



ELSEVIER

Physica A 289 (2001) 249–266

PHYSICA A

www.elsevier.com/locate/physa

# Scaling properties of foreign exchange volatility

Ramazan Gençay<sup>a,b,\*</sup>, Faruk Selçuk<sup>b</sup>, Brandon Whitcher<sup>c</sup>

<sup>a</sup>*Department of Economics, University of Windsor, Windsor, 401, Sunset ONT Canada, N9B 3P4*

<sup>b</sup>*Department of Economics, Bilkent University, Bilkent 06533, Ankara, Turkey*

<sup>c</sup>*EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

Received 14 June 2000

---

## Abstract

In this paper, we investigate the scaling properties of foreign exchange volatility. Our methodology is based on a wavelet multi-scaling approach which decomposes the variance of a time series and the covariance between two time series on a scale by scale basis through the application of a discrete wavelet transformation. It is shown that foreign exchange rate volatilities follow different scaling laws at different horizons. Particularly, there is a smaller degree of persistence in intra-day volatility as compared to volatility at one day and higher scales. Therefore, a common practice in the risk management industry to convert risk measures calculated at shorter horizons into longer horizons through a global scaling parameter may not be appropriate. This paper also demonstrates that correlation between the foreign exchange volatilities is the lowest at the intra-day scales but exhibits a gradual increase up to a daily scale. The correlation coefficient stabilizes at scales one day and higher. Therefore, the benefit of currency diversification is the greatest at the intra-day scales and diminishes gradually at higher scales (lower frequencies). The wavelet cross-correlation analysis also indicates that the association between two volatilities is stronger at lower frequencies. © 2001 Elsevier Science B.V. All rights reserved.

*PACS:* 05.20; 02.50; 05.40

*Keywords:* Foreign exchange volatility; Scaling; Wavelets; Multi-scaling

---

## 1. Introduction

In financial risk management, risk is assessed at different horizons which vary from intervals as small as a few minutes to longer horizons such as days or even months. A common practice in the risk management industry is that risk measures calculated at shorter horizons are converted into longer horizons by taking the corresponding scaling quantity into account. For instance, if a risk measure (e.g. standard deviation)

---

\* Corresponding author. Tel.: +1-519-2534232; fax: +1-519-9737096.

*E-mail address:* gencay@uwindsor.ca (R. Gençay).

is calculated at a 12 hour frequency, it is converted to a 30-day risk measure by multiplying this half-day standard deviation by  $\sqrt{60}$ . This type of scaling is only valid if the underlying data is identically and independently distributed which is not the case for financial time series.<sup>1</sup> Therefore, the type of scaling pointed out above may provide misleading inferences for financial time series. Diebold et al. [4] demonstrate this point with a simple volatility model and present a critical view of this practice.

In the literature, a number of authors have reported empirically observed scaling laws for foreign exchange and security markets. Müller et al. [5,6] and Guillaume et al. [7] report empirical evidence for scaling of absolute returns for foreign exchange rates. Mantegna and Stanley [8] and Ghashghaie et al. [9] provide examples of scaling properties for financial time series. Recently, Andersen et al. [10] provide further evidence that there are precise scaling laws for daily volatility of foreign exchange returns.<sup>2</sup>

An important issue is whether financial time series follow a single-scaling law across all horizons or whether scaling properties are themselves time varying, adhering to a multi-scaling law. This paper provides evidence that there is no unique global scaling in financial time series but rather scaling can be time varying.<sup>3</sup> In this paper, we investigate the scaling properties of high-frequency exchange rates based on a wavelet multi-scaling approach which decomposes the variance of a time series and the covariance between two time series on a scale by scale basis through the application of a discrete wavelet transformation. The wavelet methodology is specification free and therefore is robust to misspecifications which may originate from assuming a particular distribution of returns.

Our results indicate that exchange rate volatilities have different scaling law properties at different horizons. Particularly, there is a smaller degree of persistence in intra-day volatility as compared to volatility at one day or higher scales. The implication of these findings is that a multi-scale approach to scaling is more in line with the foreign exchange rates as opposed to a single-scaling constant (or exponent) for all horizons. Our findings further reinforce the fact that pragmatic scaling practices used in the financial risk management industry are inappropriate and may lead to wrong inferences and practices.

---

<sup>1</sup> As pointed out by Mandelbrot [1], Clark [2], Mantegna and Stanley [3], and many others, the return distributions of financial time series are leptokurtic. In Mandelbrot [1], the return distribution was studied to be a symmetric Lévy stable distribution. Mantegna and Stanley [3] has re-confirmed that the return distribution is Lévy stable for high-frequency returns except for tails, which are approximately exponential. The Gaussianity of the returns only reveal themselves at longer times scales over a month.

<sup>2</sup> Other authors, such as LeBaron [11] studied what might be responsible for the observed scaling laws by showing that a simple stochastic volatility model calibrated to actual data is capable of giving a power scaling type of results. LeBaron [11] indicates that this is a mere illusion since a simple stochastic volatility model is not scale invariant theoretically. Barndorff-Nielsen and Praue [12] argue that empirically observed scaling laws are largely due to the semi-heavy tailedness of the underlying distributions rather than to real scaling.

<sup>3</sup> In a recent paper, Pasquini and Serva [13] provide evidence on stock index series that volatility correlations follow power laws but the scaling exponent is not unique in supporting the multi-scaling hypothesis.

An important question for a portfolio manager or a trading manager is whether the individual volatilities move together.<sup>4</sup> In this study, estimated wavelet correlations indicate that the correlation coefficient between two volatilities increases with an increasing scale, up to one day. However, the correlation coefficient remains constant for scales one day and higher. These results imply that the benefits to portfolio diversification are the greatest at the lowest scale (intra-day horizons) and diminish with scales corresponding to one day and longer horizons.

In practice, trading managers focus on intra-day co-movements while the portfolio managers are usually interested in one day and higher dynamics as they have a longer investment horizon. Therefore, it is important to distinguish co-movements of volatilities at different scales (horizons). The wavelet multi-scale cross-correlations provide a convenient method of disentangling the cross-correlations on a scale by scale basis to determine the contribution of each scale to the overall cross-correlation dynamics. At low scales, there is hardly any significant cross-correlations, whereas middle scales are dominated by the intra-day seasonalities. At higher scales (lower frequencies), a persistent cross-correlation dynamics becomes more visible.

This paper is organized as follows: The wavelet methodology, including the maximum overlap discrete wavelet transformation, wavelet variance, wavelet covariance and wavelet correlation, is described in Section 2. Empirical results are presented in Section 3. We then conclude.

## 2. Scaling and wavelet analysis

### 2.1. Aggregation versus multiresolution

To search for a possible scaling behavior in a discrete stochastic process  $x_t$ , a common procedure<sup>5</sup> is to aggregate the process over non-overlapping time intervals (horizons) to form

$$\tilde{a}_t(\tau) \equiv \frac{1}{\tau} \int_{(t-1)\tau}^{t\tau} x_s \, ds, \quad t = 0, \pm(T+1), \pm 2(T+1), \dots \quad (1)$$

The aggregated process  $\tilde{a}_t(\tau)$  has the same distributional properties of the original process through rescaling via  $\tau^\alpha$ . The quantity of interest  $\alpha$  is the scaling parameter, allowing us to scale up or down the process to different time horizons. Since we are simply averaging across time, the aggregated process also becomes more Gaussian-like with increasing  $\tau$ . A useful alternative is to compute all possible shifted time intervals of length  $\tau$ , thus producing the process  $\tilde{a}_t(\tau)$ . This is implemented by allowing the computation in Eq. (1) to be performed for all integers instead of only integer multiples

<sup>4</sup>The existence of co-movements of different volatilities is explored in a series of multivariate volatility models such as factor models, as in Ref. [14] or common persistent components as studied by Bollerslev and Engle [15].

