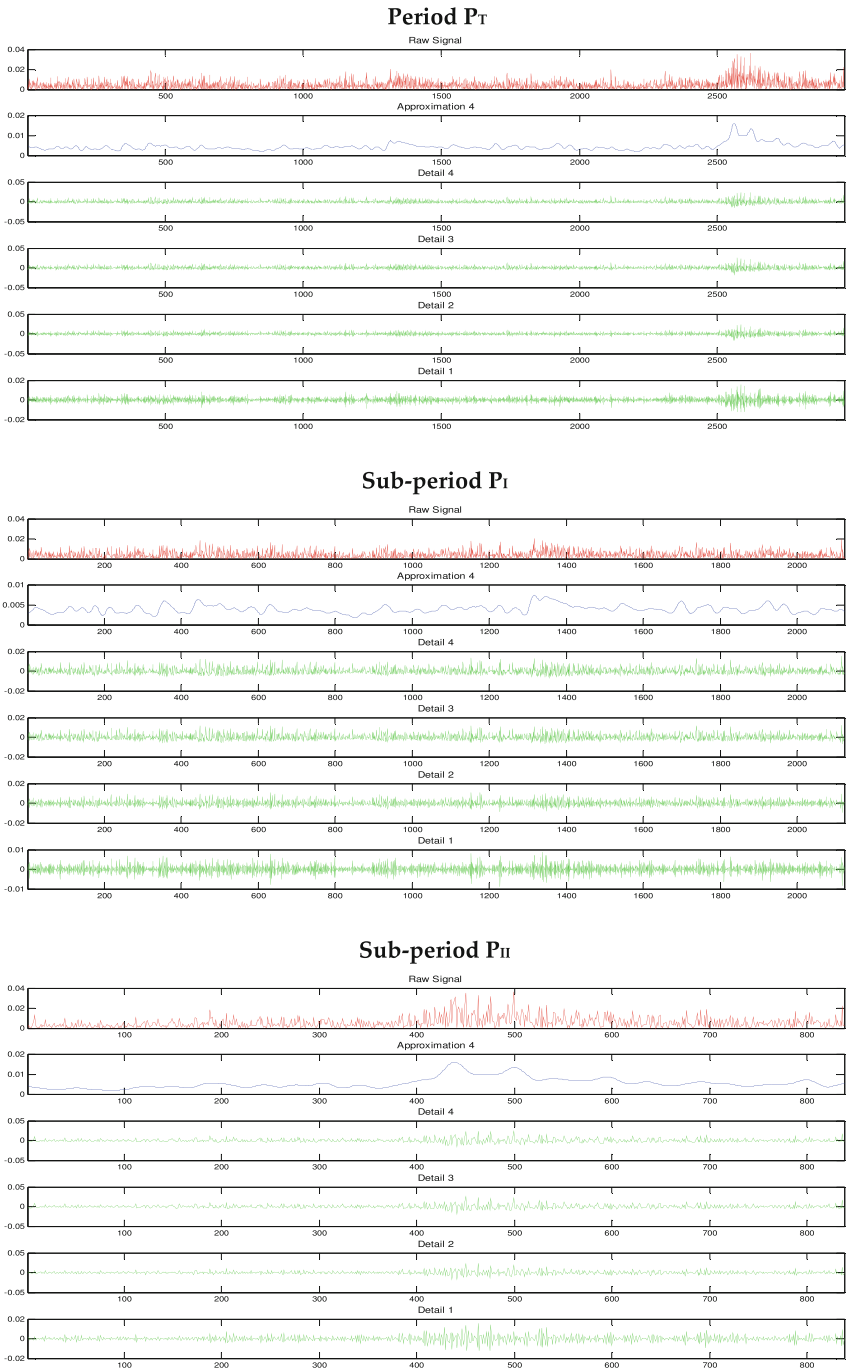

Fig. 2 SIDWT Decomposition of GBP volatility. The results of SIDWT (db8) multiresolution wavelet analysis include the D_1–D_4 wavelet details and the 7-th level approximation A_4. The raw signal is also displayed

