European University Institute

DEPARTMENT OF ECONOMICS

EUI Working Papers

ECO 2011/23
DEPARTMENT OF ECONOMICS

THE MULTISCALE CAUSAL DYNAMICS OF FOREIGN EXCHANGE MARKETS

Stelios Bekiros and Massimiliano Marcellino

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

The Multiscale Causal Dynamics of Foreign Exchange Markets

STELIOS BEKIROS

and

MASSIMILIANO MARCELLINO

EUI Working Paper **ECO** 2011/23

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

THE MULTISCALE CAUSAL DYNAMICS OF FOREIGN EXCHANGE MARKETS *

Stelios Bekiros † Massimiliano Marcellino ‡

European University Institute, Department of Economics,
Via della Piazzuola 43, I-50133 Florence, Italy

ABSTRACT

This paper relies on wavelet multiresolution analysis to capture the dependence structure of currency markets and reveal the complex dynamics across different timescales. It investigates the nature and direction of causal relationships among the most widely traded currencies denoted relative to the United States Dollar (USD), namely Euro (EUR), Great Britain Pound (GBP) and Japanese Yen (JPY). The timescale analysis involves the estimation of linear vis-à-vis nonlinear and spectral causality of wavelet components and aggregate series as well as the detection of short- vs. long-run linkages and cross-scale correlations. Moreover, this study attempts to probe into the micro-foundations of across-scale heterogeneity in the causality pattern on the basis of trader behavior with different time horizons. New stylized properties emerge in the volatility structure and the implications for the flow of information across scales are inferred. The examined period starts from the introduction of the Euro and covers the dot-com bubble, the financial crisis of 2007-2010 and the Eurozone debt crisis. Technically, this paper presents an invariant discrete wavelet transform that deals efficiently with phase shifts, dyadic-length and boundary effects. It also proposes a new entropy-based methodology for the determination of the optimal decomposition level. Overall, there is no indication of a global causal behavior that dominates at all timescales. When the nonlinear effects are accounted for, the evidence of dynamical bidirectional causality implies that the pattern of leads and lags changes over time. These results may prove useful to quantify the process of integration as well as influence the greater predictability of currency markets.

Keywords: exchange rates; wavelets; timescale analysis; causality; entropy

JEL classification: C14 ; C32 ; C51 ; F31

* We are grateful to Ramazan Gençay and Helmut Lütkepohl as well as seminar participants at various institutions for helpful comments and discussions. Stelios Bekiros also thanks the Department of Quantitative Economics at the University of Amsterdam (UvA), the Department of Economics and the Max Weber Programme at the European University Institute (EUI), for having hosted his research. Earlier versions of this paper were presented at the UvA, the Netherlands Society for Statistics and Operations Research (VVS) and at the EUI. This research is supported by the Marie Curie Intra European Fellowship (FP7-PEOPLE-2009-IEF, N° 251877) under the 7th European Community Framework Programme. The usual disclaimers apply.

† Corresponding author. *Tel.*: +39 055 4685916; *Fax*: +39 055 4685 902; *E-mail address*: stelios.bekiros@eui.eu

‡ *Tel.:* +39 055 4685956; *Fax:* +39 055 4685 902; *E-mail address*: massimiliano.marcellino@eui.eu

1. INTRODUCTION

Since the pioneering work of Grossman and Morlet (1984) wavelet methodology, a refinement of Fourier analysis, has been introduced in the literature as an alternative for analysing nonstationary data with "irregularities". Their contribution was followed by the development of multiresolution analysis by Mallat (1989) and the introduction of orthogonal wavelet bases by Daubechies (1992). Even though the wavelet methodology has widespread application in the natural sciences, it is a rather unexplored area in economics, with the exception of financial applications[4]. The wavelet multiscale decomposition, allowing for simultaneous analysis in the time and frequency domain, is a valuable means of exploring the complex dynamics of financial time series, which are inherently characterized by chaotic patterns, fat tails and long-memory, particularly at high sampling frequencies.

In this study we use the wavelet methodology to investigate the nature and direction of causality among foreign exchange (FX) markets at different timescales. During the Great Moderation period[5] and in particular during the nineties, currency markets have grown more similar and FX rate volatility decreased (Laopodis, 1998)[6]. More recently, the Euro behavior against the US dollar has seriously altered the prior state of market interrelations (Bénassy-Quéré *et al.*, 2000; Wang *et al.*, 2007). Given the status of the US dollar and Euro as anchor currencies, it is interesting to examine the nature of the causal linkages between them, as well as with other currencies[7]. The existence of causal linkages would suggest that news originating in a specific market is not country-specific and idiosyncratic, but efficiently transmitted to other foreign markets, thus providing support to the "*meteor shower*" notion introduced by Engle *et al.* (1990).

---

[4] See for example Greenblatt (1998), Jensen (2000), Davidson *et al*. (1998) and Fernandez (2005). A more detailed literature overview is provided in Appendix I.

[5] According to Stock and Watson (2003) the Great Moderation period initiated around the mid-1980s and lasted until the beginning of the 2000s. During that period, the growth variance of the G7 countries was considerably lower, from 50% to 80% in comparison to the pre- and the post –Great Moderation period.

[6] A rich empirical literature exists on the volatility spillover mechanism of the USD across other currency markets. The nature of the transmission mechanism as well as the degree of price information efficiency was already investigated during the 1980s in the beginning of the higher integration of FX rates vis-à-vis the USD (Hogan and Sharpe, 1984; Ito and Roley, 1987). Additional empirical evidence by Koutmos and Booth (1995) and Laopodis (1997) suggests that the size or sign of an innovation in USD, in response to a variation in the Federal Reserve interest rate, may significantly affect the extent of dependence across markets.

[7] The transactions involving USD-Euro amount approximately to 40% of global trading (BIS Triennial Survey, 2007).

However, it is necessary to go beyond linearity when examining the exchange rate linkages. Meese and Rogoff (1983) reported in their seminal work the failure of linear exchange rate models, and several more recent studies have provided further evidence against linearity. According to Ma and Kanas (2000) nonlinear structures may account for bubbles with self-fulfilling expectations (Blanchard and Watson, 1982), target zone models (Krugman, 1991), nonlinear monetary policies (Flood and Isard, 1989) and noise trading (Black, 1986). However, empirical studies that tested for these kinds of nonlinearities have rather failed to support them (Lindberg and Soderlind, 1994). Hence, we need to resort to a more general framework, provided by wavelet analysis.

The aim of our paper is to test for the existence of linear, nonlinear and spectral causal relationships among the three most heavily traded currencies ("FX majors") denoted relative to the United States dollar (USD), namely the Euro (EUR), Great Britain Pound (GBP) and Japanese Yen (JPY)[8]. This is implemented via the use of the wavelet methodology, which reveals the inherent dynamics across different timescales. The nature and direction of causality is investigated for each component of the time series as resulting from the wavelet analysis, and it is compared against the causality results obtained with the original "aggregate" series. The "palette" of tests utilized covers the major types of causality reported in the literature, namely the linear Granger test (Granger, 1969) the Baek and Brock (1992) parametric causality test for nonlinear dynamic causality and the frequency-domain test by Breitung and Candelon (2006). The investigated time period covers diverse regimes and "extreme" events including the rise and fall of the tech-market bubble and the financial crisis of 2007-2010. It also includes the EU debt crisis in early 2010, associated with the widening of bond yield spreads and the rise of credit default swaps, concerning Eurozone countries such as Greece, Ireland, and Portugal. This crisis had a significant effect primarily on the USD-EUR but also on the USD-GBP rate, partly due to the high UK trade deficit and debt. An attempt is made to shed light on the impact of these events on FX market linkages.

[8] The prime motivation for choosing these particular exchange rates comes from them being the most liquid and widely traded currency pairs in the world. On the spot market, according to the Bank of International Settlements (BIS, 2007), the USD was involved in 86.3% of transactions, followed by the EUR (37.0%), the JPY (17.0%) and the GBP (15.0%). Volume percentages for all individual currencies should add up to 200%, as each transaction involves two currencies..

The paper contributes to the literature in various ways. First, rather than using several datasets at different temporal frequencies, we rely on the wavelet decomposition of daily data to analyze the dependence structure of the FX markets at different time scales. The results provide evidence of complex heterogeneous dynamics across and within different scales and strongly indicate that the interactions between FX markets have different characteristics at different horizons.

Second, we attempt to probe into the micro-foundations of the detected heterogeneity across time scales, on the basis of trader behavior. We focus on the impact of the actions of long- and short-term traders, as it is plausible that they may have different time horizons for different trading decisions. Once the causality structure is identified from low-to-high frequency, this has implications for the flow of information across timescales. In one sense, this idea is just a generalization of Friedman's original concept as mentioned in Ramsey and Lampart (1998b)[9]. The propagation characteristics of the heterogeneous-driven behavior are investigated by studying the statistical properties of the information flow across scales.

Third, we introduce new practical guidelines in wavelet implementation methodology. Specifically, we propose 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 and deals effectively with "boundary effects". In addition, beyond the existing practice that has utilized either economic rationale or subjective judgement in considering the appropriate "depth" of the wavelet analysis, we introduce a new entropy-based methodology to determine the optimal level of decomposition. Overall, we show that our new approach unveils the dependence structure of FX markets in succinctly capturing the non-Gaussian dynamic features of currency rates by simultaneously modeling multiple time horizons.

The paper develops as follows: section 2 briefly describes the wavelet methodology and elaborates on our proposed shift-invariant discrete wavelet transform. It also proposes new practical guidelines in wavelet implementation. Section 3 describes the data and provides a

[9] Friedman's idea essentially was to propose a new concept of "horizon" to reconcile the short- and long-run empirical results on the permanent-income hypothesis (Friedman, 1963).

preliminary statistical analysis. Section 4 presents the empirical results of the wavelet analysis of the foreign exchange rates. Section 5 reviews causality testing and studies causality among both the aggregate and wavelet-disaggregated FX returns and volatility series. Finally, section 6 summarizes and concludes. A set of Appendices provide more technical material.

2. WAVELET MULTISCALE ANALYSIS

The multiresolution features of the wavelet decomposition can be useful in econometric analysis. Often, in financial and macroeconomic applications, the main focus is on the long-run equilibrium relationships and their interaction. Through wavelet decomposition, the low-frequency content of the data that "captures" the relevant long-run interactions can be extracted, and the high-frequency fluctuations that distort the underlying market dependencies can be removed.

Among the plethora of useful properties of wavelets, three other major facets should be highlighted namely, the ability to handle nonstationarities, the localization in time, and the decomposition in various timescales[10]. The evaluation of the effects of time scaling on the relationships among economic variables is also crucial to the present study. In the following subsection we introduce an invariant discrete wavelet transform. After that, we suggest new practical guidelines for wavelet implementation, and expand the literature that has utilized subjective reasoning in estimating the appropriate "depth" of the resolution, by proposing a new entropy-based methodology for the determination of the optimal decomposition level. An outline of necessary introductory concepts in wavelet analysis is provided in the technical Appendix I.

2.1 The Shift-Invariant Discrete Wavelet Transform (SIDWT)

The classical, decimated Discrete Wavelet Transform (DWT) involves subsampling of the output of the high- and low-pass filters h and g, to half their original length[11]. This leads to a

[10] The nonstationarity describes a broader notion than merely the existence of a unit-root, such as time variation, structural breaks, singularities as well as locally temporal effects.

[11] The $h = \left(h_0, ..., h_{M-1}\right)$ finite length wavelet filter and the $g = \left(g_0, ..., g_{M-1}\right)$ complement low-pass (scaling) filter satisfy the quadrature mirror relationship $g_m = \left(-1\right)^{m+1} h_{M-1-m}$ for $m = 0, ..., M-1$ (see technical Appendix I for details)

serious drawback, namely the transform is not invariant in the *real-axis*. Specifically, the DWT of a shifted signal is not the shifted version of the DWT of the signal[12]. Alternatively, an undecimated DWT can be implemented without the subsampling technique[13]. According to Coifman and Donoho (1995) undecimated versions of the DWT present some advantages compared to DWT. Primarily, they can handle any sample size T, while the J-th order DWT restricts the sample size to a multiple of 2^J. In addition, since the coefficients of an undecimated DWT are associated with zero phase filters, the original time series is properly aligned with the wavelet components. Moreover, they are invariant to circularly shifting the time series, a property that does not hold for the DWT. Finally, the undecimated wavelet variance estimator is asymptotically more efficient than the DWT estimator (Percival, 1995). In this study, we propose a new variation of the undecimated DWT, namely the Shift-Invariant DWT (SIDWT).

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 $\mathbf{w} = \mathbf{W}\mathbf{y}$ according to the DWT (see Appendix I). Formally, the SIDWT is defined as follows: 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

$$\tilde{\mathbf{w}} = \left[\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, ..., \tilde{\mathbf{w}}_J, \tilde{\mathbf{s}}_J\right]^T \qquad (1)$$

where $\tilde{\mathbf{w}}_j$ 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 $t = 0, ..., T/2^j - 1$, and the scaling coefficients

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

[13] Some undecimated versions of the DWT are encountered in the statistical literature, such as the "maximal overlap DWT" (Percival and Mofjeld, 1997; Allan, 1966) and the "stationary DWT" (Nason and Silverman, 1995).