<sup>5</sup>Aggregation is used in the recent literature by, for example, Pasquini and Serva [13] and LeBaron [11].

of  $T + 1$ . These two competing processes will be investigated further when we define the wavelet variance.

Aggregation is closely related to multiresolution analysis via the discrete wavelet transform (DWT). The key to multiresolution analysis is to represent a process through a series of coarse approximations and details [16]. This approximation is performed by first defining an orthonormal basis of shifted and translated functions.

$$\{\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k)\} \quad (2)$$

based on a single function  $\phi$ . The scaling coefficients  $v_{1,k}$  are defined via

$$v_{1,k} \equiv \int_{-\infty}^{\infty} x_t \phi_{1,k}(t) dt = \int_{-\infty}^{\infty} x_t \sqrt{2} \phi(t/2 - k) dt .$$

There is an inherent downsampling (removal of every other value) present that guarantees orthogonality of this transform. As with the aggregation process, we may ignore the downsampling and perform additional rescaling of the wavelet function in order to produce a “maximally overlapped” sequence of scaling coefficients  $\tilde{v}_{1,k}$ . Coarser approximations are obtained by simply projecting the process  $x_t \equiv v_{0,k}$  onto higher-order scaling functions.

The information contained in the difference between two adjacent approximations, say  $v_{0,k}$  and  $v_{1,k}$ , is given by the wavelet coefficients  $w_{1,k}$ . Under a multiresolution framework, the wavelet coefficients are obtained through a projection of the process onto shifted and translated versions of a wavelet function  $\psi$ . Combining the two sub-series, one scaling coefficients and the other wavelet coefficients, we arrive at the next finer approximation to the original process.

If we consider the Haar scaling function

$$\phi^{(H)}(t) = \begin{cases} 1, & -1 < t \leq 0, \\ 0, & \text{otherwise,} \end{cases}$$

then it is easy to see that the aggregate process  $\bar{a}_t(\tau)$  and the scaling coefficients  $v_{j,k}$  are identical under the condition  $\tau = 2^j$ . This is an indication of how aggregation and multiresolution analysis are related. We can go beyond this simple identity and utilize the wavelet coefficients instead of the scaling coefficients to characterize self-similar behavior in observed stochastic processes. We may also use higher order Daubechies wavelet functions to compute the transform. If the scaling coefficients for level  $j$  are associated with averages of length  $2^j$ , then the level  $j$  wavelet coefficients (which are differences of averages half this length) are associated with changes at scale  $\tau_j \equiv 2^{j-1}$ . This is true for all families of Daubechies compactly supported wavelets.

## 2.2. The wavelet variance

The wavelet variance is introduced here in order to estimate the scaling parameter of a stochastic process. Abry et al. [17] and Percival and Walden [18] both advocate using such an estimator for investigating scaling behavior. We restrict ourselves to analyzing realizations of stationary processes with finite length  $N$ . This brings the

issue of boundary conditions into the computation procedure. We choose to simply reflect the time series at the final observation, thus producing a series of length  $2N$ , and computing the wavelet transform on this series using periodic boundary conditions.

Let  $\tilde{h}_l = \{\tilde{h}_0, \tilde{h}_1, \dots, \tilde{h}_{N-1}\}$  denote the rescaled wavelet filter coefficients from a Daubechies compactly supported wavelet family (rescaled such that  $\tilde{h} = h/2^{1/2}$ ) and let  $\tilde{g}_l$  be the corresponding rescaled scaling filter coefficients, defined via the quadrature mirror relationship  $\tilde{g}_{1,m} = (-1)^{m+1}\tilde{h}_{1,L-1-m}$ . For an observed series  $\mathbf{X}$  of length  $N$ , applying these filter coefficients, and not downsampling the output, produces  $J \equiv \lceil \log_2 N \rceil$  vectors of wavelet coefficients

$$\tilde{\mathbf{W}}_j = \{\tilde{W}_{j,0}, \tilde{W}_{j,1}, \dots, \tilde{W}_{j,N-1}\}, \quad j = 1, \dots, J, \tag{3}$$

computed via

$$\tilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t = 0, \dots, N-1,$$

and one vector of scaling coefficients  $\tilde{\mathbf{V}}_J = \{\tilde{V}_{J,0}, \tilde{V}_{J,1}, \dots, \tilde{V}_{J,N-1}\}$  through

$$\tilde{V}_{J,t} \equiv \sum_{l=0}^{L_J-1} \tilde{g}_{J,l} X_{t-l \bmod N}, \quad t = 0, \dots, N-1.$$

This is known as the maximal overlap discrete wavelet transform (MODWT).<sup>6</sup> As discussed when comparing aggregation and multiresolution analysis, the term “maximal overlap” is used to indicate that all possible shifted time intervals were computed. Thus, orthogonality of the transform is lost but it has been shown that the wavelet variance utilizing MODWT coefficients is more efficient than the one obtained through the orthonormal DWT. Percival [19] gives the asymptotic relative efficiencies for the wavelet variance estimator based on the orthonormal DWT compared to the estimator based on the MODWT using a variety of power law processes.

The wavelet variance is defined to be the variance of the wavelet coefficients at scale  $\tau_j$ ; i.e.,  $v_X^2(\tau_j) \equiv \text{Var}\{\tilde{W}_{j,t}\}$ . In fact, due to the inherent differencing associated with wavelet functions, if a sufficiently long wavelet filter is used the wavelet coefficients have mean zero and thus  $v_X^2(\tau_j) \equiv E\{\tilde{W}_{j,t}^2\}$ . To form an unbiased estimator of the wavelet variance using the MODWT, we remove all coefficients affected by the periodic boundary conditions to yield

$$\tilde{v}_X^2(\tau_j) \equiv \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t}^2,$$

where  $\tilde{N}_j = N - L_j + 1$  and  $L_j \equiv (2^j - 1)(L - 1) + 1$  is the length of the scale  $\tau_j$  wavelet filter. The wavelet variance decomposes the variance of a process on a scale by scale basis, thus allowing us to inspect how different horizons behave relative to one another. This is related to the fact that the spectrum (modulus squared of its Fourier

<sup>6</sup>For details on the MODWT and its relationship to the orthonormal DWT see Percival and Walden [18].

transform) decomposes the variance of a process on a frequency by frequency basis. For example, if the process we observe is a power law process with spectrum given by  $S_X(f) \propto |f|^\alpha$ , then the wavelet variance has the following approximate relation to wavelet scale:

$$v_X^2(\tau_j) \propto \tau_j^{-\alpha-1}.$$

Observing a linear relationship between  $\log(v_X^2(\tau_j))$  and  $\log(\tau_j)$  would thus indicate scaling behavior. We use this property when analyzing 10 year records of 5-min foreign exchange rates.