Firstly, the low- and high-frequency components of GBP returns and volatility are examined. In case of the return series for all periods there are no significant differences in high- and low-frequency dynamics. All components display a fairly low oscillation amplitude. Basically, there is no notable "activity" in high scales at all levels as a direct result of the trend-removal procedure, albeit in all periods the return fluctuations are amplified after the beginning of 2008, i.e. after entering the financial crisis period, where a regime switch occurs. The increased volatility is mostly manifested in detail $\mathbf{D}_1$ of the GBP volatility series in P_{II}, which is associated with oscillations of 2–4 days period length, but also in the second, third and fourth scale corresponding to oscillations with a period of approximately 1 week, 1.6–3.2 weeks and 0.8–1.6 months respectively. Furthermore, in P_T of the volatility series there is a turbulent pattern in the high frequency harmonic associated with 2–4 days, which might be attributed to traders with short-term trading horizons. The switching regime appearing in the GBP return details immediately after the crisis outbreak is also depicted in P_{II} details of the volatility.[12] It is notable that persistent oscillations are present in all detail components of the volatility in P_I, indicating a near-cyclical pattern in low scales for the pre-crisis period and probably "neutral" mean-reverting trading behavior. This is also depicted in the $\mathbf{A}_4$ approximation in the volatility series in P_I, but not in P_{II} and P_T, where possibly a split in the long-run trend component signifies the entry in the high volatility regime of the Eurozone debt crisis. One important aspect is the SIDWT scale-dependent duration of regime switches. Specifically, a high volatility regime initiated by a market information flow appears to persist longer at the lower frequency associated also with longer trading horizons, as opposed to a high-frequency horizon. This is demonstrated for the GBP volatility. Overall, the duration of regimes seems to be longer for high-scale horizons whereas low-scale behavior results in frequent regime changes.

Secondly, the "vertical" heterogeneity in the variability pattern is examined across scales. In case of GBP returns all scales seem to contribute to the raw series variance. Likewise, the $\mathbf{D}_1$ detail volatility in all periods dominates over the aggregate raw signal oscillation amplitude, albeit other frequency components incorporate lower information. It appears that a low frequency shock embedded in the long-run approximation wavelet coefficients, might lead a high frequency response by a short time period, as in the case of the crisis emergence shown in the volatility series. Consequently, vertical heterogeneity demonstrates trader behavior with different time horizons. The highest approximation scale of the trading mechanism "involves" fundamentalists who trade on longer time horizons, whilst at low scales short-term traders operate with time horizons of a few days up to a week or month. Each trader class possesses a homogeneous behavior, but the combination of these classes in all time-scales generates the aggregate time series. In financial markets, a high-scale shock penetrates through all scales, while a low-scale shock fades out quickly and has no impact in the long-run dynamics. The most characteristic example of a low-frequency shock that affected all

[12] The structural changes in the wavelet approximations mentioned throughout this section have also been tested with the Chow' and CUSUM tests as well as with the Bai and Perron (2003) and Zivot and Andrews (1992) tests. For the details, the switching regimes have been verified via a Markov-switching model with two regimes, using an AR(1) specification in each regime.

scales and market agents is the sovereign debt crisis. Moreover, the "across-scales" interrelationship of the various regimes is one-directional, in that a low regime volatility state at low frequencies affects the oscillation state at higher frequencies. On the contrary, high variability at a low frequency does not necessarily entail a high volatility at higher frequencies. This follows empirical evidence that markets "cool off" after a shock at low scales in a much shorter period than after a high-scale "fundamental" or institutional regime switch.