$s_{J,t} = 2^{J/2}\tilde{s}_{J,2^J(t+1)-1} \quad t = 0,...,T/2^J - 1$. 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 $\tilde{\mathbf{W}} = \left[\tilde{\mathbf{W}}_1 \ \tilde{\mathbf{W}}_2 \ ... \tilde{\mathbf{W}}_J \ \tilde{\mathbf{S}}_J\right]^T$. 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 $\left(j = 1,...,J\right)$. 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., $\tilde{\mathbf{W}}_1 = \left[\tilde{\mathbf{h}}_1^{(1)}, \tilde{\mathbf{h}}_1^{(2)}, \tilde{\mathbf{h}}_1^{(3)}, ..., \tilde{\mathbf{h}}_1^{(T-2)}, \tilde{\mathbf{h}}_1^{(T-1)}, \tilde{\mathbf{h}}_1\right]^T$ and it can be interpreted as the circularly shifted version of DWT submatrix $\mathbf{W}_1$. The other matrices $\tilde{\mathbf{W}}_2, ..., \tilde{\mathbf{W}}_J$ are similarly constructed through replacing $\tilde{\mathbf{h}}_1$ by $\tilde{\mathbf{h}}_j$.

The SIDWT implementation algorithm starts with the data y_t, that is no longer limited to dyadic length as opposed to the classical "pyramid algorithm" introduced by Mallat (1989), 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[14]. 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 $t = 0,1,...,T-1$. 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., $\tilde{w}_{2,t} = \sum_{m=0}^{M-1} \tilde{h}_m \tilde{s}_{1,t-m \bmod T}$ and $\tilde{s}_{2,t} = \sum_{m=0}^{M-1} \tilde{g}_m \tilde{s}_{1,t-m \bmod T}$ with $t = 0,1,...,T-1$. 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 $J = \log_2\left(T\right)$ times in order to provide the full vector of SIDWT coefficients in Eq. (1). In the Inverse transform the final-level wavelet and scaling coefficients are convolved with their respective filters and the resulting vectors

[14] See Appendix I for a detailed description of Mallat's algorithm.

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 $t = 0,1,...,T-1$. 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 produces 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 $t = 0,1,...,T-1$.

The SIDWT, in the same way as the classical DWT, results in the additive decomposition of the time series. Let $\tilde{\mathbf{D}}_j = \tilde{\mathbf{W}}_j^T \tilde{\mathbf{w}}_j$ define the *wavelet detail* for the SIDWT corresponding to changes in the time series $\mathbf{y}$ at scale a_j for the level $j = 1,...,J$. The multiscale decomposition is defined for each observation y_t as the linear combination of wavelet SIDWT detail coefficients, i.e., $y_t = \sum_{j=1}^{J+1} \tilde{D}_{j,t}, \quad t = 0,...,T-1$, where $\tilde{D}_{j,t}$ is the t-th element of $\tilde{\mathbf{D}}_j$ for $j = 1,...,J$. Similarly, $\tilde{\mathbf{A}}_j = \sum_{k=j+1}^{J+1} \tilde{\mathbf{D}}_k$ is the cumulative sum of the variations of the details and is defined as the j-th level SIDWT *wavelet approximation* for $0 \le j \le J$ with $\tilde{\mathbf{A}}_{J+1}$ being a vector of zeros. The j-th level *wavelet rough* $\tilde{\mathbf{R}}_j = \sum_{k=1}^{j} \tilde{\mathbf{D}}_k$, $1 \le j \le J+1$ incorporates the remaining lower-scale details. Overall, the vector of observations may be decomposed as

$$\mathbf{y} = \tilde{\mathbf{A}}_j + \sum_{k=1}^{j} \tilde{\mathbf{D}}_j = \tilde{\mathbf{A}}_j + \tilde{\mathbf{R}}_j \quad (2)$$

It is emphasized that the SIDWT associates the wavelet coefficients with zero-phase filters, thus the details and approximations correspond directly to the original sample in perfect alignment. Percival and Mofjeld (1997) proved that undecimated transforms are energy (variance) preserving transforms. Thus, SIDWT is an efficient transform and the total variance of the time series is given by $\|\mathbf{y}\|^2 = \sum_{j=1}^{J} \|\tilde{\mathbf{w}}_j\|^2 + \|\tilde{\mathbf{s}}_J\|^2$.

2.2 New practical guidelines in wavelet implementation

It was shown that SIDWT is time-invariant as opposed to the classical DWT which exhibits some translation in time even after applying a "signal extension" process[15]. Furthermore, 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. Because the SIDWT entails zero-phase filtering, the details and approximation at each timescale contain the same number of observations and line up in time with the original series. This property makes SIDWT a particularly useful tool in the analysis of time-dependent processes.

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. If the spectral density is dynamic, longer filters should be employed in order to distinguish the frequency activity between scales. Optimally, in most data sets a balance between frequency localization and time localization should be pursued. According to Gençay *et al.* (2001) and Gençay *et al.* (2002), a moderate length wavelet filter (e.g., length eight) adequately captures the stylized features of financial data. Moreover, in case the wavelet filter bears no "similarity" with the underlying features, then the decomposition will be quite inefficient. Given that the wavelet basis functions are used to represent the information contained in the time series, they should "mimic" its underlying features[16]. Usually, smoothness and (a)symmetry are the most crucial factors in selecting suitable wavelet basis functions (Gençay *et al.*, 2002; Ramsey and Lampart, 1998b). The SIDWT coefficients in this study are calculated from the Daubechies family of compactly supported wavelet filters, which are well localized in time (Daubechies, 1992). Specifically, the Daubechies filter of length eight, (db8) is selected in order to balance smoothness, length and symmetry (Jensen and Whitcher, 2000; Gençay *et al.*, 2001). This is a widely used wavelet and is applicable in a wide variety of data types. It achieves an "ideal compromise"

[15] In order to deal with time series of non-dyadic length, a "signal extension" process is usually employed for DWT, 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, using higher-order polynomials, periodic extension, and numerical integration.

[16] 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 (see Appendix I for a description of Haar and Daubechies filter families).

between competing requirements in that it has reasonably narrow support, is fairly smooth, is twice differential, nearly symmetric and has a moderate degree of flexibility[17].

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[18]. To deal with boundary effects in this study, SIDWT employs a specialized "periodic extension" pattern. Specifically, 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 (Pesquet *et al.*, 1996). The extension mode used for the inverse SIDWT is the same to ensure a perfect reconstruction. In addition, using these boundary coefficients, the SIDWT retains its numerically stability (Herley, 1995).

Finally, in the literature the depth (level) of the multiscale wavelet decomposition is usually determined arbitrarily or based on some subjective (economic) rationale with regard to the examined time scales. Alternatively in this study the optimal level of multiscale decomposition is pursued with respect to the *minimization* of the Shannon entropy-related criterion[19]. It is estimated on the basis of the sample length, the selected wavelet class and the boundary-distortion method. 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 (Coifman and Wickerhauser, 1992). The optimal level is determined at the minimum value of the entropy-related criterion[20]. In the following expressions y is the signal and c_i represents the details and the j-th level approximation

[17] Alternative choices of wavelet classes were also applied in the empirical study, but the results were very robust to such changes and the current selection appeared to be the most balanced.

[18] Although various theoretical methods are available to tackle with this issue, they are rather inefficient from a practical viewpoint (Cohen *et al.*, 1993). A common technique applied in Fourier analysis involves the entire series to be duplicated around the last obs. This may be reasonable for some series with strong seasonal effects, but cannot be applied in general (Strang and Nguyen, 1996).

[19] Classical entropy-based criteria describe information-relevant properties for an accurate representation of a given signal.

[20] The Shannon 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 "translation" of the wavelet scales into economic time horizons, as mentioned in the empirical section.

coefficient of y for scales $j = 1,\ldots,J$ in an orthonormal basis[21]. The entropy *E* must be an additive function such that $E(0) = 0$ and $E(y) = \sum_j E(c_j)$. The Shannon entropy for the coefficients in each level is defined as

$$E_{Shannon}(c_j) = -c_j^2 \cdot \log(c_j^2) \qquad (3)$$

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

3. DATA DESCRIPTION AND PRELIMINARY ANALYSIS

The data comprise three time series of daily closing currency rates denoted relative to United States dollar (USD), namely Euro (EUR), Great Britain Pound (GBP) and Japanese Yen (JPY). The exact ratios represent EUR/USD, GBP/USD and USD/JPY[22]. The foreign exchange returns are defined as $r_t = \log(P_t) - \log(P_{t-1})$, where P_t is the closing level on day t, while the volatility series is defined as the absolute value of the returns $u_t = |r_t|$ as in Jensen and Whitcher (2000) and Gencay *et al.* (2002) . The data span a time period from January 5, 1999 to May 10, 2010 (2960 observations), namely from the introduction of the Euro until the EU ministers and the ECB agreed on a program of bond purchases and an unprecedented defence package of 750 billion€, with the contribution of the IMF, in order to deal with the 2010 Eurozone sovereign-debt crisis. Moreover, the robustness of the results is examined in sub-periods based on economic rationale as well as identified by the application of stability tests for structural breakpoints. Specifically, three breakpoints are initially considered for the identification of the sub-periods, hence setting a platform for departure for causality tests. The first structural break is *March 10, 2000* and corresponds to the date when the dot-com bubble "burst" (Greenspan, 2007). On that day 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. Moreover, the financial crisis of 2007-2010 is examined. It was

[21] Based on the wavelet decomposition, the reconstructed signal comprises the j-th level wavelet approximation and the details in all levels. Consequently, these are used to estimate the Shannon entropy criterion.

[22] These are the most liquid and widely traded currency pairs in the world ("FX majors") with 27% market turnover share for EUR/USD, 13% for USD/JPY and 12% for GBP/USD (BIS Triennial Survey, 2007).

triggered by a liquidity shortfall in the US banking system, which resulted in the collapse of large financial institutions, the "bail out" of banks by national governments and turbulence 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, the first major subprime related loss to be reported. This particular date is used as the second breakpoint[23]. Finally, the EU sovereign debt crisis in the end of 2009 is also investigated. It led to a crisis of confidence as well as the widening of bond yield spreads and rise of credit default swaps for Eurozone countries such as Greece, Ireland and Portugal, which further intensified the fear of a global contagion[24]. The crisis deepened towards the end of 2009 when there was an abrupt increase in the spreads due to the downgrading of Greece's credit rating by all three major international credit agencies (Fitch, Moody's and S&P). In this paper *December 8, 2009* is set as the third breakpoint, corresponding to the first Greek rating cut by Fitch.

In addition to the economic rationale, 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 sequentially applied both on return series and on absolute returns of each currency to investigate also for volatility breaks (McConnell and Perez-Quiros, 2000; van Dijk *et al.*, 2005). The results from the Chow's test, indicate that in the majority of cases for the FX returns one breakpoint is statistically identified at the 5% level of significance, namely *February 22, 2007*. The CUSUM test also detects parameter instability at the 5% significance level mostly around the same breakpoint. In case of all volatility series, the null of no structural change can be rejected for both Chow's and CUSUM test at the 1% level. The structural break of *February 22, 2007* is finally selected for the return and volatility series

[23] On September 15, 2008, the Lehman Brothers Holdings filed for bankruptcy following drastic losses in its stock and devaluation of its assets by credit rating agencies. The filing marked the largest bankruptcy in U.S. history.

[24] At the beginning of 2010, a failed bond auction in Portugal increased the fear of a Portuguese debt default (Blackstone *et al.*, 2010). The euro was weakened and a global stock and commodity sell off occurred in February 2010 and the following months. Greece was the focal point of the crisis. The Greek government searched for a bailout plan in case it failed to raise the money to fill its budget gap through the credit markets. The provided plan failed to reassure investors, thus leading to an agreement on a defence package of 750 billion€ by the EU and the IMF, in order to prevent the relentless speculative attacks on the euro and eventually restore stability.

of all FX rates, thus combining statistical and economic rationale[25]. Overall, the examined sub-periods are the following: P_1: January 5, 1999 to February 21, 2007 (2122 obs.), P_2: February 22, 2007 to May 10, 2010 (838 obs.). In addition, the entire sample period P_{Total}: January 5, 1999 to May 10, 2010 (2960 obs.) is comparatively investigated.