### 2.3. Confidence intervals for the wavelet variance

The asymptotic variance of the MODWT-based estimator of the wavelet variance is given by

$$\lim_{N \rightarrow \infty} \tilde{N}_j \text{Var}\{\tilde{v}_X^2(\tau_j)\} = 2 \int_{-1/2}^{1/2} S_{W_j}^2(f) df \equiv 2\mathcal{V}_j,$$

where  $S_{W_j}(f)$  is the spectrum of the scale  $\tau_j$  wavelet coefficients. It follows that the random interval

$$\left[ \tilde{v}_X^2(\tau_j) - \Phi^{-1}(1-p) \left( \frac{2\mathcal{V}_j}{\tilde{N}_j} \right)^{1/2}, \tilde{v}_X^2(\tau_j) + \Phi^{-1}(1-p) \left( \frac{2\mathcal{V}_j}{\tilde{N}_j} \right)^{1/2} \right],$$

captures the true wavelet variance and corresponds to a  $100(1-2p)\%$  confidence interval for  $v_X^2(\tau_j)$ . For Gaussian processes, the quantity  $\mathcal{V}_j$  will have to be estimated using the autocovariance sequence of the scale  $\tau_j$  MODWT wavelet coefficients to obtain an approximate confidence interval.<sup>7</sup>

The last sentence of the previous paragraph made the important assumption of Gaussianity with regard to the MODWT coefficients at all scales. We may relax this assumption and slightly modify the results from above to develop an approximate confidence interval valid for non-Gaussian processes. Let  $\tilde{\mathbf{U}}_j$  be the vector of mean-corrected squared wavelet coefficients; i.e.,  $\tilde{U}_{j,t} = \tilde{W}_{j,t}^2 - E\{\tilde{W}_{j,t}^2\}$  for  $t=0, 1, \dots, N-1$ . Serroukh et al. [20] showed that the asymptotic variance of  $\tilde{v}_X^2(\tau_j)$  is given by

$$\lim_{N \rightarrow \infty} \tilde{N}_j \text{Var}\{\tilde{v}_X^2(\tau_j)\} = S_{W_j^2 - v_X^2(\tau_j)}(0) = S_{U_j}(0),$$

and further proposed a multitaper spectral estimator for the spectrum evaluated at frequency zero on the right-hand side. This approach enables us to produce an approximate confidence interval applicable to non-Gaussian processes.

### 2.4. The wavelet covariance and correlation

The wavelet scale analysis of univariate time series may be easily generalized to multiple time series by defining the concept of the wavelet covariance between  $\mathbf{X}$

<sup>7</sup> See Percival [19] and Percival and Walden [18] for an extensive discussion of approximate confidence intervals for the wavelet variance.

and  $\mathbf{Y}$ . The wavelet covariance is defined to be the covariance between the scale  $\tau_j$  wavelet coefficients of  $\mathbf{X}$  and  $\mathbf{Y}$ , i.e.,  $\gamma_{XY}(\tau_j) \equiv \text{Cov}\{\tilde{W}_{j,t,X}, \tilde{W}_{j,t,Y}\}$ . As with the wavelet variance, assuming a sufficient length for the wavelet filter guarantees each series of wavelet coefficients has mean zero and therefore  $\gamma_{XY}(\tau_j) \equiv E\{\tilde{W}_{j,t,X}\tilde{W}_{j,t,Y}\}$ .

An unbiased estimator of the wavelet covariance using the MODWT is given by

$$\tilde{\gamma}_{XY}(\lambda_j) \equiv \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1} \tilde{W}_{j,t,X} \tilde{W}_{j,t,Y},$$

where all wavelet coefficients affected by the boundary are removed. The wavelet covariance decomposes the covariance between two processes on a scale by scale basis and provides a unique method for attributing levels of association between the processes with different horizons. Whereas complicated associations may be present in the observed processes, the wavelet covariance is able to determine which scales (horizons) are significantly contributing to these associations.<sup>8</sup>

Just as the usual correlation coefficient is a function of the covariance between two series of observations, standardized by their standard deviations, the wavelet correlation is composed of the wavelet covariance between  $\mathbf{X}$  and  $\mathbf{Y}$  and the square root of their wavelet variance; i.e.,  $\rho_{XY}(\tau_j) \equiv \gamma_{XY}(\tau_j)/[v_X(\tau_j)v_Y(\tau_j)]$ . An unbiased estimator of the wavelet correlation using the MODWT is given by simply replacing the true wavelet covariance and wavelet variances by their unbiased estimators via

$$\tilde{\rho}_{XY}(\tau_j) \equiv \frac{\tilde{\gamma}_{XY}(\tau_j)}{\tilde{v}_X(\tau_j)\tilde{v}_Y(\tau_j)}.$$

### 2.5. Confidence intervals for the wavelet covariance and correlation

The asymptotic variance of the MODWT-based estimator of the wavelet covariance is given by

$$\lim_{N \rightarrow \infty} \tilde{N}_j \text{Var}\{\tilde{\gamma}_{XY}(\tau_j)\} = \int_{-1/2}^{1/2} S_{W_{j,x}}(f)S_{W_{j,y}}(f) df + \int_{-1/2}^{1/2} S_{W_{j,x},W_{j,y}}^2(f) df \equiv \Upsilon_j,$$

where  $S_{W_{j,x}}(f)$  and  $S_{W_{j,y}}(f)$  are the spectra for the scale  $\tau_j$  wavelet coefficients of  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively, and  $S_{W_{j,x},W_{j,y}}(f)$  is the cross spectrum between the two series of scale  $\tau_j$  wavelet coefficients. We can construct the random interval

$$\left[ \tilde{\gamma}_{XY}(\tau_j) - \Phi^{-1}(1-p) \left( \frac{\Upsilon_j}{\tilde{N}_j} \right)^{1/2}, \tilde{\gamma}_{XY}(\tau_j) + \Phi^{-1}(1-p) \left( \frac{\Upsilon_j}{\tilde{N}_j} \right)^{1/2} \right],$$

that captures the true wavelet covariance and forms a  $100(1-2p)\%$  confidence interval. In order to calculate an approximate confidence interval for Gaussian processes,  $\Upsilon_j$  can be estimated using the autocovariance and cross-covariance sequences of the scale  $\tau_j$  wavelet coefficients from  $\mathbf{X}$  and  $\mathbf{Y}$  which is similar to the construction of an approximate confidence interval for the wavelet variance.

<sup>8</sup> See Whitcher [21] and Whitcher et al. [22] for a more thorough introduction to the wavelet covariance and its application to observed multivariate time series.

We may relax the assumed Gaussianity in the two series of wavelet coefficients to obtain approximate confidence intervals applicable to non-Gaussian time series. Let  $\tilde{U}_{j,XY}$  be the vector of mean-corrected products of the wavelet coefficients, i.e.,  $\tilde{U}_{j,t,XY} = \tilde{W}_{j,t,X} \tilde{W}_{j,t,Y} - E\{\tilde{W}_{j,t,X} \tilde{W}_{j,t,Y}\}$ . Serroukh and Walden [23] showed that the asymptotic variance of  $\tilde{\gamma}_{XY}(\tau_j)$  is given by

$$\lim_{N \rightarrow \infty} \tilde{N}_j \text{Var}\{\tilde{\gamma}_{XY}(\tau_j)\} = S_{W_j^{(X)} W_j^{(Y)}} - \gamma_{XY(\tau_j)}(0) = S_{U_{j,XY}}(0),$$

and proposed a multitaper spectral estimator for  $S_{U_{j,XY}}(0)$  in order to produce an approximate confidence interval applicable to non-Gaussian processes.