Next, the impact of the global financial crisis can be observed from February 22, 2007 onwards (obs. 2,123). A meticulous consideration of Figs. 1 and 2 reveals that the effect of the subprime crisis is not evident in the GBP exchange rate market in the extent at which affected stock markets globally. Also in this case there is no clear evidence of a contagion effect between stock and currency markets. Finally, the Eurozone sovereign debt crisis is investigated (December 8, 2009).[13] The estimated wavelet components at all scales clearly indicate that the GBP currency market entered into a high volatility state towards the beginning of the third quarter of 2008, that is before the end of 2009 where the Eurozone crisis was more formally acknowledged by the EU official bodies and IMF. Closer inspection of Figs. 1 and 2 reveals that low- and high-frequency coefficients (corresponding to daily-yearly horizons) as well as the long-run persistent component (mostly in the volatility series), exhibit a high volatility regime after approximately the third quarter of 2008 (near obs. 400 in P_{II}). Via the power of the SIDWT analysis, these signs could have been safely considered as precursor signals of an imminent crisis. Moreover, the high volatility state is not uniform across the scales; at lower scales and especially at the finest scale a_1, the time span of the regime becomes wider. For return and volatility detail $\mathbf{D}_1$ (approximately 2–4 days) in periods P_T and P_{II}, the wavelet components display a high volatility regime through the end of the sample. At scales $\mathbf{D}_2$–$\mathbf{D}_4$ the volatility state shows a smaller amplitude. Consequently, for short-term traders the turbulence continues within 2010, whereas for investors, the turmoil mostly lasts until the beginning of 2010. In all cases, a high volatility regime across all scales has been in effect since the end of 2008.

Overall, the empirical investigation highlighted the advantages of the SIDWT in comparison to the classical DWT and other undecimated DW transforms as well as against Fourier analysis. The invariability of the SIDWT lead to an aligned signal association among all scales, while the assumption of a "dyadic-length" time series that was relaxed allowed for a robustness analysis in any sub-sample regardless of the number of observations. The new SIDWT coped effectively with "boundary effects" as opposed to the other methods where the solution involved either taking observations from the initial part of a T-length periodic series to finish computations at the end, or the entire series being duplicated/reflected about the last observation. This may be reasonable for some time series with strong seasonal effects but cannot be applied universally in practice, as in the case of the GBP series. This problem was surmounted via the use of the specialized "periodic extension" pattern of the SIDWT. Finally, beyond the existing practice of the DWT, the new

[13] The first Greek credit rating cut by Fitch corresponds to obs. 2,851 in P_T.

SIDWT entropy-based methodology lead to the optimal selection of the decomposition level.

4 Conclusions

In contrast to simple aggregation/disaggregation at different time horizons, this study relied on wavelet multiresolution to analyze the inherent dynamics of the GBP FX market across different timescales (frequencies). This study attempted to reveal the micro-dynamics of across-scale heterogeneity in the second moment (volatility), on the basis of market agent behavior with different trading preferences and information flows across scales. In addition, the scale-dependent duration of regime switches was highlighted. Specifically, a high volatility regime initiated by a market information flow appeared to persist longer at the lower frequency associated also with longer trading horizons, as opposed to a high-frequency horizon. An asymmetry in volatility dependence across different time horizons was identified as an important stylized property. In that a low regime variability state at high scales identically affected the oscillation state at lower scales, while on the contrary high volatility at a high scale did not inevitably cause a high volatility at lower scales.

Technically this paper expanded the literature via the introduction of the Shift-Invariant Discrete Wavelet Transform that allowed for non-dyadic-length time series analysis and multi-scale point-to-point (aligned) comparison with respect to the initial series, both of which are of utmost importance in modern econometric applications. SIDWT utilized a new entropy-based methodology for the determination of the optimal "depth" of the multiresolution analysis, instead of subjective or ad-hoc approaches.

The application of wavelet analysis in economic modelling is still in its infancy and many properties are not yet fully explored. Multiscale wavelet decomposition could become a valuable means of exploring the complex dynamics of economic time series, as it allows for temporal and frequency analysis at the same time.

Acknowledgments This research is supported by the Marie Curie Fellowship (FP7-PEOPLE-2011-CIG, No. 303854) under the 7th European Community Framework Programme. I am grateful to the Editor Hans Amman for valuable comments. The usual disclaimers apply.

References

Almasri, G., & Shukur, A. (2003). An illustration of the causality relation between government spending and revenue wavelet analysis on Finnish data. *Journal of Applied Statistics*, *30*(5), 571–584.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*, 1–22.

Brock, W. (1999). Scaling in economics: a reader's guide. *Industrial and Corporate Change*, *8*(3), 409–446.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, *37*, 149–192.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, *28*, 591–605.

Cohen, A., Daubechies, I., Jawerth, B., & Vial, P. (1993). Multiresolution analysis, wavelets and fast wavelet transform on an interval. *CRAS Paris*, *316*, 417–421.