The descriptive statistics for the return and volatility series of each currency are presented in Table 1. The return series are zero mean-reverting with low corresponding variance. The Jarque-Bera statistic for all FX rates 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 for all currencies. As indicated by skewness, EUR return series are symmetric while JPY and GBP have a longer left tail. The volatility series are not normally distributed with a fat right 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. In addition, 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_1 and P_2 are quite evident in Table 1, where a significant increase in variance can be observed in P_2 for all exchange rates as well as increased fat-tailedness of the return and volatility distributions reflected in the higher kurtosis. Additionally, P_2 witnessed many occasional negative spikes as it can be inferred from the

[25] In Chow's breakpoint testing, 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 differences in the estimated equations. 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 EUR, GBP and JPY, 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 the 5% (for GBP is rejected at 10% significance level). The CUSUM test is based on the cumulative sum of the recursive residuals. In case of EUR and GBP rates it detects parameter instability at 5% level though marginally, around the region of the February 22, 2007 break (e.g., 1950 – 2150 observations). For JPY no structural change is observed. Regarding all 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 for all currencies. In addition, for EUR it rejects the null for all date combinations not including March 10, 2000 and for GBP and JPY, not including December 8, 2009. Finally, the CUSUM test strongly detects parameter instability for all currencies around the February 22, 2007 breakpoint. The selected breakpoints have also been verified with the Bai and Perron (2003) and Andrews and Zivot (1992) tests for unknown dates. These results are available upon request.

skewness of GBP and JPY returns, as opposed to a longer right tail for the EUR. Volatility series also present more spikes in P_2.

Table 1 also reports the contemporaneous correlation matrix for each period. Significant sample cross-correlations are noted for EUR and GBP returns, indicating a high interrelationship between the two markets. JPY shows a low negative correlation (or uncorrelatedness in P_2) with EUR and GBP returns, while all volatility series are positively correlated for all currencies in both periods. However, since linear correlations cannot be expected to fully capture the linear/nonlinear linkages in a reliable way, these results should be interpreted with caution. Consequently, what is needed is a detailed causality analysis, conducted both at the aggregate foreign exchange rates and on each of the wavelet components. In the following section we present the results of the wavelet analysis, while in section 6 we focus on the causality analysis.

4. A WAVELET ANALYSIS OF THE FOREIGN EXCHANGE RATES

As the evaluation of the "scaling" effects on the relationships among FX markets is fundamental to the present study, in this section we offer a thorough investigation of the FX market dynamics across and within all scales both for returns and volatility via the wavelet multiscale decomposition analysis introduced in section 2. As mentioned, this technique allows for locally temporal effects, sharp cusps, structural breaks, time variation and regime switches. We then identify the impact of changes in long term dynamics, and analyze the implications for the flow of information across time scales.

4.1 Minimum entropy wavelet decomposition

The results of the optimal minimum-entropy decomposition for the FX returns and volatility are presented in Table 3. The entropy in each level is compared to that of the raw time series and to that estimated at the previous level. In most cases for the FX returns the optimal decomposition level is the seventh, while for the volatility series the minimum value of the Shannon entropy criterion is calculated at the fourth scale[26].

[26] The 8th scale is the optimal for EUR returns in period P_1 and the 6th for GBP in P_2 and JPY in P_{total}. Regarding the volatility series, the 3rd scale provides the minimum entropy for JPY in P_2.

Additionally, economic reasons can also be identified in considering the appropriate "depth" of the wavelet analysis, based also on the "translation" of the wavelet scales into time horizons. Table 4 "translates" the wavelet scales into appropriate time horizons, thus providing insight on the relation between SIDWT levels and time scales for the time series. Each scale corresponds to a frequency interval and it is associated with a range of time horizons that span from several days to one year. For instance, the detail $\mathbf{D}_2$ is associated with a frequency range of 4-8 days (0.8-1.6 weeks), while $\mathbf{D}_4$ with approximately one month. Scale level $j = 7$ corresponds to a cycle length between 2.1 to 4.3 quarters, or equally between a semester and a yearly variation. Thereafter the notation $\mathbf{D}_j$ (and not the $\tilde{\mathbf{D}}_j$ used in Section 2) corresponds to the SIDWT details, to enhance readability.

To sum up, the FX 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. Also in economic terms it is reasonable to investigate causality relationships for the returns from daily to yearly frequencies, whereas up to monthly variations for the volatility. In real world applications, quarterly or yearly volatility is not interesting for the economic analysis of high-frequency (daily) FX series, nor "traded" in FX markets, as opposed to daily, weekly and monthly volatility. On the contrary, the causality analysis of the returns up to yearly variations can be very useful in detecting FX market linkages with macroeconomic fundamentals and in producing multi-step ahead return or price forecasts.

4.2 Scale-dependent descriptive analysis

Figures 1-2 present the SIDWT wavelet *approximation* and *details* estimated from the Daubechies (db8) class, for the returns and volatility series on EUR. Figures III.1-4 in the Appendix III provide the same information for the other currencies. To distill information from the wavelet components, it is crucial to recall that nonzero wavelet coefficients correspond to activity in a particular range of frequencies over time. Consequently, when the details are rapidly changing,

this implies that its corresponding frequency interval contains important information about the original process[27].

For the EUR return series, for all periods there are no significant differences in high- and low-scale dynamics. All components display a non-cyclical pattern with fairly low oscillation amplitude. Essentially, there is no notable "activity" in high scales at all levels, which can be interpreted as a direct result of the trend-removal procedure, albeit in P_2 the return fluctuations are slightly amplified after the first quarter of 2008, i.e. after entering the financial crisis period. A similar regime switch is also observed in the details of the P_{Total} and P_1. The increased variability is mostly evident in detail $\mathbf{D}_1$ of the EUR absolute return series in P_2, 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. Additionally, in P_{Total} for the EUR volatility there is an increased variability pattern in the high frequency harmonic (first detail) associated with 2-4 days, which might be interpreted as the dominant market dynamics and can be attributed to traders with short-term trading horizons.

The regime switch appearing in the EUR return details immediately after the crisis burst is also depicted in the P_2 details of the EUR volatility[28]. Interestingly, persistent oscillations are present in all detail components of the EUR volatility in P_1, 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_1, but not in P_2 (and consequently not at the end of P_{Total}), where a break in the long-run trend component signifies the entry in the high volatility regime of the 2007-2010 financial crisis. The same applies for the $\mathbf{A}_7$ low frequency component of the EUR return series, yet in a smaller extent.

---

[27] When viewing the SIDWT decomposition, it is evident that the wavelet details display a complicated structure that cannot be attributable to an oscillation at a single frequency. This is due to the fact that the underlying spectrum of these processes is rapidly and dynamically changing within the frequency intervals induced by the wavelet transform.

[28] The occurrence of the structural changes in the wavelet approximations mentioned throughout Section 9 have also been tested with the Chow's test (Chow, 1960) for known (imposed) breaks, with the cumulative sum (CUSUM) test (Brown et al., 1975) for unknown points as well as with the Bai and Perron (2003) and Andrews and Zivot (1992) tests. In addition for the details, the switching regimes have been verified via a Markov-switching model with two regimes. In each regime an AR(1) specification was used.

The qualitative characteristics of the low- and high-frequency components of GBP returns and volatility are directly comparable and analogous to that of EUR. Finally, the Japanese currency market has almost no high-scale dynamics and only the wavelet scales ($\mathbf{D}_1$-$\mathbf{D}_4$) of the JPY volatility exhibit some activity in P_2. The latter demonstrates that agents with short trading horizons (daily-monthly) are mostly affected by the crisis break.

4.3 Heterogeneous market dynamics and micro-foundations

The motivation behind the investigation of "vertical" (across scales) heterogeneity in the variability pattern comes from comparative observation. For example, for JPY returns (Figure III.3) the first two finest scales mostly affect the dynamics appearing in the raw data, while for EUR (Figure 1) and GBP (Figure III.1) all scales seem to contribute to the raw series variability. Likewise, the detail $\mathbf{D}_1$ of the GBP volatility (Figure III.2) in all periods dominates the aggregate raw series oscillation amplitude, whereas other frequency components embed lower information. It is also noticeable that a low frequency shock (displayed in the long-run approximation wavelet component), might lead to a high frequency response by a short time span, as in the case of the crisis emergence depicted in the GBP and JPY volatility series (Figures III.2 and III.4). This vertical heterogeneity suggests the presence of trader behavior with different time horizons. At the highest approximation scale the trading mechanism "comprises" fundamentalists who trade on longer time horizons. Then, at low scales short-term traders and market makers operate with time horizons of a few days up to a month. Each trader class may possess a homogeneous behavior, but it is the combination of these classes in all scales that generates the aggregate time series. Therefore, the underlying dynamics are heterogeneous due to the interaction of all trader classes at different time scales. In such a market, a low-frequency shock infiltrates through all scales, while a high-frequency shock runs out quickly and might have no impact whatsoever in the long-run dynamics. Probably, a characteristic example of a low-frequency shock which penetrated all scales and market "behaviors" is that of the Eurozone sovereign debt crisis observed in all wavelet components of the examined currency series.

Another aspect worthy of investigation is the 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. For example, this is demonstrated for the JPY volatility in P_{Total} and P_2. Overall, the duration of regimes seems to be longer for high-scale (low-frequency) trading horizons, whereas low-scale behavior results in short and frequent regime switching.

Moreover, the "vertical" causality of the regime structure is one-way, in that a low regime variability state at low frequencies identically affects the oscillation state at higher frequencies. Indeed, the results in section 4.2 indicate that if e.g., a low volatility regime is observed at a monthly frequency, it is more likely that there is also a low volatility pattern at the weekly or daily scale. On the contrary, high variability at a low frequency does not necessarily entail a high volatility at higher frequencies (e.g., as in the case of JPY market in P_1). This result is in accordance with the empirical evidence that markets "cool off" after a shock at higher frequencies in a much shorter period than after a lower-frequency, "structural" change.

4.4 Impact of extreme events and structural breakpoints across time scales

The impact of the stock market dot-com bubble is initially investigated, in particular after the breakpoint of March 10, 2000 (obs. 309 of P_{Total}). The estimated sequence of wavelet approximations and details, as depicted in Figures 1-2 for the EUR and III.1-4 (Appendix III) for GBP and JPY, indicates that the currency markets were not affected seriously by the tech-bubble "burst" which partly coincided with the after-Euro era. The volatility regime in all FX rates is relatively "flat" across the scales.

Another extreme event was the terrorist attack at the World Trade Center on September 11, 2001 (obs. 701), which lead to a sharp drop in stock prices worldwide. The analysis of the impact on the currency markets in all scales (both on returns and volatility series) reveals that it was a relatively short one with practically imperceptible consequences, thus no evidence of contagion across scales was observed.

Next, the financial crisis of 2007-2010 occurred, starting from the HSBC write-down in February 22, 2007 (obs. 2123). A close assessment of the wavelet components in Figures 1-2 and III.1-4 indicates that the effect of the subprime crisis is not as evident in currency markets as in the stock markets, except for the case of the JPY volatility. Due to the structural break in the long-run component ($\mathbf{A}_4$), the Japanese market is gradually entering into a high volatility regime (associated with all four scales, a_1 - a_4).

Finally, the Eurozone sovereign debt crisis after December 8, 2009 and during 2010 is analyzed. The first Greek credit rating cut by Fitch corresponds to obs. 2851 in P_{Total}. Interestingly, it can be inferred that international FX markets are experiencing since then a large ongoing turmoil. The estimated wavelet components at all scales clearly indicate that the EUR, GBP and JPY markets entered into a high volatility state since the end of 2008. Moreover, the high volatility state is not uniform across the scales; at lower scales and especially at the finest scale, the time span of the regime becomes wider[29]. This could be safely considered as a warning, precursor signal of an escalating crisis.

5. TIMESCALE CAUSALITY INVESTIGATION

5.1 Causality testing

We perform causality detection via the Granger test, the modified Baek-Brock test and the Breitung-Candelon test. The conventional approach of causality testing is based on the Granger test (Granger, 1969), which assumes a parametric, linear model for the conditional mean. This specification is simple and appealing as the test is reduced to determining whether the lags of one examined variable significantly enter into the equation of the other, albeit it requires the linearity assumption. Baek and Brock (1992) noted that the parametric linear Granger causality test has low power against certain nonlinear alternatives or higher moments. As a result, nonparametric

[29] For example, for return and volatility detail $\mathbf{D}_1$ (approximately 2-4 days) of all currencies in periods P_{Total} and P_2, the estimated wavelet coefficients display a high volatility regime all through the end of the sample. At scales 2-4 (0.8-1.6 weeks to 0.8-1.6 months) the high state is observed with smaller amplitude. In other words, for short-term traders the currency turbulence continues within 2010, whereas for longer horizon traders or investors, the turmoil mostly lasts from 2009 until the first quarter of 2010. In fact, there were several bursts of different oscillation range and amplitude since the end of 2008 and during the period of the Eurozone crisis.

causality tests have been introduced in the literature, directly focusing on predictive power without imposing a linear functional form. Hiemstra and Jones (1994) proposed a causality-in-probability test for nonlinear dynamic relationships which is applied to the residuals of vector autoregressions and it is based on the conditional correlation integrals of lead–lag vectors of the variables. This test relaxes Baek and Brock's assumption of i.i.d time series and instead allows each series to display weak (or short-term) temporal dependence. It can detect the nonlinear causal relationship between variables by testing whether past values influence present and future values. Finally, a test for causality-in-frequency (spectral causality) is also applied. Geweke (1982) and Hosoya (1991) originally proposed a causality measure based on the decomposition of the spectral density, while Yao and Hosoya (2000) developed a Wald-type test procedure based on a complicated set of nonlinear restrictions on the parameters of vector autoregressions. In a latter study, Yao and Hosoya (2000) applied a numerical method to estimate the nonlinear function of the autoregressive parameters and the asymptotic covariance matrix. Recently, Breitung and Candelon (2006) proposed a simplified test procedure that is based on a set of linear hypotheses on the parameters of a bivariate vector autoregressive model. It allows testing for short- and long-run causality at a specified range of frequencies. The test by Breitung and Candelon (2006), as well as those by Geweke (1982) and Hosoya (1991) upon which the Breitung and Candelon test is based, provides very good causality results – in terms of size and power properties – only at some pre-specified frequency range, which depends on the input data frequency. Applying wavelet analysis could provide an efficient means of overcoming the constraint of reaching a threshold in the lowest possible frequency investigated, namely probing further in the "long-run" behavior (i.e., in this study reaching closer to business cycle fluctuations). The three causality tests are formally described in Appendix II.