We now turn our attention to the wavelet correlation. Given the inherent non-normality of the correlation coefficient for small sample sizes, a nonlinear transformation is sometimes required in order to construct a confidence interval. Let  $h(\rho) \equiv \tanh^{-1}(\rho)$  define Fisher’s  $z$ -transformation. For the estimated correlation coefficient  $\hat{\rho}$ , based on  $N$  independent samples,  $\sqrt{N-3}[h(\hat{\rho}) - h(\rho)]$  is approximately distributed as a Gaussian with mean zero and unit variance. The random interval

$$\left[ \tanh \left\{ h[\tilde{\rho}_{XY}(\lambda_j)] - \frac{\Phi^{-1}(1-p)}{\sqrt{\hat{N}_j - 3}} \right\}, \tanh \left\{ h[\tilde{\rho}_{XY}(\lambda_j)] + \frac{\Phi^{-1}(1-p)}{\sqrt{\hat{N}_j - 3}} \right\} \right]$$

captures the true wavelet correlation and provides an approximate  $100(1-2p)\%$  confidence interval. The quantity  $\hat{N}_j$  is the number of wavelet coefficients associated with scale  $\tau_j$  computed via the DWT – not the MODWT. This assumption of uncorrelated observations in order to use Fisher’s  $z$ -transformation is only valid if we believe no systematic trends or non-stationary features exist in the wavelet coefficients at each scale. The DWT is known to approximately decorrelate a wide range of power-law processes and thus provides a reasonable measure of the scale-dependent sample size.

Notice that the approximate confidence interval for the estimated wavelet correlation does not utilize any information regarding the distribution of the wavelet coefficients. Hence, no adjustment is made regarding the distribution of the incoming wavelet coefficients; they may be Gaussian or non-Gaussian.

### 3. Empirical findings

Before we examine the results, we provide some basic properties of the possible long-memory data generating processes which may underlie the foreign exchange markets. These data generating processes are the fractional Gaussian noise process, pure power law process and the fractionally differenced process.<sup>9</sup>

<sup>9</sup>An extensive discussion of long-memory processes and their properties can be found in Percival and Walden [18].



Table 1

The properties of the fractional Gaussian noise, pure power law and fractionally differenced processes for a range of  $H$ ,  $\alpha$  and  $d$  values.  $d$  is related to  $\alpha$  such that  $d = -\alpha/2$ , Ref. [18]

Process type	Nonstationary long memory process	Stationary long memory process	White noise process	Stationary short memory process
Fractional Gaussian noise	—	$0.5 < H < 1$	$H = 0.5$	$0 < H \leq 0.5$
Pure power law	$\alpha \leq -1$	$-1 < \alpha < 0$	$\alpha = 0$	$\alpha \geq 0$
Fractionally differenced	$d \geq 0.5$	$0 < d < 0.5$	$d = 0$	$d \leq 0$

Mandelbrot and Van Ness [24] studied the *fractional Gaussian noise* process whose spectral density function is given by

$$S(f) = 4\sigma_X^2 C_H \sin^2(\pi f) \sum_{j=-\infty}^{\infty} \frac{1}{|f + j|^{2H+1}}, \quad -0.5 \leq f \leq 0.5, \tag{4}$$

where  $\sigma_X^2$  is the variance of the process,  $C_H > 0$  and  $H$  is the so-called Hurst exponent. This process exhibits stationary long-memory dynamics when  $0.5 < H < 1$  and reduces to white noise when  $H = 0.5$ .

On the other hand, the spectral density of a *pure power-law process* is given by

$$S(f) = C_S |f|^\alpha, \quad -0.5 \leq f \leq 0.5, \tag{5}$$

where  $C_S > 0$  and  $\alpha$  is the scaling parameter. When  $-1 < \alpha < 0$ , this process has stationary long memory features. It is nonstationary for  $\alpha \leq -1$  and can be made stationary through differencing.

Another type of long-memory process is the *fractionally-differenced* type, which was introduced by Granger and Joyeux [25] and Hosking [26]. The fractionally differenced process is defined by  $(1 - L)^d X_t = \varepsilon_t$  where  $\varepsilon_t$  is a Gaussian white noise. The spectral density of this process is given by

$$S(f) = \frac{\sigma_\varepsilon^2}{(4 \sin^2(\pi f))^d}, \quad -0.5 \leq f \leq 0.5. \tag{6}$$

For  $0 < d < 0.5$ , the fractional difference process has the stationary long memory dynamics and becomes a white noise for  $d = 0$ . The link between  $H$ ,  $\alpha$  and  $d$  is summarized in Table 1.

The studied data sets are the five-minute Deutschemark–US dollar (DEM-USD) and Japanese Yen–US dollar (JPY-USD) price series for the period from December 1, 1986 to December 1, 1996.<sup>10</sup> Bid and ask prices at each 5 min interval are obtained by linear interpolation over time as in Müller et al. [5] and Dacorogna et al. [27]. Prices are computed as the average of the logarithm of the bid and ask prices

$$P_t = \frac{1}{2} [\log P(\text{bid})_t + \log P(\text{ask})_t], \quad t = 1, \dots, 751, 645. \tag{7}$$

Olsen and Associates applied data cleaning filters to the price series (as received from Reuters) in order to correct for data errors and to remove suspected outliers. We also

<sup>10</sup> The data set is provided by Olsen and Associates in Zürich, Switzerland.

removed the weekend quotes from Friday 21:05 GMT to Sunday 21:00 GMT. Apart from this, we did not apply any further filtering to the data set nor did we exclude any data points.<sup>11</sup> Continuously compounded 5-min returns are calculated as the log difference of the prices:

$$r_{t_5} = (\log P_t - \log P_{t-1}) \times 100 \text{ and } t_5 = 1, \dots, 751, 644. \quad (8)$$

It is often argued that price changes observed at very high frequencies can be overly biased by the buying and selling intentions and the quoting institutions [7]. We therefore decided to work with 20-min aggregated returns:<sup>12</sup>

$$r_{t_{20}} = \sum_{i=0}^3 r_{t_5-i}, \quad t_{20} = 1, \dots, 187, 911. \quad (9)$$

Therefore, our sample covers 2610 business days in 10 years with 72 observations per day.<sup>13</sup> The 20-min volatilities are defined to be the 20-min absolute returns,  $|r_{t,20}|$ .

We use  $|r_{t,20}|$  to obtain the MODWT decomposition of the variances of absolute return series on a scale by scale basis for both the DEM-USD and JPY-USD series. This study covers scales from 20 min to approximately one month. This coverage is achieved with a 12 level MODWT decomposition. The Daubechies least asymmetric family of wavelets ( $LA(8)$ ) was utilized in the MODWT (see Fig. 1). We have tried other wavelets such as Daubechies  $D(4)$  as well. The results are not sensitive to the choice of the wavelet family as long as the underlying process is stationary or an integer difference of the process is stationary.

The wavelet variance for each absolute return series and confidence intervals under the assumption of Gaussianity<sup>14</sup> is shown in Fig. 2, plotted on a log–log scale. Note that lower scales correspond to higher frequency bands. For example, the first scale is associated with 20 min changes, the second scale is associated with  $2 \times 20 = 40$  min changes and so on. The first six scales capture the frequencies  $1/28 \leq f \leq \frac{1}{2}$ ; i.e., oscillations with a period length of 2 (40 min) to 128 (2560 min). Since there are

<sup>11</sup> Andersen et al. [10] utilizes the same sample of DEM-USD and JPY-USD series. However, they removed weekends and several (mostly North American) holidays from the sample. They have also excluded the days containing “fifteen longest zero and constant runs”. Andersen and Bollerslev [28,29] analyzed a shorter sample of the same data set. They also removed the weekend quotes from their sample. See Bollerslev and Domowitz [30] for a detailed analysis of quote activity in the interbank market and a justification for the weekend definition above.