Coifman, R.R., & Donoho, D. (1995). Time-invariant wavelet de-noising. In A. Antoniadis & G. Oppenheim (Eds.), *Lecture Notes in Statistics*, (Vol. 103, pp. 125–150). Springer-Verlag, NY.

Coifman, R. R., & Wickerhauser, M. V. (1992). Entropy-based algorithms for best basis selection. *IEEE Transaction on Information Theory*, *38*(2), 713–718.
Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, *41*(7), 909–996.
Daubechies, I. (1992). *Ten lectures on wavelets. CBMS-NSF regional conference series in applied mathematics (SIAM)* (Vol. 61). Philadelphia, USA: Society for Industrial and Applied Mathematics.
Fernandez, V. (2005). The international CAPM and a wavelet-based decomposition of value at risk. *Studies of Nonlinear Dynamics & Econometrics*, *9*(4), Article 4.
Gençay, R., Whitcher, B., & Selçuk, F. (2001). Differentiating intraday seasonalities through wavelet multiscaling. *Physica A*, *289*(3–4), 543–556.
Gençay, R., Whitcher, B., & Selçuk, F. (2002). *An introduction to wavelets and other filtering methods in finance and economics*. San Diego: Academic Press.
Goffe, W. L. (1994). Wavelets in macroeonomics: An introduction. In D. Belsley (Ed.), *Computational techniques for econometrics and economic analysis* (pp. 137–149). The Netherlands: Kluwer Academic.
Greenspan, A. (2007). *The age of turbulence: Adventures in a new world*. Penguin Press.
Herley, C. (1995). Boundary filters for finite-length signals and time-varying filter banks. *IEEE Transactions on Circuits and Systems-II*, *42*, 102–114.
Hong, Y., & Kao, C. (2004). Wavelet-based testing for serial correlation of unknown form in panel models. *Econometrica*, *72*(5), 1519–1563.
Jensen, J., & Whitcher, B. (2000). *Time-varying long-memory in volatility: detection and estimation with wavelets*,technical report. Columbia, MO: Department of Economics, University of Missouri.
Krugman, P. (2009). *The return of depression economics and the crisis of 2008*. Norton.
Lee, J., & Hong, Y. (2001). Testing for serial correlation of unknown form using wavelet methods. *Econometric Theory*, *17*(2), 386–423.
Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intellignce*, *11*(7), 674–693.
McConnell, M. M., & Perez-Quiros, G. (2000). Output fluctuations in the United States: What has changed since the early 1980s? *The American Economic Review*, *90*, 1464–1476.
McCoy, E. J., & Walden, A. T. (1996). Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics*, *5*, 26–56.
Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, *61*(4), 631–653.
Nikkinen, J., Sahlström, P., & Vähämaa, S. (2006). Implied volatility linkages among major European currencies. *Journal of International Financial Markets, Institution and Money*, *16*(2), 87–103.
Ogden, R. T. (1997). On preconditioning the data for the wavelet transform when the sample size is not a power of two. *Communications in Statistics: Simulation and Computation*, *26*(2), 467–486.
Percival, D. B. (1995). On estimation of the wavelet variance. *Biometrika*, *82*(3), 619–631.
Percival, D. B., & Mofjeld, H. O. (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association*, *92*, 868–880.
Percival, D. B., & Walden, A. T. (2000). *Wavelet methods for time series analysis*. Cambridge, UK: Cambridge University Press.
Priestley, M. B. (1996). Wavelets and time-dependent spectral analysis. *Journal of Time Series Analysis*, *17*(1), 85–103.
Ramsey, J. B., Usikov, D., & Zaslavsky, G. M. (1995). An analysis of US stock price behavior using wavelets. *Fractals*, *3*(2), 377–389.
Ramsey, J. B., & Lampart, C. (1998a). The decomposition of economic relationships by time scale using wavelets: Money and income. *Macroeconomic Dynamics*, *2*, 49–71.
Ramsey, J. B., & Lampart, C. (1998b). The decomposition of economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics and Econometrics*, *3*(1), 23–42.
van Dijk, D., Osborn, D. R., & Sensier, M. (2005). Testing for causality in variance in the presence of breaks. *Economics Letters*, *89*(2), 193–199.
Zivot, E., & Andrews, D. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, *10*(3), 251–270.