5.2 Empirical analysis

The proposed empirical analysis involves three steps. In the first step, the short- and long-run spectral causal relationships are explored at a pre-specified range of frequencies applying the Breitung-Candelon test on the aggregate log-differenced time series. Next, the Granger causality

test is employed on the original and on the "disaggregated" wavelet components in order to investigate the linear dynamic linkages at various scales. Lastly, bivariate vector autoregressive filtering is implemented on the raw and decomposed series and the residuals are examined pairwise by the nonparametric modified Baek-Brock test. Thus, the nature and direction of causality explored for each scale component of the return and volatility series is compared against the causality results of the aggregate series.

Linear and spectral causality are investigated in a VAR representation. VAR modelling is also applied for the volatility series in accordance with previous results derived by Nikkinen *et al.* (2006)[30]. The results from the SIC criterion, taking into consideration many lag specifications for the bivariate VAR modelling as in Engel and West (2005), indicate in most of the cases two lags for the FX return series and their wavelet components in all periods. Similarly, four lags are chosen for the volatility series and the corresponding components. Finally, for the nonlinear causality test, in what follows the common lag lengths used are $\ell_X = \ell_Y = 1$. The test is applied on the VAR residuals derived from the pairwise linear causality testing and the distance measure is set to $\varepsilon = 1.5$, as suggested by Hiemstra and Jones (1994)[31].

The FX multiscale causality results from all tests employed in the study are reported in Tables 5 and 6. The simplifying notation " ** " is used to indicate that the corresponding *p*-value of a particular causality test is smaller than 1% and " * " that the *p*-value of a test is in the range 1-5%. This was necessitated in order to overcome the difficulty of presenting large tables with numbers.

---

[30] The results from testing nonstationarity with the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests show that FX log-levels log-levels are $I(1)$ processes, whilst the returns and volatility series are stationary. In addition, the trace and maximum eigenvalue statistics (Johansen, 1988; Johansen and Juselius, 1990) applied on the log-prices series did not identified any cointegrating vectors and the null of no cointegration was not rejected (Table 2). The lag lengths for testing nonstationarity 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. ADF and PP tests indicate that the null of a unit root cannot be rejected at 1% for all currency log-levels 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 currencies and periods. 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. Based on these results and in order to identify the correct model specification for the investigation of linear and spectral causality (i.e., VAR or VECM), the trace and maximum eigenvalue statistics were further applied to the log-prices series to explore possible effects of cointegration. For all pairs the Johansen tests did not identified any cointegrating vectors and the null of no cointegration was not rejected (Table 2). Thus, linear and spectral causality are investigated with a VAR representation.

[31] In the estimation $\varepsilon = 0.5$ and $\varepsilon = 1$ were also considered, with no qualitative difference in the results. In addition, evidence from the second and third common lag lengths did not significantly modified the nonlinear causality results.

Directional causalities in the text are denoted by the functional representation " → ". The causality linkages are also depicted diagrammatically in Figures 3 and 4 where strong unidirectional or bidirectional causality (" ** ") is denoted by a double one-sided or two-sided arrow respectively. In addition, the corresponding results on the cross-correlation of the return and volatility wavelet components are displayed in Table 7, while the spectral causal relationships of the Breitung-Candelon test are graphically illustrated in Figures III.5 and III.6 (Appendix III). These figures report the test statistics along with their 5% and 1% critical values (broken lines) for all frequencies in the x-axis interval $[0,\pi]$ as in Breitung and Candelon (2006)[32]. This interval, based on the frequency of the input raw data (i.e., daily), corresponds to a frequency range from 1 day to 16 days, or $[0.0-0.8]$ months. The right value represents the lowest frequency upon which the Breitung-Candelon test can infer on causality. Hence, the results of the frequency-domain test for the aggregate series can be comparatively analyzed against those of up to the a_4 scale for the return and volatility series.

The hypothesis of no causality for the Breitung-Candelon spectral test in case of the unidirectional EUR→GBP return relationship is rejected at the 5% level in the period P_1 for frequencies in the *x*-axis $[2.4,3.2]$, corresponding to a range of 1-2 days or roughly to scale a_1 in the wavelet decomposition. In the same period JPY causes GBP for return series in $[1.2,2.0]$, which corresponds to a frequency range of 3-6 days or scale a_2. The volatility series for P_1 reveal a strong causal linkage of JPY→EUR (1% significance level) in the range of 8-16 days directly associated with the a_3 wavelet scale. In all other periods (P_2 and P_{Total}) after investigating volatility series, strong causal relationships are detected (1% level) among all currencies mostly in the frequency range of $[0.0,1.2]$, or 6-16 days (comparable to a_3). Instead, inferring upon the returns, the results vary across periods. In P_2, a bidirectional linkage GBP↔JPY emerges at the 1% level at the range of 2-8 days and a univariate relationship JPY→EUR (3-16 days), while in P_{Total} causalities run from

[32] It follows from Breitung and Candelon (2006) that for frequencies in $(0,\pi)$ the effect of the coefficient c (section 6.3) in the size and power of the test is minimal and the empirical power is very close to the asymptotic power.

GBP to JPY (3-16 days) and EUR (6-8 days) as well as from JPY to EUR (3-8 days) at weaker significance levels.

Next, the analysis is unfolded on the basis of each wavelet scale. The long-run linkages exhibit the following characteristics: in case of returns, linear and nonlinear feedback relationships are observed for the $\mathbf{A}_7$ component in all periods except for the absence of the EUR-JPY causality in P_2. The volatility component $\mathbf{A}_4$ provides strong statistical evidence on linear and nonlinear bidirectional linkages in all periods, with the exception of P_2 when only EUR↔GBP is observed. The correlation of the long-run component is stronger (positively or negatively higher) for plain and absolute returns compared to the one of the details, for the majority of the examined currency pairs. The low-frequency components $\mathbf{D}_5$-$\mathbf{D}_7$ for the return series present identical interdependencies. Specifically, the causalities run from EUR to GBP in P_1, in both directions for GBP-JPY in P_2, while they "add-up" in P_{Total}. There is also evidence of unidirectional nonlinear relationships from EUR to GBP and JPY in P_1, and strong feedback nonlinear links in P_2 and P_{Total} for all pairs. Instead, cross-correlation in these high scales is positively high only for EUR-GBP, whereas around -0.3 for the other two pairs. Furthermore, the third and fourth wavelet scales (a_3 and a_4) display the same features in terms of investigated causalities. Firstly regarding returns, linear unidirectional linkages are observed from EUR to JPY and GBP in P_1 and a bidirectional relationship for the GBP-JPY pair in P_2. In P_{Total} all links observed in P_1 and P_2 exist. The nonlinear feedback causalities for all pairs observed in P_2 are identical to the ones in P_{Total}, while in P_1 a sequential pattern emerges from EUR to JPY, to GBP and back to EUR Secondly, in terms of absolute returns, the a_3 and a_4 scales reveal absence of linear causality in P_1, and only a strong bidirectional link EUR↔JPY in P_2 and P_{Total}. The nonlinear linkages EUR→GBP and EUR→JPY are observed in P_1, while all currencies present bidirectional causalities in the other two periods. The correlation at a_3 and a_4 scales is similar to the one observed for raw series, but different compared to the long-run component. At the second wavelet scale a_2 linear bidirectional relationships exist in all periods, as well as nonlinear bidirectional causalities for all pairs. In volatility series, no

currency Granger causes the other in P_1 and only a strong EUR↔JPY link is observed in the other periods. Again, nonlinear links dominate P_{Total} and P_2 , while EUR causes GBP and JPY in P_1. Finally, at the finest scale a_1 (highest frequency) the same as in a_2 applies for the volatility series regarding the nature and direction of the interdependencies. In the case of returns, a feedback relationship emerges for the EUR-JPY and GBP-JPY pairs in P_2 and P_{Total}, whereas nonlinear linkages appear for all pairs, with the exception of GBP-JPY in P_1. In addition, the cross-correlation of the wavelet components for the two finest scales both for returns and volatility, has approximately the same value as in a_3, albeit somewhat lower for the more "noisy" scale a_1 (highest frequencies).

Overall, the evidence provided in this section leads to the conclusion that spillovers and interactions between FX markets have different characteristics at different timescales[33]. Especially when the nonlinear effects are accounted for, the evidence of dynamical bidirectional causality implies that the pattern of leads and lags changes over time. The market agents filter information relevant to their positions as new information arrives and, at any time point, one FX market may lead the other and vice versa. Overall, there is no indication of a *"global causality"* behavior or a "*prevailing pattern*" of interdependencies dominating at all scales.

6. CONCLUSIONS

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. In contrast to simple disaggregation at different time horizons, this study relied on wavelet multiresolution to analyze the controversial issue of the dependence structure of the FX markets. The aim of the paper was to test for the existence of causal relationships among the most liquid and widely traded currencies in the world ("FX majors"), namely the EUR, GBP and JPY. The nature and direction of causality was investigated for each component of the raw return and volatility series corresponding to a different sampling frequency and was compared against the

[33] The inferred results seem also to corroborate with Gençay *et al* (2002) and Ramsey and Lampart (1998a) on multi-scale linkages between macroeconomic variables.

results of the original aggregate series. Causality detection was performed with the use of the linear Granger test, the modified Baek-Brock test for nonlinear causality and the frequency domain Breitung-Candelon test. The explored period, starting from the introduction of the Euro, covers diverse regimes including the rise and fall of the "dot-com" bubble, the financial crisis of 2007-2010 and the Eurozone sovereign debt crisis in early 2010. The timescale causality investigation, including wavelet cross-scale correlation, proceeded complementarily albeit distinctly both on the basis of period as well as of wavelet scale. It involved the comparative examination of the aggregate versus component-based causality, of linear vis-à-vis nonlinear and spectral causality as well as of short- versus long-run linkages.

Moreover, this study attempted to probe into the micro-foundations of across-scale heterogeneity in the variability pattern, on the basis of trader behavior with different time horizons and information flow across time scales. The trading pattern of fundamentalists is reflected at the highest approximation wavelet scale, while at lower scales short-term traders and market makers operate. Each trader class may possess a homogeneous behavior, but the aggregate underlying market dynamics are heterogeneous due to the interaction of all trader classes at different time scales. In such a market, a low-frequency shock infiltrates through all scales, while a high-frequency shock runs out quickly and might have no impact whatsoever in the long-run dynamics. The propagation properties of this heterogeneous-driven behavior were investigated, the causality structure from low-to-high frequency was identified, and the implications for the flow of information across time scales in the FX markets were inferred. 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. Finally, an asymmetry in volatility dependence across different time horizons was identified as an important stylized property. The across-scale causality of the various regime structures is one-way, in that a low regime variability state at low frequencies identically affected the oscillation state at higher frequencies. On the contrary, high variability at a low frequency did not necessarily entail a high

volatility at higher frequencies. This result is in accordance with the empirical evidence that markets "cool off" after a shock at higher frequencies in a much shorter period than after a significant structural change.

In technical terms, the present work introduced new practical guidelines for wavelet implementation and expanded the relevant literature by presenting an invariant discrete wavelet transform that contains no phase shifts, relaxes the assumption of a "dyadic-length" time series, enables multi-scale point-to-point comparison and copes effectively with "boundary effects". Beyond the existing practice that has utilized subjective judgement or economic reasoning in estimating the appropriate "depth" of the wavelet analysis, a new entropy-based methodology was introduced for the determination of the optimal decomposition level.

Overall, the results strongly indicate that interactions between currency markets have different characteristics at different timescales and that there is no "*global causal behavior*" that prevails at all time horizons. When the nonlinear effects are accounted for, neither FX market leads or lags the other consistently, namely the pattern of leads and lags changes over time. Given that causality can vary from one direction to the other, a finding of bidirectional causality over the sample period may be taken to imply a changing pattern of leads and lags over time. In particular, market participants filter information relevant to their positions as new information arrives and, at any time point, one FX market may lead the other and vice versa.