<sup>12</sup> We could have worked with the raw data set and simply ignored wavelet scales associated with higher frequencies. However, the aggregation greatly reduces the computational burden.

<sup>13</sup> The last day in the sample is a Friday. Since we removed the weekend quotes starting from Friday 21:05, there are only 63 20-min return observations for the last day in the sample. Therefore, the sample size is  $(72 \times 209) + 63 = 187,911$ .

<sup>14</sup> The alternative confidence intervals calculated by relaxing the Gaussianity assumption as suggested in Section 2.3. These do not differ significantly and are not plotted in Fig. 2.

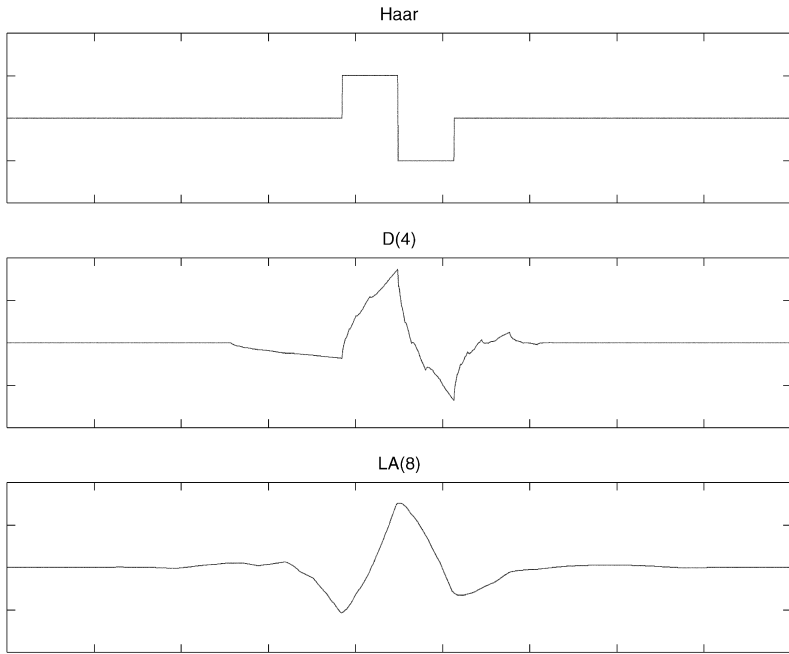


Fig. 1. Daubechies compactly supported wavelet functions: Haar wavelet based on two non-zero coefficients (also corresponds to the extremal phase wavelet based on two non-zero coefficients), the extremal phase wavelet based on four non-zero coefficients (D(4)) and the least asymmetric wavelet based on eight non-zero coefficients (LA(8)).

$72 \times 20 = 1440$  min in one day, we conclude that the first six scales are related with intra-day dynamics of our sample.

The seventh scale, where we observe an apparent break in the variance for both series, is associated with  $64 \times 20 = 1280$  min changes. Since there are 1440 min in one day, the seventh scale corresponds to 0.89 day. Therefore, the seventh and higher scales are taken to be related with one day and higher dynamics.

The relationship between wavelet variance and scale is given by  $v_{\chi}^2(\tau_j) \propto \tau_j^{-\alpha-1}$  where  $\alpha$  is the scaling parameter in a pure power law process. An estimate of  $\alpha$  is obtained from the ordinary least-squares (OLS) regression of  $\log v_{\chi}^2(\tau_j)$  on  $\log \tau_j^{-\alpha-1}$ . Fig. 3 plots the OLS fits of the sample points for two different regions. Estimated slopes,  $-(1 + \hat{\alpha})$ , for the smallest six scales are  $-0.48$  and  $-0.59$  for DEM-USD and JPY-USD series, respectively. This result implies that estimated scaling parameter is  $\hat{\alpha} = -0.52$  for the DEM-USD and  $\hat{\alpha} = -0.40$  for the JPY-USD volatilities for the first six scales (intra-day).

The estimated slopes for the higher scales (from the seventh to the twelfth) are  $-0.20$  and  $-0.16$  for the DEM-USD and JPY-USD volatilities, respectively. The wavelet variance estimators strongly indicate the presence of different scaling parameters for volatility at one day and higher scales where  $\hat{\alpha}$  are  $-0.80$  and  $-0.84$  for DEM-USD and JPY-USD volatilities, respectively.

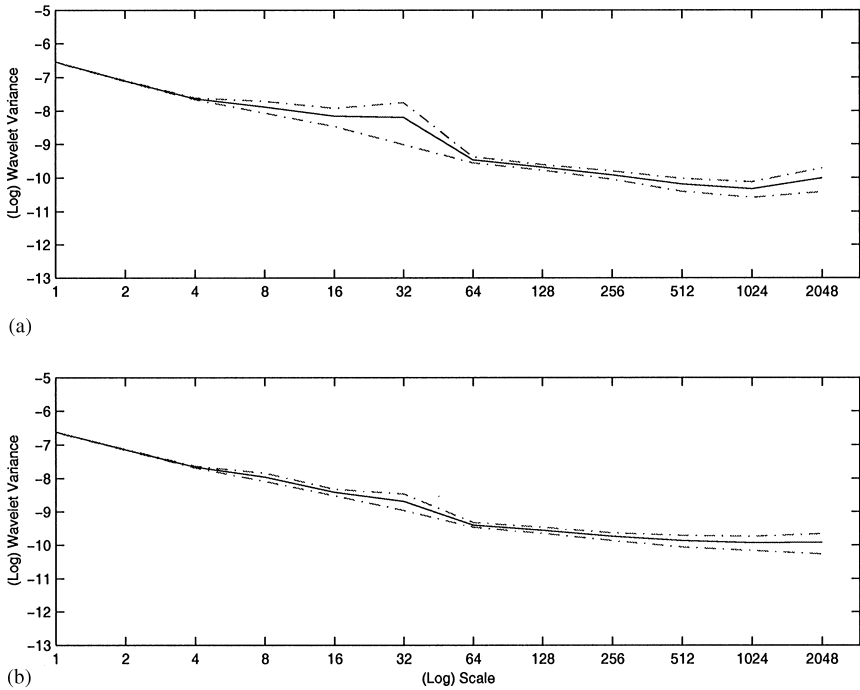


Fig. 2. Wavelet variance for 20-min volatilities of (a) DEM-USD and (b) JPY-USD from December 1, 1986 to December 1, 1996 at log–log scale. The straight line is the estimated variance and dashed lines are approximate 95% confidence intervals under the assumption of Gaussianity. The alternative confidence intervals calculated by relaxing the Gaussianity assumption as suggested in Section 2.3. These do not differ significantly and are not plotted here. Each scale is associated with a particular time period. For example the first scale is 20 min, the second scale is  $2 \times 20 = 40$  min, the third scale is  $4 \times 20 = 80$  min and so on. The seventh scale is  $64 \times 20 = 1280$  min. Since there are 1440 min per day, the seventh scale corresponds to approximately one day. The last scale is approximately 28 days.

As presented in Table 1, the relationship between  $d$  and  $\alpha$  is such that  $d = -\alpha/2$ . The results imply that the fractional integration parameters are  $d = 0.26$  for the DEM-USD and  $d = 0.20$  for the JPY-USD volatilities at the lower scales (intra-day). At higher scales (one day and more) the estimated  $d$  values for the DEM-USD and JPY-USD volatilities are 0.40 and 0.42 indicating less persistence in the intra-day horizon as compared to the volatilities of one day and higher scales. Note that a fractionally differenced process becomes more persistent as  $d$  takes larger positive values.