An interesting subject for future research is the nature and source of the nonlinear linkages, as it was shown that volatility effects might partly induce nonlinear causality. Conditional volatility or statistically significant higher-order moments may account for a part of the nonlinearity in daily exchange rates, but only in some cases as it is already known by many studies, including Scheinkman and LeBaron (1989) and Bekiros and Diks (2008). In general, the detailed knowledge of the nature of interdependency between the currency markets and the degree of their integration at different timescales will expand the information set available to practitioners and policymakers. The results of this study, apart from offering a much better understanding of the dynamic heterogeneous relationships underlying the major currency

markets, may have important implications for market efficiency. In that, they may be useful in future research to quantify the process of financial integration or may influence the greater predictability of these markets.

REFERENCES

Allan, D.W., 1966, Statistics of atomic frequency standards. Proceedings of the IEEE, 54 (2), 221-230.

Andersen, T.G. and T. Bollerslev, 1997, Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4 (2-3), 115-158.

Baek, E. and W. Brock, 1992, A general test for non-linear Granger causality: bivariate model. *Working paper*, Iowa State University and University of Wisconsin, Madison, WI.

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

Bank for International Settlements (BIS), 2007, Triennial Central Bank Survey: Foreign exchange and derivatives market activity in 2007. Basel, Switzerland.

Bessler, D.A. and J. Yang, 2003, The structure of interdependence in international stock markets. *Journal of International Money and Finance* 22, 261–287.

Bekiros S.D. and C.G.H. Diks, 2008, The nonlinear dynamic relationship of exchange rates: Parametric and nonparametric causality testing. *Journal of Macroeconomics* 30, 1641-1650

Bénassy-Quéré, A. and A. Lahrèche-Révil, 2000, The Euro as a Monetary Anchor in the CEECs. *Open Economies Review* 11 (4), 303-321.

Black, F., 1986, Noise. *Journal of Finance* 41, 529–543.

Blackstone, B., T. Lauricella and N. Shah, 2010, Global Markets Shudder: Doubts About U.S. Economy and a Debt Crunch in Europe Jolt Hopes for a Recovery. *The Wall Street Journal*.

Blanchard, O. and M.W. Watson, 1982, *Bubbles, rational expectations, and financial markets.* in: Wachtel, P. (Ed.) Crises in the economic and financial structure, Lexington, MA.

Breitung, J. and B. Candelon, 2006, Testing for short- and long-run causality: A frequency-domain approach. *Journal of Econometrics* 132, 363-378.

Brown, R.L., J. Durbin and J.M. Evans, 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., I. Daubechies, B. Jawerth and P. Vial, 1993, *Multiresolution analysis, wavelets and fast wavelet transform on an interval.* CRAS Paris 316, 417–421.

Coifman, R.R. and M.V. Wickerhauser, 1992, Entropy-based algorithms for best basis selection. *IEEE Transaction on Information Theory* 38 (2), 713–718.

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

Daubechies, I., 1992, *Ten Lectures on Wavelets*. Regional Conference Series in Applied Mathematics (SIAM), Society for Industrial and Applied Mathematics Vol. 61. Philadelphia, USA.

Engel, C., and K.D. West, 2005, Exchange Rates and Fundamentals. *Journal of Political Economy* 113, 485-517

Engle, R.F., T. Ito and W-L. Lin, 1990, Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. *Econometrica* 58(3), 525–542.

Flood, R.P. and P. Isard, 1989, Monetary policy strategies. *IMF Staff Papers* 36, 612–632.

Friedman, M., 1963, *Windfalls, the 'horizon,' and related concepts in the permanent-income hypothesis,* in: Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld, Carl Christ, et al. (Eds.), Stanford, CA, pp. 3–28.

Gençay, R., B. Whitcher and F. Selçuk, 2001, Differentiating intraday seasonalities through wavelet multi-scaling. *Physica A* 289(3-4), 543–556.

_____, 2002, *An introduction to wavelets and other filtering methods in finance and economics.* Academic Press, San Diego.

Geweke, J., 1982, Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association* 77, 304–324.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3), 424-438.

Greenspan, A., 2007, *The Age of Turbulence: Adventures in a New World*. Penguin Press.

Grossmann, A. and J. Morlet, 1984, Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J Math Anal* 15(4), 723–736.

Herley, C., 1995, Boundary filters for finite-length signals and time-varying filter banks. *IEEE Transactions on Circuits and Systems-II* 42, 102-114.

Hiemstra, C. and J.D. Jones, 1994, Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance* 49(5), 1639-1664.

Hogan, W.P. and I.G. Sharpe, 1984, On the relationship between the NY closing spot U.S.$/$A rate and the reserve bank of Australia's official rate. *Economics Letters* 14(1), 73–79.

Hosoya, Y., 1991, The decomposition and measurement of the interdependency between second-order stationary processes. Probability Theory and Related Fields 88(4), 429–444.

Ito, T. and V.V. Roley, 1987, News from the U.S. and Japan: which moves the yen/dollar exchange rate? *Journal of Monetary Economics* 19(2), 225–277.

Jensen, J. and B. Whitcher, 2000, Time-Varying Long-Memory in Volatility: Detection and Estimation with Wavelets. *Technical Report*, Department of Economics, University of Missouri.

Johansen, S., 1988, Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12(2-3), 231–254.

Johansen, S. and K. Juselius, 1990, Maximum likelihood estimation and inference on cointegration with application to the demand for money. *Oxford Bulletin of Economics and Statistics* 52, 169–209.

Koutmos, G. and G.G. Booth, 1995, Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance* 14(6), 747–762.

Krugman, P., 1991, Target zones and exchange rate dynamics. *The Quarterly Journal of Economics* 106, 669–682.

_____, 2009, *The Return of Depression Economics and the Crisis of 2008*. Norton.

Laopodis, N., 1997, US dollar asymmetry and exchange rate volatility. *Journal of Applied Business Research* 13(2), 1–8.

_____, 1998, Asymmetric volatility spillovers in deutschemark exchange rates. *Journal of Multinational Financial Management* 8, 413–430.

Lindberg, H. and P. Soderlind, 1994, Testing the Basic Target Zone Models on the Swedish Data 1982–1990. *European Economic Review* 38, 1441–1469.

Ma, Y. and A. Kanas, 2000, Testing for a nonlinear relationship among fundamentals and exchange rates in the ERM. *Journal of International Money and Finance* 19(1), 135–152.

MacKinnon, J., Haug, A. and L. Michelis, 1999, Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. *Journal of Applied Econometrics* 14, 563-577.

Mallat, S., 1989, A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7), 674–693.

McConnell, M.M. and G. Perez-Quiros, 2000, Output fluctuations in the United States: What has changed since the early 1980s? *American Economic Review* 90, 1464–1476.

Meese, R.A. and K. Rogoff, 1983, Empirical exchange rate models of the seventies: Do they fit out of sample?. *Journal of International Economics* 14(1-2), 3–24.

Nason, G. P. and B. W. Silverman, 1995, *The Stationary Wavelet Transform and Some Statistical Applications*, in: A. Antoniadis and G. Oppenheim (Eds.), Wavelets and Statistics, New York: Springer, pp. 281-299.

Newey, W.K. and K.D. *West, 1994, Automatic Lag Selection in Covariance Matrix Estimation. Review* of Economic Studies 61(4), 631-653

Nikkinen, J., P. Sahlström and S. Vähämaa, 2006, Implied volatility linkages among major European currencies. *Journal of International Financial Markets, Institution & Money* 16, 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 & Computation* 26, 467–486.

Percival, D.B., 1995, On estimation of the wavelet variance. *Biometrika* 82(3), 619-631.

Percival, D.B. and H.O. Mofjeld, 1997, Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association* 92, 868-880.

Pesquet, J.-C., H. Krim and H. Carfantan, 1996, Time-invariant orthonormal wavelet representations. *IEEE Transactions on Signal Processing* 44, 1964-1970.

Ramsey, J.B. and C. Lampart, 1998b, The decomposition of economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics & Econometrics* 3(1), 23-42.

Scheinkman, J., and B. LeBaron, 1989, Nonlinear Dynamics and Stock Returns. The *Journal of Business* 62 (3), 311-337.

Stock J. and M. Watson, 2003, Has the Business Cycle Changed and Why?, *NBER Macroeconomics Annual* 17, 159-218

van Dijk, D., D.R. Osborn and M. Sensier, 2005, Testing for causality in variance in the presence of breaks. *Economics Letters* 89(2), 193–199.

Wang, Z., J. Yang and Q. Li, 2007, Interest rate linkages in the Eurocurrency market: Contemporaneous and out-of-sample Granger causality tests. *Journal of International Money and Finance* 26(1), 86-103.

Yao, F. and Y. Hosoya, 2000, Inference on one-way effect and evidence in Japanese macroeconomic data. *Journal of Econometrics* 98(2), 225–255.

Zivot, E., and D. Andrews, 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.

TABLE 1: DESCRIPTIVE STATISTICS

Return Statistics

Statistic	EUR			GBP			JPY		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Std. dev.	0.007	0.006	0.007	0.006	0.005	0.008	0.007	0.006	0.008
Skewness	0.128	0.074	0.218	-0.379	-0.008	-0.553	-0.105	-0.017	-0.141
Kurtosis	4.447	3.722	5.255	5.844	3.640	5.547	6.742	5.576	6.456
JB test	266.18*	48.10*	184.28*	1068.48*	36.28*	269.17*	1732.19*	586.74*	419.92*
Q(12)	12.23	14.85	10.74	11.38	9.24	16.52	33.91*	4.64	37.05*

Return Correlation matrix

	P_{Total}			P_1			P_2		
	EUR	**GBP**	**JPY**	**EUR**	**GBP**	**JPY**	**EUR**	**GBP**	**JPY**
EUR	*1*			*1*			*1*		
GBP	*0.658*	*1*		*0.681*	*1*		*0.638*	*1*	
JPY	*-0.232*	*-0.126*	*1*	*-0.356*	*-0.345*	*1*	*-0.039*	*0.139*	*1*

Volatility Statistics

Statistic	EUR			GBP			JPY		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Mean	0.005	0.005	0.005	0.004	0.004	0.006	0.005	0.005	0.006
Std. dev.	0.004	0.004	0.005	0.004	0.003	0.005	0.005	0.004	0.006
Skewness	1.588	1.293	1.883	2.042	1.250	1.978	2.270	1.955	2.207
Kurtosis	6.669	4.862	7.852	10.191	4.720	8.316	12.650	9.995	11.409
JB test	2905.27*	898.17*	1317.05*	8433.69*	814.37*	1532.96*	14027.58*	5678.53*	3149.45*
Q(12)	365.54*	53.49*	378.50*	832.81*	42.96*	465.12*	403.88*	60.91*	209.61*

Volatility Correlation matrix

	P_{Total}			P_1			P_2		
	EUR	**GBP**	**JPY**	**EUR**	**GBP**	**JPY**	**EUR**	**GBP**	**JPY**
EUR	*1*			*1*			*1*		
GBP	*0.485*	*1*		*0.483*	*1*		*0.495*	*1*	
JPY	*0.291*	*0.307*	*1*	*0.226*	*0.198*	*1*	*0.376*	*0.387*	*1*

Notation: The FX exchange returns are defined as $r_t = \log\left(P_t\right) - \log\left(P_{t-1}\right)$, where P_t is the closing level on day t, while the volatility series as the absolute value of the returns $u_t = \left|r_t\right|$ as in Jensen and Whitcher (2000) and Gencay *et al.* (2002). (*) denotes significance at 5% confidence level. The periods are P_1: 01/05/1999-02/21/2007, P_2: 02/22/2007-05/10/2010 and P_{Total}: 01/05/1999-05/10/2010.