There are two major implications of these findings. The first is that the foreign exchange volatility is a stationary long-memory process whether it is governed by a pure power law or a fractionally differenced process. The second is that the foreign exchange volatility has a multi-scaling behavior. There are two different scaling parameters corresponding to intra-day and higher scales, respectively. This apparent multi-scaling behavior occurs approximately at the daily horizon where the intra-day seasonality is strongly present. The removal of the intra-day seasonality would not eliminate this

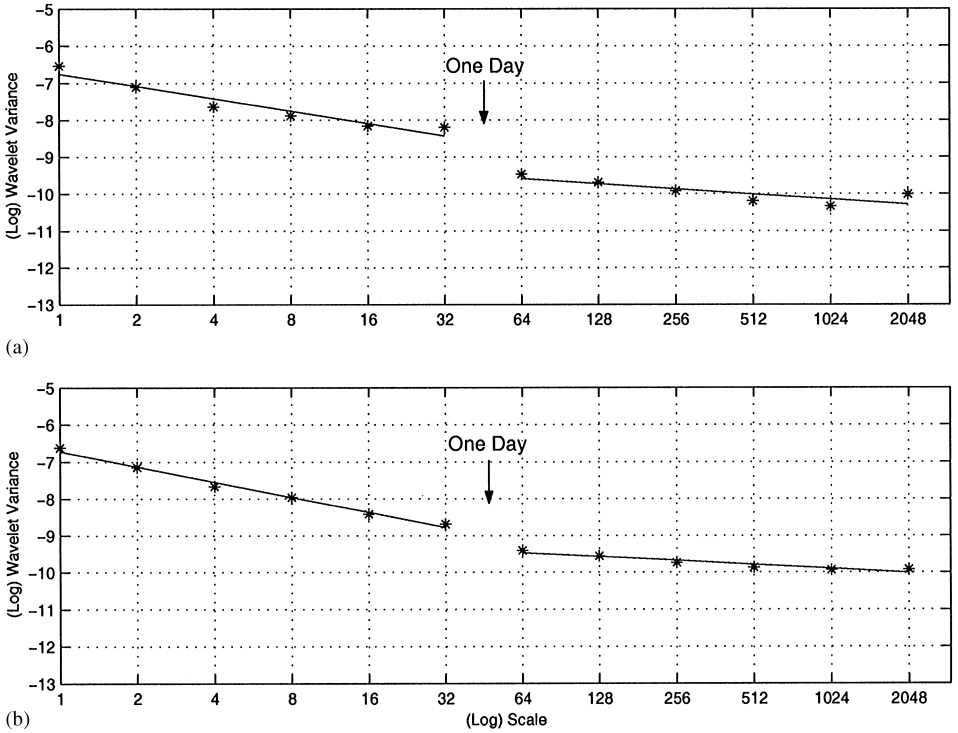


Fig. 3. Wavelet variance for 20-min volatilities of (a) DEM-USD and (b) JPY-USD from December 1, 1986 to December 1, 1996 at log–log scale. The \*’s are the estimated variances for each scale. The straight lines are OLS fits. Each scale is associated with a particular time period. For example the first scale is 20 min, the second scale is  $2 \times 20 = 40$  min, the third scale is  $4 \times 20 = 80$  min and so on. The seventh scale is  $64 \times 20 = 1280$  min. Since there are 1440 min per day, the seventh scale corresponds to approximately one day. The last scale is approximately 28 days.

multi-scaling but the transition between two scaling regions would be more gradual by exhibiting a concave scaling behavior (see Corsi et al. [31]).

The wavelet correlations between two volatilities are given in Fig. 4. The wavelet correlation is significantly different from zero for all scales. The correlation coefficient at the lowest scale (20 min) is  $0.415 \pm 0.005$ <sup>15</sup> and increases at higher scales. It reaches to a maximum at the sixth scale (640 min) with  $0.695 \pm 0.18$ . At large scales (one day and higher), the correlation coefficient between two series remains constant at around 0.70. The result is in contrast with the findings in Andersen et al. [10]. They found that volatility correlation between two exchange rates drops at longer horizons, suggesting that the benefits to currency diversification may be greatest over longer investment horizons. Our findings indicate that the benefits to currency diversification

<sup>15</sup> As pointed out in Section 2.5, the approximate confidence interval for the estimated wavelet correlation does not utilize any information regarding the distribution of the wavelet correlation. Therefore, these confidence intervals are robust to non-Gaussianity.

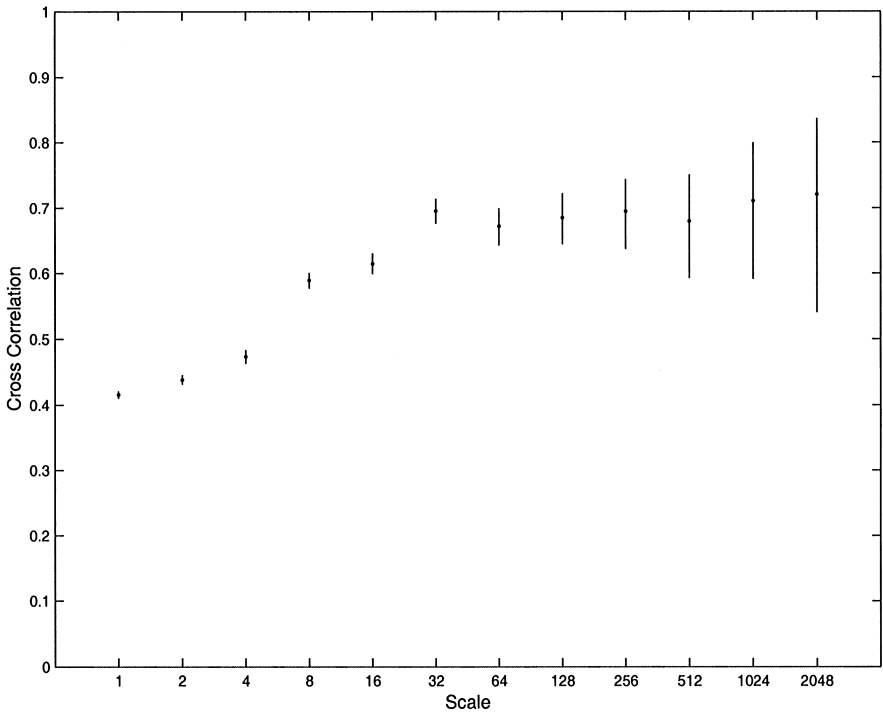


Fig. 4. Wavelet correlation for 20-min volatilities of DEM-USD and JPY-USD from December 1, 1986 to December 1, 1996. The vertical lines form an approximate 95% confidence intervals. As pointed out in Section 2.5, the approximate confidence interval for the estimated wavelet correlation does not utilize any information regarding the distribution of the wavelet correlation. Therefore, these confidence intervals are robust to non-Gaussianity. Each scale is associated with a particular time period. For example, the first scale is 20 min, the second scale is  $2 \times 20 = 40$  min, the third scale is  $4 \times 20 = 80$  min and so on. The seventh scale is  $64 \times 20 = 1280$  min. Since there are 1440 min per day, the seventh scale corresponds to approximately one day. The last scale is approximately 28 days.

are the greatest at the lowest scale and decreases with increasing scale during the day.

Fig. 5 plots the 20 min cross-correlations of two volatilities for 10 days where the strong intra-day seasonality is evident. As Gençay et al. [32] demonstrates, the presence of seasonalities in a long memory process obscures the underlying low frequency dynamics. With the wavelet analysis, we disentangle the cross-correlations on a scale by scale basis which enables us to determine which scales are contributing to the overall association between two volatilities. This is demonstrated in Fig. 6 which

Fig. 6. Wavelet cross-correlations for 20-min volatilities of DEM-USD and JPY-USD from December 1, 1986 to December 1, 1996. The dashed lines are the approximate confidence intervals. As pointed out in Section 2.5, the approximate confidence interval for the estimated wavelet correlation does not utilize any information regarding the distribution of the wavelet correlation. Therefore, these confidence intervals are robust to non-Gaussianity. Each scale is associated with a particular time period. For example the first scale is 20 min, the second scale is  $2 \times 20 = 40$  min, the third scale is  $4 \times 20 = 80$  min and so on. The sixth scale is  $32 \times 20 = 640$  min.

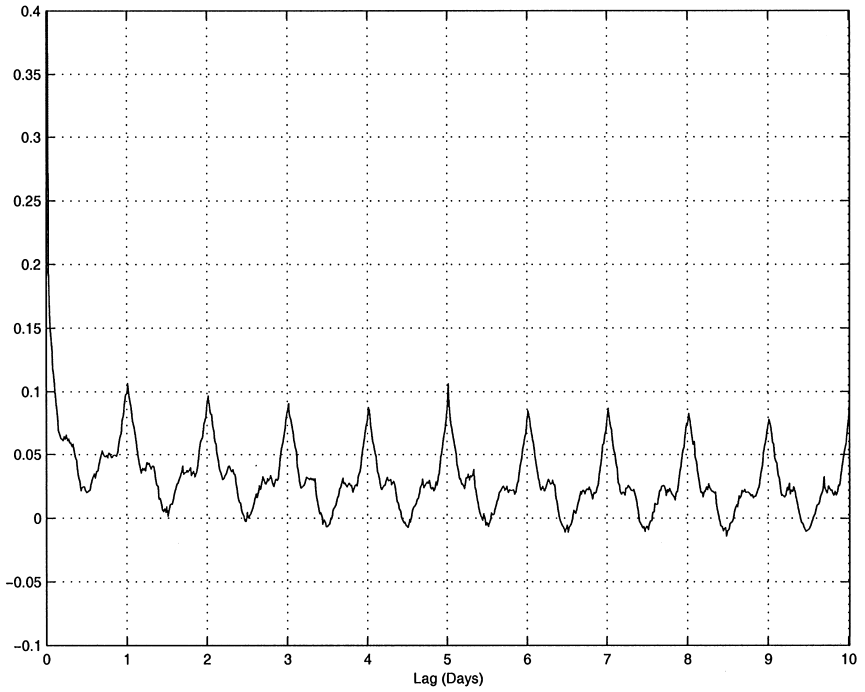
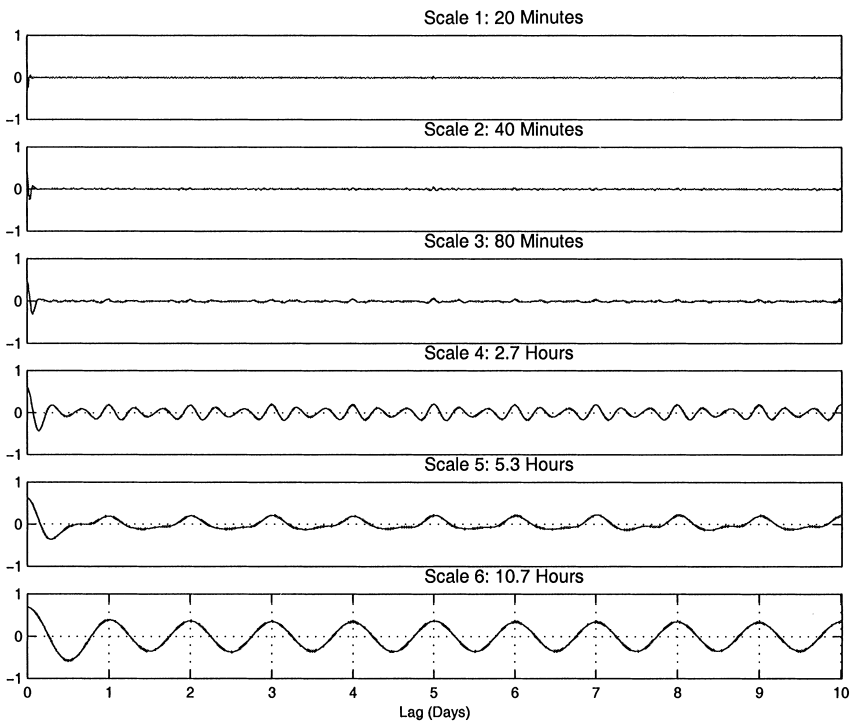


Fig. 5. Cross-correlations of 20-min volatilities for DEM-USD and JPY-USD from December 1, 1986 to December 1, 1996.



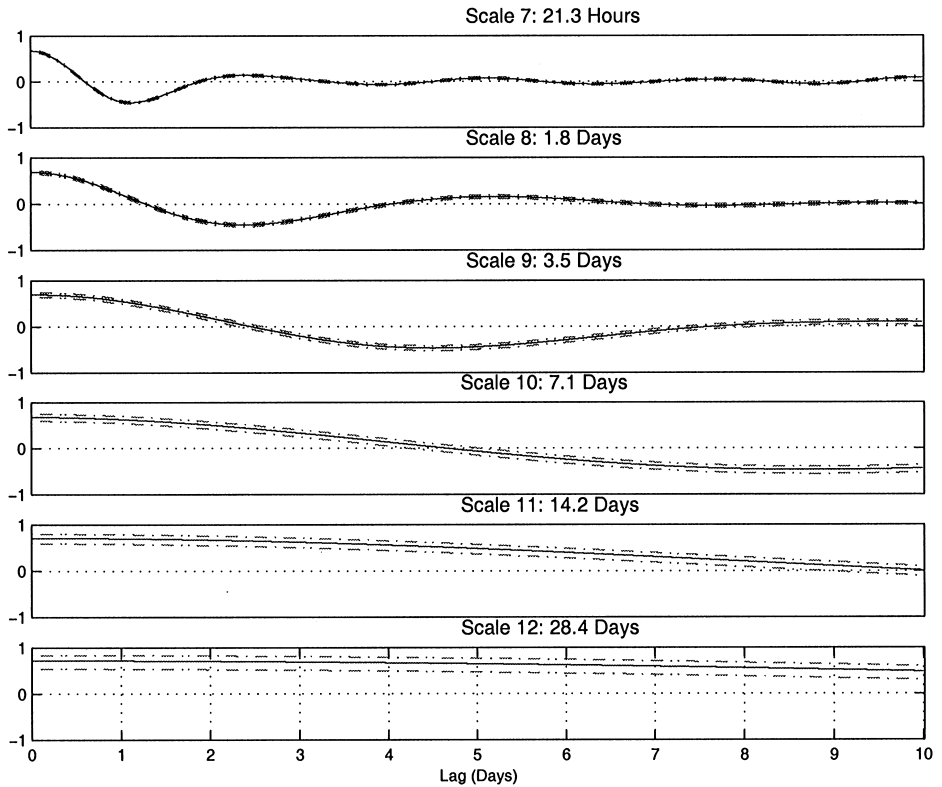


Fig. 7. Wavelet cross-correlations for 20-min volatilities of DEM-USD and JPY-USD from December 1, 1986 to December 1, 1996. The dashed lines are the approximate confidence intervals. As pointed out in Section 2.5., the approximate confidence interval for the estimated wavelet correlation does not utilize any information regarding the distribution of the wavelet correlation. Therefore, these confidence intervals are robust to non-Gaussianity. Each scale is associated with a particular time period. For example, the seventh scale is  $64 \times 20 = 1280$  min, the eight scale is  $128 \times 20 = 2560$  min and so on. The last scale is approximately 28 days.