TABLE 2: UNIT ROOT AND COINTEGRATION TESTS

Unit Root tests

Periods		P_{Total}				P_1				P_2			
Variables		**ADF**		**PP**		**ADF**		**PP**		**ADF**		**PP**	
		ADF_c	ADF_τ	PP_c	PP_τ	ADF_c	ADF_τ	PP_c	PP_τ	ADF_c	ADF_τ	PP_c	PP_τ
EUR	P_t	0.76	0.23	0.77	0.23	0.88	0.21	0.87	0.17	0.49	0.77	0.47	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*
GBP	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*
JPY	P_t	0.57	0.49	0.49	0.44	0.21	0.49	0.23	0.53	0.52	0.18	0.51	0.25
	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*

Cointegration tests

Pair		Trace statistic						Maximum Eigenvalue statistic					
		P_{Total}		P_1		P_2		P_{Total}		P_1		P_2	
X	**Y**	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$
EUR	GBP	0.993	0.489	0.501	0.343	0.766	0.457	0.997	0.489	0.004	0.343	0.761	0.459
EUR	JPY	0.473	0.152	0.622	0.716	0.801	0.319	0.629	0.152	0.551	0.716	0.851	0.319
GBP	JPY	0.641	0.243	0.825	0.800	0.645	0.362	0.729	0.243	0.769	0.800	0.665	0.362

Notation: 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. Reported numbers for the trace and max. eigenvalue statistics are the MacKinnon-Haug-Michelis (1999) *p*-values. (*) denotes significance at 1% confidence level. The periods are P_1: 01/05/1999-02/21/2007, P_2: 02/22/2007-05/10/2010 and P_{Total}: 01/05/1999-05/10/2010

TABLE 3: OPTIMAL MINIMUM-ENTROPY DECOMPOSITION

WL Level	EUR/USD						GBP/USD						USD/JPY					
	Returns			**Volatility**			**Returns**			**Volatility**			**Returns**			**Volatility**		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Raw	***1.172***	***0.780***	***0.392***	***1.172***	***0.780***	***0.392***	***0.966***	***0.538***	***0.428***	***0.966***	***0.538***	***0.428***	***1.228***	***0.726***	***0.502***	***1.228***	***0.726***	***0.502***
1	1.274	1.164	0.436	0.551	0.508	0.187	0.998	0.789	0.446	0.429	0.328	0.199	1.404	1.131	0.628	0.616	0.528	0.248
2	1.215	1.074	0.458	0.519	0.475	0.181	1.033	0.768	0.503	0.453	0.357	0.205	1.318	1.087	0.564	0.622	0.522	0.264
3	1.066	0.964	0.422	0.474	0.441	0.184	0.988	0.698	0.530	0.448	0.364	0.204	1.278	1.057	0.567	0.575	0.519	0.277
4	1.180	1.109	0.383	0.454	0.387	0.141	1.062	0.777	0.523	0.392	0.289	0.193	1.091	1.057	0.549	0.683	0.474	0.293
5	1.071	0.859	0.523	0.456	0.448	0.174	1.000	0.863	0.435	0.424	0.295	0.227	1.283	1.012	0.494	0.717	0.483	0.404
6	1.291	1.044	0.546	0.507	0.414	0.207	0.942	0.912	0.367	0.573	0.494	0.287	0.962	0.954	0.506	1.261	0.673	0.898
7	0.995	0.942	0.298	1.345	0.920	0.956	0.715	0.456	0.407	1.377	0.643	1.203	0.985	0.622	0.201	2.238	1.525	1.480
8	1.468	0.719	0.887	2.514	0.708	1.796	1.623	0.533	1.224	3.009	0.512	2.500	1.200	1.347	0.328	2.110	1.707	1.039
9	1.199	1.256	0.724	8.573	3.789	5.997	1.244	0.788	0.803	7.482	1.320	6.648	1.116	0.937	0.379	4.308	2.964	4.160
10	2.353	1.433	1.111	10.867	5.428	8.625	1.629	0.939	1.296	15.405	2.029	9.764	1.536	0.974	0.379	7.491	2.914	6.820

Notation: Shaded 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 4: TRANSLATION/CONVERSION OF WAVELET SCALES INTO TIME HORIZONS

	Time Horizons				
WL Scale	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

Notation: 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.

TABLE 5: CAUSALITY RESULTS (RETURN SERIES)

WL Component	EUR ↔ GBP												EUR ↔ JPY												GBP ↔ JPY											
	Spectral Causality						**Linear Granger Causality**						**NonLinear Causality**						**Spectral Causality**						**Linear Granger Causality**						**NonLinear Causality**					
	X→Y			Y→X			X→Y			Y→X			X→Y			Y→X			X→Y			Y→X			X→Y			Y→X			X→Y			Y→X		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Raw	*		*				*	**					**	*	**	**		**		*	**				**	**	**	**		**	**		*	**		*	**		**	**
A_7							**	**	**	**	**	**	**	**	*	**	**	**				**	**	**	**	**	**				**	**	**	**	**	**		**	*	**	**
D_1								*					**	*	**	**	*	**				*		*	**	**					**		**	**		**	**		**	**
D_2							**	**		*	*		**		**	**	*	**				**		**	*	**					*	*	**	**		**	**		**	*	**
D_3							*	**					**		**	**	*	**				*	*		**	**						**	**	**		**	**		**	*	**
D_4							*	**					**		**	**	*	**							**		**					*	**	**		**	**		**	*	**
D_5							*	*					**	*	**	**		**							**	*	**					**	**	**		**	**		**		**
D_6							*	*					**	*	**	**		**							**	*	**					**	**	**		**	**		**		**
D_7							*	**					**	*	**	**		**							**	*	**					**	**	**		**	**		**		**

Notation: X→Y: r_x does not Granger cause r_y. Statistical significance represents *5% (*)* and *1% (**)*. The foreign exchange rates Euro (EUR), Great Britain Pound (GBP) and Japanese Yen (JPY) are denoted relative to United States dollar (USD). The exact ratios represent EUR/USD, GBP/USD and USD/JPY respectively. The periods are P_1: 01/05/1999-02/21/2007, P_2: 02/22/2007-05/10/2010 and P_{Total}: 01/05/1999-05/10/2010. The spectral causality is tested only on the raw series as in Breitung and Candelon (2006). For all pairs the Johansen tests did not identified any cointegrating vectors and the null of no cointegration was not rejected (Table 2). Thus, linear and spectral causality are investigated with a VAR representation. The results from the SIC criterion, taking into consideration many lag specifications for the bivariate VAR modelling, as in Engel and West (2005), indicate in most of the cases two lags for the FX return series and their wavelet components in all periods. Finally, for the nonlinear causality test in what follows, the common lag lengths used are $\ell_x = \ell_y = 1$. The nonlinear test is applied on the VAR residuals derived from the pairwise linear causality testing and the distance measure is set to $\varepsilon = 1.5$, as suggested by Hiemstra and Jones (1994).

TABLE 6: CAUSALITY RESULTS (VOLATILITY SERIES)

WL Component	EUR ↔ GBP												EUR ↔ JPY												GBP ↔ JPY											
	Spectral Causality						**Linear Granger Causality**						**NonLinear Causality**						**Spectral Causality**						**Linear Granger Causality**						**NonLinear Causality**					
	X→Y			Y→X			X→Y			Y→X			X→Y			Y→X			X→Y			Y→X			X→Y			Y→X			X→Y			Y→X		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Raw	**		**	**		**	**		**	**		**	**		**	*		**	*		**	**	**	**		**	**		**	**		**	*	**	**		**
A_4							**	**	**	**	**	**	**	*	**	**	**	*				**	**	**	**	**	**	**	**	**	**	**	**		**	**	
D_1										*			**	**	**	**		**	*		**	*		**	**	**	**	**		**							
D_2									**				**	**	**	**	**	**	**		**	*		**	**	**	**	**		**				**	*	**	**	
D_3													**	*	**	**		**	**		**	**		**	**	*	**	**		**				**	*	**	**	
D_4													**	*	**	**		**	**		**	**	**	**	*			**		**				**		**	**	

Notation: X→Y: r_x does not Granger cause r_y. Statistical significance represents *5% (*)* and *1% (**)*. The foreign exchange rates Euro (EUR), Great Britain Pound (GBP) and Japanese Yen (JPY) are denoted relative to United States dollar (USD). The exact ratios represent EUR/USD, GBP/USD and USD/JPY respectively. The periods are P_1: 01/05/1999-02/21/2007, P_2: 02/22/2007-05/10/2010 and P_{Total}: 01/05/1999-05/10/2010. The spectral causality is tested only on the raw series as in Breitung and Candelon (2006). For all pairs the Johansen tests did not identified any cointegrating vectors and the null of no cointegration was not rejected (Table 2). Thus, linear and spectral causality are investigated with a VAR representation. The results from the SIC criterion, taking into consideration many lag specifications for the bivariate VAR modelling, as in Engel and West (2005), indicate in most of the cases four lags for the FX volatility series and their wavelet components in all periods. Finally, for the nonlinear causality test in what follows, the common lag lengths used are $\ell_X = \ell_Y = 1$. The nonlinear test is applied on the VAR residuals derived from the pairwise linear causality testing and the distance measure is set to $\varepsilon = 1.5$, as suggested by Hiemstra and Jones (1994).

TABLE 7: CROSS-SCALE CORRELATION RESULTS (RETURN AND VOLATILITY SERIES)

Returns

WL Component	EUR - GBP			EUR - JPY			GBP - JPY		
	Cross-correlation			**Cross-correlation**			**Cross-correlation**		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Raw	0.658	0.681	0.638	-0.232	-0.356	-0.039	-0.126	-0.345	0.139
A_7	0.831	0.831	0.891	-0.289	-0.465	-0.026	-0.085	-0.624	0.372
D_1	0.671	0.689	0.652	-0.238	-0.389	-0.001	-0.144	-0.383	0.155
D_2	0.654	0.675	0.635	-0.227	-0.362	-0.020	-0.137	-0.354	0.129
D_3	0.655	0.676	0.638	-0.230	-0.358	-0.032	-0.125	-0.344	0.140
D_4	0.655	0.677	0.638	-0.225	-0.353	-0.026	-0.122	-0.342	0.143
D_5	0.655	0.677	0.637	-0.229	-0.353	-0.038	-0.125	-0.341	0.137
D_6	0.656	0.679	0.636	-0.231	-0.353	-0.040	-0.128	-0.343	0.134
D_7	0.656	0.680	0.634	-0.231	-0.354	-0.039	-0.127	-0.343	0.137

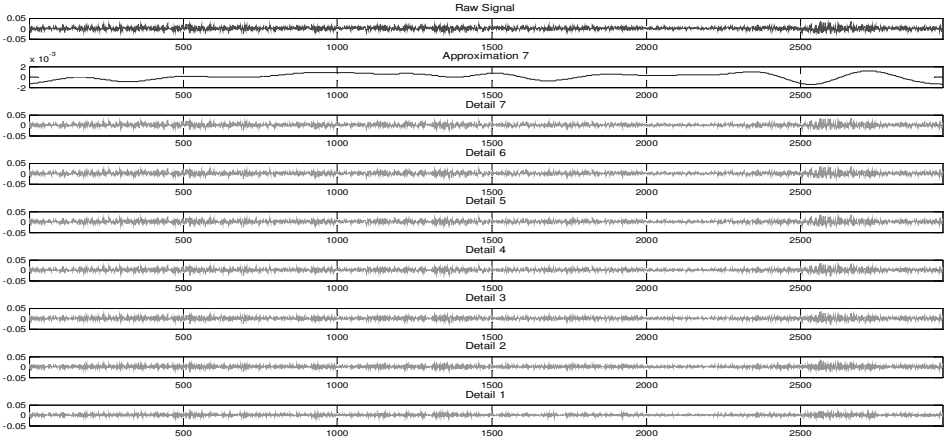
Volatility

WL Component	EUR - GBP			EUR - JPY			GBP - JPY		
	Cross-correlation			**Cross-correlation**			**Cross-correlation**		
	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2	P_{Total}	P_1	P_2
Raw	0.485	0.483	0.495	0.291	0.226	0.375	0.307	0.198	0.387
A_4	0.768	0.533	0.921	0.585	0.302	0.746	0.569	0.076	0.672
D_1	0.382	0.475	0.246	0.172	0.201	0.121	0.207	0.183	0.241
D_2	0.408	0.479	0.306	0.210	0.208	0.219	0.229	0.199	0.267
D_3	0.418	0.479	0.333	0.218	0.204	0.247	0.237	0.199	0.287
D_4	0.424	0.478	0.349	0.235	0.216	0.273	0.248	0.208	0.302

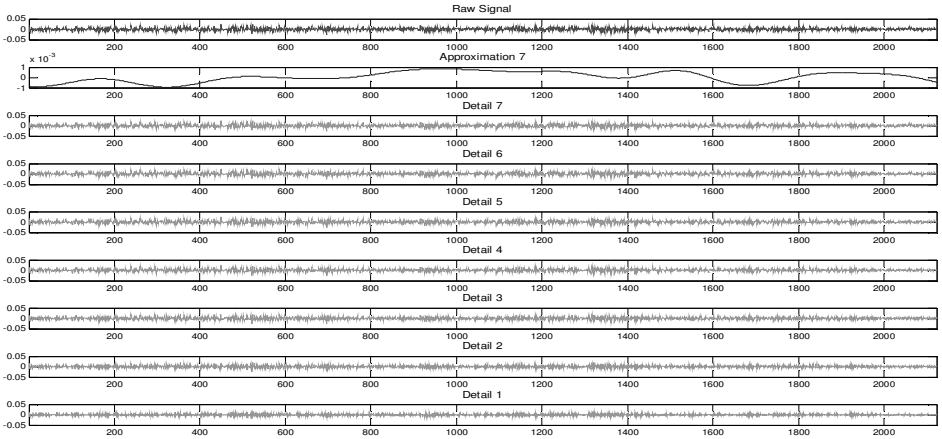
Notation: Reported values indicate the wavelet cross-correlation between all pairs of exchange rate returns and volatility.

FIGURE 1: WAVELET MULTI-SCALE ANALYSIS (EUR RETURNS)

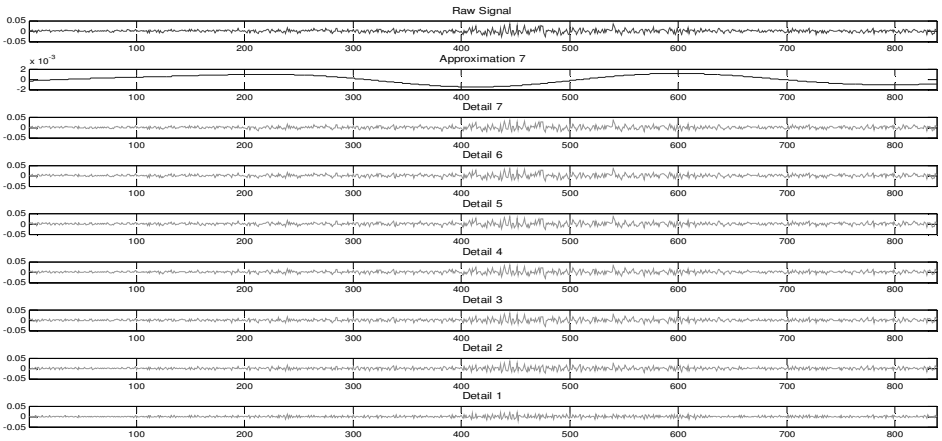
P_{Total}



P_1



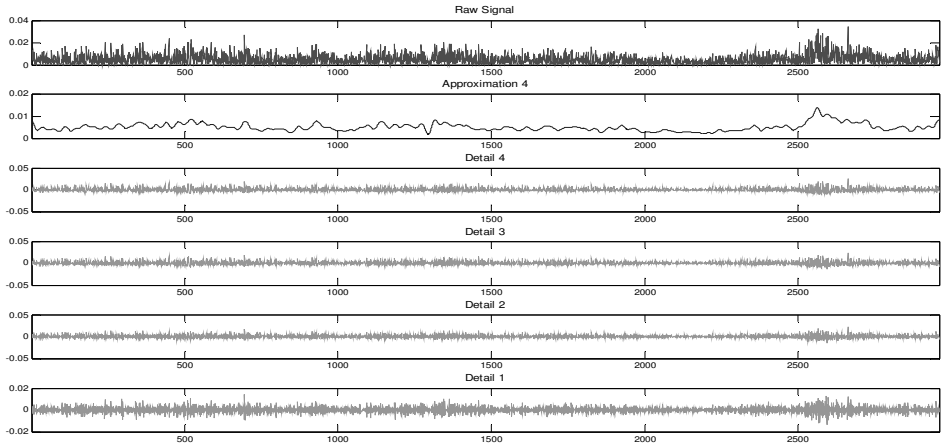
P_2



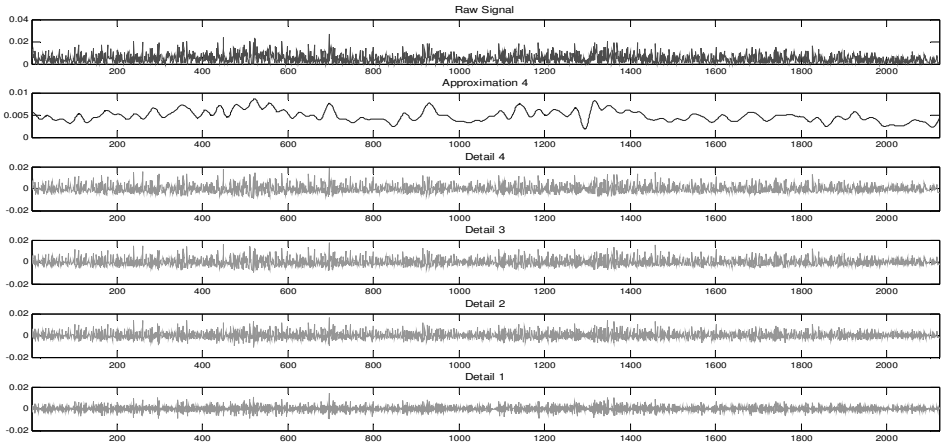
Notation: The results of SIDWT (db8) multiresolution wavelet analysis include the $D_1 - D_7$ wavelet details and the 7-th level approximation A_7. The raw signal is also displayed.

FIGURE 2: WAVELET MULTI-SCALE ANALYSIS (EUR VOLATILITY SERIES)

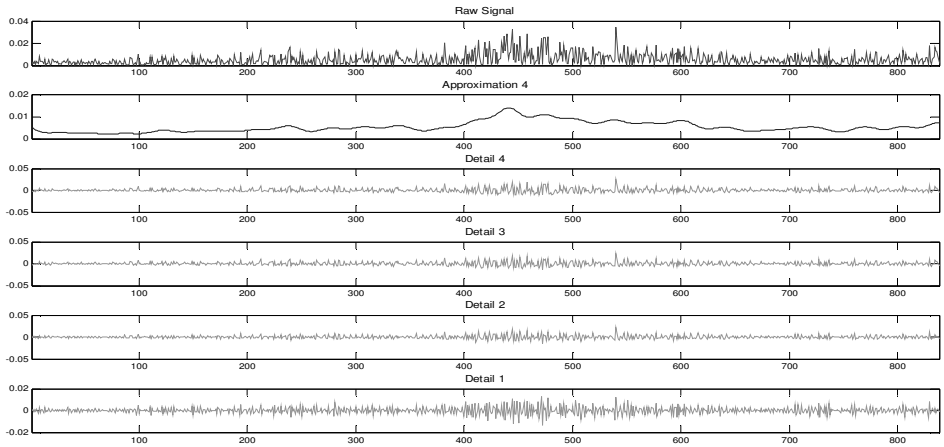
P_{Total}



P_1



P_2



Notation: The results of SIDWT (db8) multiresolution wavelet analysis include the $D_1 - D_4$ wavelet details and the 4-th level approximation A_4. The raw signal is also displayed.

FIGURE 3: DIAGRAMMATICAL REPRESENTATION OF DIRECTIONAL CAUSALITIES (RETURNS)

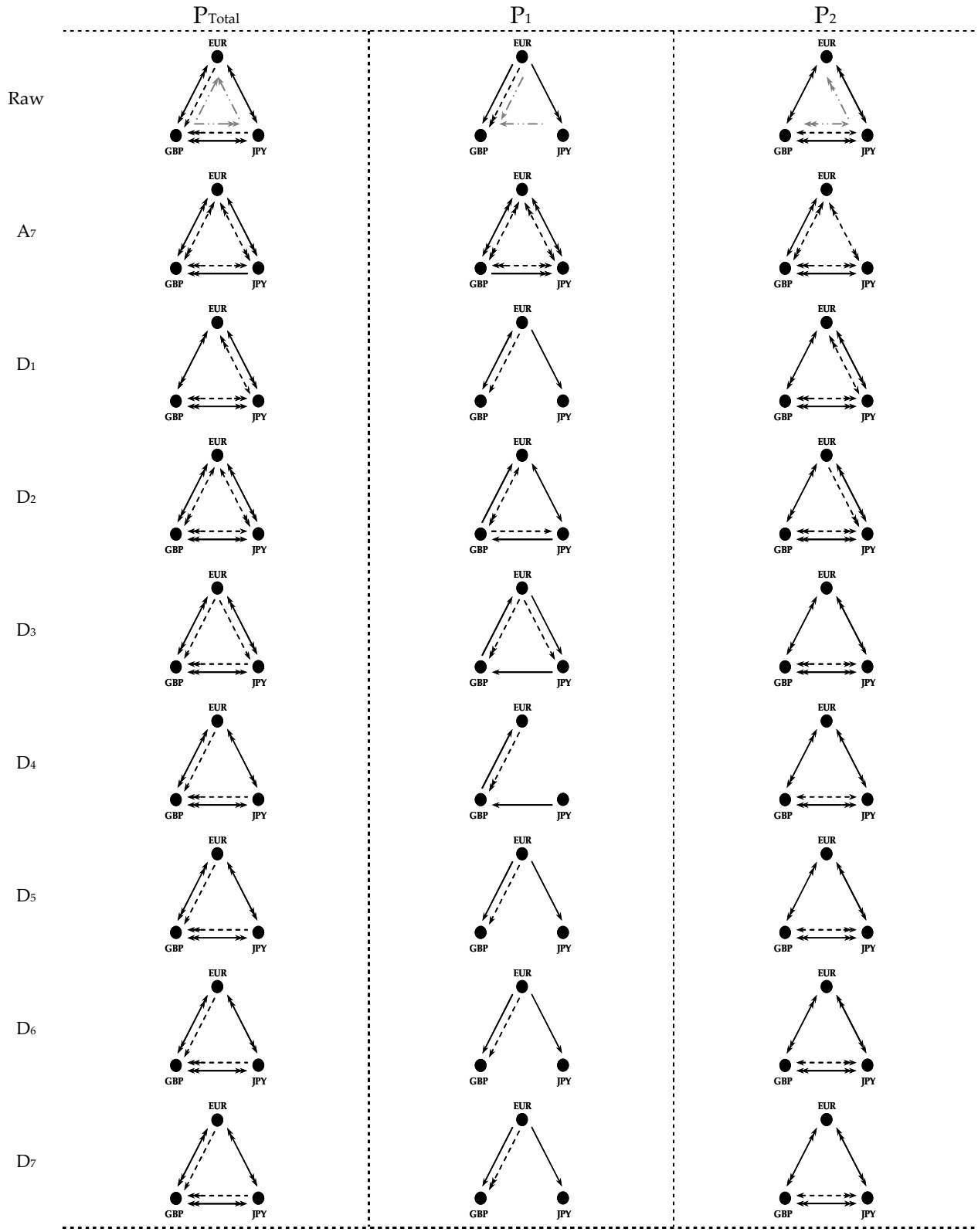
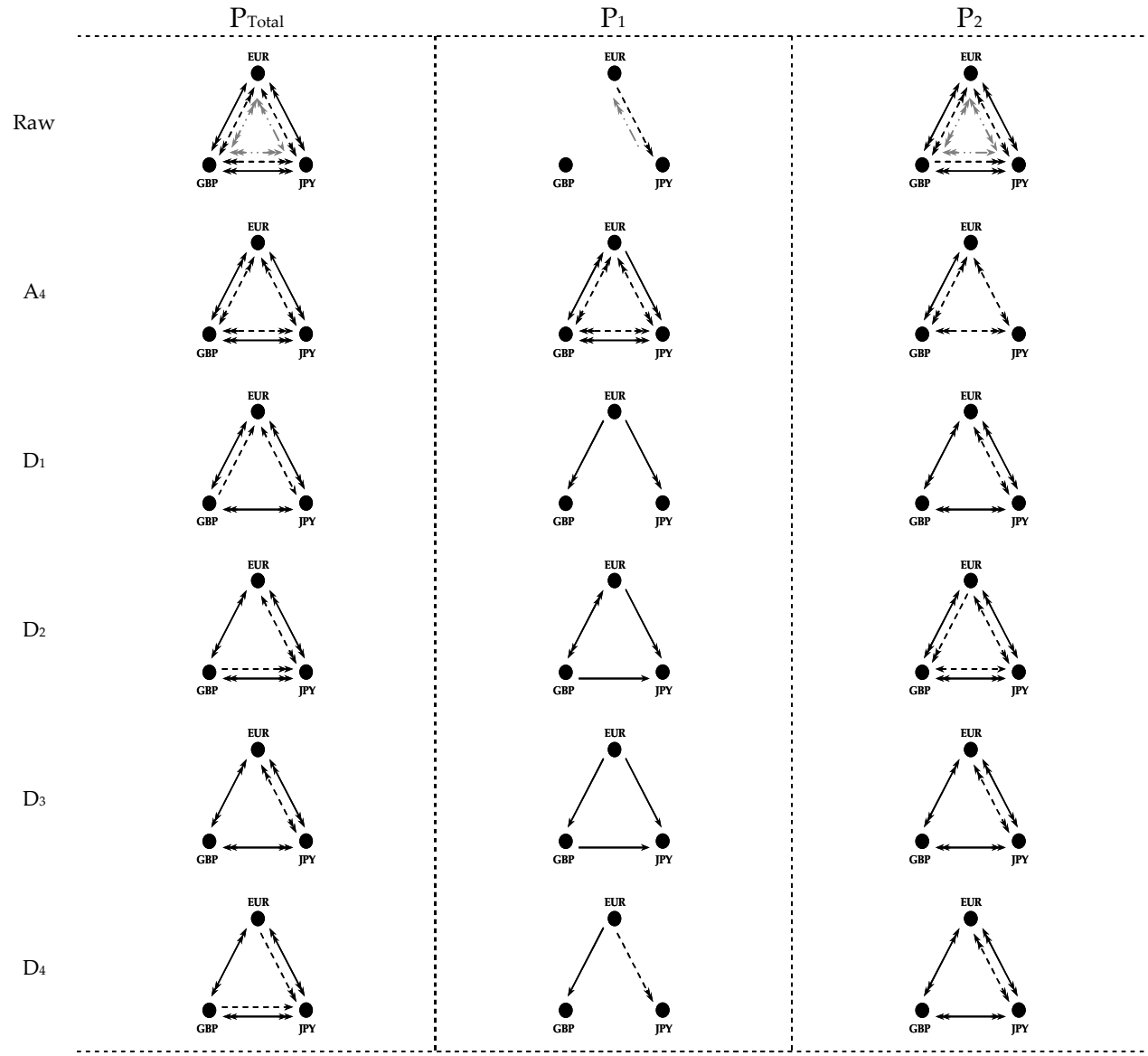


FIGURE 4: DIAGRAMMATICAL REPRESENTATION OF DIRECTIONAL CAUSALITIES (VOLATILITY)



Notation (Figure 3 and Figure 4):

The light grey line represents spectral casual relationships; the dotted and solid lines depict linear and nonlinear causal linkages respectively.