provides wavelet cross-correlations<sup>16</sup> and the corresponding approximate confidence intervals<sup>17</sup> for the first six scales for 10 days. For scales 1–3 (from 20 to 80 min), there is a small cross-correlation corresponding to intra-day horizon after which there is no significant cross-correlation persistence. At scales 4–6, a seasonal cross-correlation gradually evolves and a daily seasonality dominates in the sixth scale ( $\approx 10.7$  h). In Fig. 7, the cross-correlations of the scales 7–12 are presented. At these higher scales, a persistent cross-correlation behavior becomes more visible. At the highest scale ( $\approx 28.4$  days), the rate of the decay in the cross-correlations is more in line with a hyperbolic decay than an exponential decay.

<sup>16</sup> The cross-correlations are roughly symmetric about zero. Therefore, the negative lags are not plotted to conserve space.

<sup>17</sup> As pointed out in Section 2.5, the approximate confidence interval for the estimated wavelet correlation does not utilize any information regarding the distribution of the wavelet correlation. Therefore, these confidence intervals are robust to non-Gaussianity.



## 4. Conclusions

We have proposed a simple method for identifying the scaling laws in financial time series. The proposed methodology is based on a wavelet multi-scaling approach which decomposes the variance of a time series and the covariance between two time series on scale by scale basis through the application of a non-decimated discrete wavelet transformation. It is simple to calculate and can easily be implemented as it does not depend on a particular model selection criterion and model specific parameter choices.

It is shown that exchange rate volatility has different scaling properties at different horizons. The intra-day persistence in volatility is significantly less than the volatility at one day or higher scales. The correlation between two volatility series increases within the day but remains constant at one day or higher scales. The wavelet cross-correlation analysis indicates that the association between two volatilities is stronger at higher scales (low frequencies).

## Acknowledgements

Ramazan Gençay gratefully acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada. We are grateful to Olsen & Associates for providing the data used in this paper. We would like to thank Abdurrahman Ulugulyağci for his research assistance.

## References

- [1] B.B. Mandelbrot, The variation of certain speculative prices, *J. Bus.* 36 (1963) 394–419.
- [2] P.K. Clark, A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* 41 (1973) 135–155.
- [3] R.N. Mantegna, H.E. Stanley, Scaling behavior in the dynamics of an economic index, *Nature* 376 (1995) 46–49.
- [4] F.X. Diebold, A. Hickman, A. Inoue, T. Schuermann, Scale models, *Risk* 11 (1998) 104–107.
- [5] U.A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, C. Morgenegg, Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis, *J. Banking Finance* 14 (1990) 1189–1208.
- [6] U.A. Müller, M.M. Dacorogna, R.D. Dave, O. Pictet, R.B. Ward, Fractals and intrinsic time, a challenge to econometricians, Olsen & Associates Discussion Paper, 1995.
- [7] D.M. Guillaume, M.M. Dacorogna, R.D. Dave, U.A. Müller, R.B. Olsen, O.V. Pictet, From the bird's eye to the microscope: a survey of new stylized facts of the Intra-daily Foreign Exchange Markets, *Finance Stochastics* 1 (1997) 95–129.
- [8] R.N. Mantegna, H.E. Stanley, Turbulence and financial markets, *Nature* 383 (1996) 587–588.
- [9] S. Ghoshghaie, W. Brewmann, J. Peinke, P. Talkner, Y. Dodge, Turbulent cascades in foreign exchange markets, *Nature* 381 (1996) 767–770.
- [10] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association*, forthcoming.
- [11] B. LeBaron, Volatility persistence and apparent scaling laws in finance, Brandeis University Graduate School of International Studies, Discussion Paper, 1999.
- [12] O.E. Barndorff-Nielsen, K. Prause, Apparent scaling, *Finance and Stochastics*. (2000), to appear.

- [13] M. Pasquini, M. Serva, Multiscale behaviour of volatility autocorrelations in a financial market, *Econ. Lett.* (1999) 275–279.
- [14] F.X. Diebold, M. Nerlove, The dynamics of exchange rate volatility: a multivariate latent-factor ARCH model, *J. Appl. Econometrics* 4 (1989) 1–22.
- [15] T. Bollerslev, R.F. Engle, Common persistence in conditional variances, *Econometrica* 61 (1993) 166–187.
- [16] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [17] P. Abry, D. Veitch, P. Flandrin, Long-range dependence: revisiting aggregation with wavelets, *J. Time Ser. Anal.* 19 (1998) 253–266.
- [18] D.B. Percival, A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, 2000.
- [19] D.B. Percival, On estimation of the wavelet variance, *Biometrika* 82 (1995) 619–631.
- [20] A. Serroukh, A.T. Walden, D.B. Percival, Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series, *J. Am. Stat. Assoc.* 95 (2000) 184–196.
- [21] B. Whitcher, *Assessing nonstationary time series using wavelets*, Ph.D. Thesis, University of Washington, 1998.
- [22] B. Whitcher, P. Guttorp, D.B. Percival, Wavelet analysis of covariance with application to atmospheric time series, *J. Geophys. Res. – Atmos.* 105 (2000) 14941–14962.
- [23] A. Serroukh, A.T. Walden, Wavelet scale analysis of bivariate time series II: statistical properties for linear processes, *J. Nonparametric Stat.* (2000), in press.
- [24] B.B. Mandelbrot, J. Van Ness, Fractional Brownian motion, fractional noises and applications, *SIAM Rev.* 10 (1968) 422–437.
- [25] C. Granger, R. Joyeux, An introduction to long memory time series models and fractional differencing, *J. Time Ser. Anal.* 1 (1980) 15–29.
- [26] J. Hosking, Fractional differencing, *Biometrika* 68 (1981) 165–176.
- [27] M.M. Dacorogna, U.A. Müller, R.J. Nagler, R.B. Olsen, O.V. Pictet, A geographical model for the daily and weekly seasonal volatility in the foreign exchange markets, *J. Int. Money Finance* 12 (1993) 413–438.
- [28] T.G. Andersen, T. Bollerslev, Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns, *J. Finance* 52 (1997) 975–1005.
- [29] T.G. Andersen, T. Bollerslev, DM-Dollar volatility: intraday activity patterns, macroeconomic announcements, and longer-run dependencies, *J. Finance* 53 (1998) 219–265.
- [30] T. Bollerslev, I. Domowitz, Trading patterns and prices in the interbank foreign exchange market, *J. Finance* 48 (1993) 1421–1443.
- [31] F. Corsi, M. Dacorogna, U. Müller, G. Zumbach, High frequency data do improve volatility and risk estimation, *Olsen & Associated Discussion Paper*, Zurich, Switzerland, 2000.
- [32] R. Gençay, F. Selçuk, B. Whitcher, Differentiating intraday seasonalities through wavelet multi-scaling, manuscript, 1999.