⟶ ⟷ denote unidirectional and bi-directional causality corresponding to 5% $\leq$ *p*-value < 1%

⟶» «⟷» denote unidirectional and bi-directional causality corresponding to *p*-value $\leq$ 1%

APPENDIX I: WAVELET ANALYSIS

I1. LITERATURE REVIEW

The interest in economic applications of wavelets emerged in the mid-90s, mostly stimulated by the work of Ramsey and his collaborators. Ramsey and Lampart (1998a, 1998b) used a wavelet-based scaling method to investigate the relationship and causality between money, income and expenditure. Further work on long memory processes and fractional integration in financial data can be found in Greenblatt (1998) and Jensen (2000). Davidson *et al.* (1998) used wavelets in introducing a semiparametric approach for analysing commodity prices. Wong *et al.* (2003) provided an example of using wavelets in forecasting exchange rates, wherein other conventional time series models were also used in order to compare against the wavelet-based methodology. 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 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, such as Pan and Wang (2000), Stengos and Sun (2001), Lee and Hong (2001), Hong and Kao (2004) and Fan and Gençay (2010).

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 the wavelet literature the concept of time scaling is quite different from that in economics. Based on the selected function space, the time series are analysed into "fine" and "coarse" resolution components, namely into low- and high-frequency parts of a signal respectively. 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 can be provided by narrow-support components. However, the link between

scale and frequency can be directly interpretable only when the data is stationary. In general, a multiscale decomposition could be estimated by a bank of filters with varying frequencies and widths. Yet, selecting the proper filters and their parameters so as not to discard important features of the underlying series is a difficult, subjective task, lacking a solid methodology. The wavelet analysis provides a sound mathematical framework for designing filters which eventually provide an adaptive partition of the time-frequency domain.

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 $L^2\left(0,2\pi\right)$ i.e., the set of square integrable functions in the interval $\left(0,2\pi\right)$. 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 $L^2\left(\mathbb{R}\right)$ 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. The most widely used classes of wavelets are the orthogonal ones namely the *haar* wavelets, *daubechies*, *symlets* and *coiflets* (Percival and Walden, 2000).

I2. TECHNICAL OVERVIEW

I2.1 Preliminaries

A function $\psi\left(.\right)$ that is real-valued and continuous such that $\int_{-\infty}^{\infty}\psi\left(t\right)dt=0$ and $\int_{-\infty}^{\infty}\psi\left(t\right)^2\,dt=1$ defines a wavelet. Considering that $h=\left(h_0,\ldots,h_{M-1}\right)$ is a finite wavelet filter with length M, 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 $\sum_{m=0}^{M-1}h_m^2=1$. If $g=\left(g_0,\ldots,g_{M-1}\right)$ denotes the complement low-pass (scaling) filter of the wavelet (high-pass) filter then according to Gençay *et al.* (2002) and Percival and Walden (2000), the scaling filter coefficients are estimated

based on the *quadrature mirror relationship* $g_m = \left(-1\right)^{m+1} h_{M-1-m}$ for $m = 0,...,M-1$. The properties of the scaling filter are $\sum_{m=0}^{M-1} g_m = \sqrt{2}$ and $\sum_{m=0}^{M-1} g_m^2 = 1$. The T-length vector of the wavelet coefficients $\mathbf{w}$ for a time series $\mathbf{y} = \left\{y_t\right\}_{t=1}^{T}$ 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 further decomposed into $J+1$ vectors

$$\mathbf{w} = \left[\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_J, \mathbf{s}_J\right]^T \qquad \text{(I.1)}$$

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 $\mathbf{h}_1 = \left[h_{1,T-1}, h_{1,T-2}, ..., h_{1,1}, h_{1,0}\right]^T$. If $\mathbf{h}_1$ is circularly shifted by factors of two e.g., $\mathbf{h}_1^{(2)} = \left[h_{1,1}, h_{1,0}, h_{1,T-1}, h_{1,T-2}, ..., h_{1,3}, h_{1,2}\right]^T$ etc., 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 $\mathbf{W}_1 = \left[\mathbf{h}_1^{(2)}, \mathbf{h}_1^{(4)}, ..., \mathbf{h}_1^{(T/2-1)}, \mathbf{h}_1\right]^T$. 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 of 2^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 $\mathbf{W} = \left[\mathbf{W}_1\ \mathbf{W}_2\ ...\ \mathbf{W}_J\ \mathbf{S}_J\right]^T$. From matrix $\mathbf{W}$ the wavelet filter coefficients for scales $1,...,J$ are computed via the Inverse Discrete Fourier Transform (IDFT).

I2.2 Implementation of the Discrete Wavelet Transform (DWT)

The implementation algorithm of the DWT was introduced by Mallat (1989). The time series y_t is 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 the g_1 filter. Specifically, the first step of the pyramid algorithm begins by convolving the data with each filter to obtain the wavelet $w_{1,t} = \sum_{m=0}^{M-1} h_m y_{2t+1-m \bmod T}$ and scaling coefficients $s_{1,t} = \sum_{m=0}^{M-1} g_m y_{2t+1-m \bmod T}$, $t = 0,1,...,T/2-1$. This also includes a downsampling operation. 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. By saving all wavelet coefficients and the final level of scaling coefficients the decomposition becomes $\mathbf{w} = \left[\mathbf{w}_1 \mathbf{w}_2 \mathbf{s}_2\right]^T$. This procedure is repeated up to $J = \log_2\left(T\right)$ times and provides the vector of wavelet coefficients in Eq. (I.1). 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 $s_{J-1,t} = \sum_{m=0}^{M-1} h_m w_{J,t+m \bmod 2}^0 + \sum_{m=0}^{M-1} g_m s_{J,t+m \bmod 2}^0$ 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., $y_t = \sum_{m=0}^{M-1} h_m w_{1,t+m \bmod T}^0 + \sum_{m=0}^{M-1} g_m s_{1,t+m \bmod T}^0 \;\; t = 0,1,...,T-1$.

The DWT results in the additive decomposition of the time series. Let $\mathbf{D}_j = \mathbf{W}_j^T \mathbf{w}_j$ define the *wavelet detail* corresponding to changes in the time series $\mathbf{y}$ at scale a_j for the level $j = 1,...,J$.

The coefficients $\mathbf{W}_j^T\mathbf{w}_j$ represent the part of the signal 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). Considering that for each observation, $y_t = \sum_{j=1}^{J+1} D_{j,t} \quad t = 0,\ldots,T-1$ is the linear combination of wavelet detail coefficients, then $\mathbf{A}_j = \sum_{k=j+1}^{J+1} \mathbf{D}_k$ is the cumulative sum of the variations of the details defined as the j-th level *wavelet approximation* for $0 \le j \le J$ with $\mathbf{A}_{J+1}$ being a vector of zeros. The j-th level *wavelet rough* $\mathbf{R}_j = \sum_{k=1}^{j} \mathbf{D}_k$, $1 \le j \le J+1$ incorporates the remaining lower-scale details. The time series may be decomposed as

$$\mathbf{y} = \mathbf{A}_j + \sum_{k=1}^{j} \mathbf{D}_j = \mathbf{A}_j + \mathbf{R}_j \qquad \text{(I.2)}$$

Orthonormality of the matrix $\mathbf{W}$ implies, as in the case of Discrete Fourier Transform, that the DWT is an efficient, variance preserving transform i.e., $\|\mathbf{y}\|^2 = \mathbf{y}^T\mathbf{y} = (\mathbf{W}\mathbf{w})^T\mathbf{W}\mathbf{w} = \mathbf{w}^T\mathbf{W}^T\mathbf{W}\mathbf{w} = \mathbf{w}^T\mathbf{w} = \|\mathbf{w}\|^2$. 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 $1 \le j \le J$ (due to orthonormality of $\mathbf{W}$ and $\mathbf{S}$), an equal decomposition is $\|\mathbf{y}\|^2 = \sum_{j=1}^{J} \|\mathbf{D}_j\|^2 + \|\mathbf{A}_J\|^2$.

I2.3 Wavelet classes: Haar and Daubechies

In case of a time series $\mathbf{y} = \{y_t\}_{t=1}^T$, the Haar wavelet and scaling coefficients are $w_{1,t} = \left(1/\sqrt{2}\right)\cdot\left(y_{2t} - y_{2t-1}\right)$ and $s_{1,t} = \left(1/\sqrt{2}\right)\cdot\left(y_{2t} + y_{2t-1}\right)$ respectively. Although the Haar filter is easy to visualize and implement, it is inadequate for real-world applications as it provides a poor approximation to an ideal band-pass filter (Gençay *et al.*, 2002). Instead the Daubechies wavelets improve the frequency domain characteristics of the Haar and also are compactly supported. In

general 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 $t = M/2,\ 1 + M/2,...,T/2$. For example, the Daubechies with length $M = 4$ have as wavelet coefficients $h_0 = 1 - \sqrt{3}/4\sqrt{2}$, $h_1 = -3 + \sqrt{3}/4\sqrt{2}$, $h_2 = 3 + \sqrt{3}/4\sqrt{2}$ and $h_3 = -1 - \sqrt{3}/4\sqrt{2}$.

APPENDIX II: CAUSALITY TESTS

II1. LINEAR CAUSALITY (PARAMETRIC TEST)

The linear Granger causality test (Granger, 1969) is based on a reduced-form vector autoregression (VAR) model. If $\mathbf{y}_t = \left[y_{1t},...,y_{\ell t}\right]$ is the vector of endogenous variables and ℓ the number of lags, the VAR(ℓ) model is given by

$$\mathbf{y}_t = \sum_{s=1}^{\ell} \Phi_s \mathbf{y}_{t-s} + \varepsilon_t \qquad \text{(II.1)}$$

where Φ_s is the $\ell \times \ell$ parameter matrix and ε_t the residual vector, for which $E(\varepsilon_t) = \mathbf{0}$ and $E(\varepsilon_t \varepsilon_s^{'}) = \begin{Bmatrix} \varepsilon_\varepsilon & t = s \\ \mathbf{0} & t \neq s \end{Bmatrix}$. In case of two stationary time series $\left\{x_t\right\}$ and $\left\{y_t\right\}$ the bivariate VAR model is given by

$$\begin{aligned} x_t &= \Phi(\ell)x_t + \mathrm{X}(\ell)y_t + \varepsilon_{x,t} \\ y_t &= \Psi(\ell)x_t + \Omega(\ell)y_t + \varepsilon_{y,t} \end{aligned} \qquad t = 1,2,...,N \qquad \text{(II.2)}$$

where $\Phi(\ell), \mathrm{X}(\ell), \Psi(\ell)$ and $\Omega(\ell)$ are lag polynomials with roots outside the unit circle and the error terms are i.i.d. processes with zero mean and constant variance. The test whether y strictly Granger causes x is simply a test of the joint restriction that all coefficients of the lag polynomial $\mathrm{X}(\ell)$ are zero, whilst a test of whether x strictly Granger causes y is a test regarding $\Psi(\ell)$. In the unidirectional case the null hypothesis of no Granger causality is rejected if the exclusion restriction is rejected, whereas if both $\mathrm{X}(\ell)$ and $\Psi(\ell)$ joint tests for significance are different from zero the series are bi-causally related.

However, in order to explore possible effects of cointegration a vector autoregression model in error correction form (Vector Error Correction Model-VECM) is estimated using the methodology developed by Engle and Granger (1987) and expanded by Johansen (1988) and Johansen and Juselius (1990). The bivariate VECM model has the following form

$$\begin{aligned} \Delta x_t &= -p_1\left(\left[1 \quad -\lambda\right]\cdot\left[y_{t-1} \;\; x_{t-1}\right]^T\right) + \Phi(\ell)\Delta x_t + \mathrm{X}(\ell)\Delta y_t + \varepsilon_{\Delta x,t} \\ \Delta y_t &= -p_2\left(\left[1 \quad -\lambda\right]\cdot\left[y_{t-1} \;\; x_{t-1}\right]^T\right) + \Psi(\ell)\Delta x_t + \Omega(\ell)\Delta y_t + \varepsilon_{\Delta y,t} \end{aligned} \qquad t = 1,2,...,N \qquad \text{(II.3)}$$

where $\begin{bmatrix} 1 & -\lambda \end{bmatrix}$ the cointegration row-vector and λ the cointegration coefficient. Thus, in case of cointegrated time series $\{x_t\}$ and $\{y_t\}$ linear Granger causality should be investigated on $\mathrm{X}(\ell)$ and $\Psi(\ell)$ via the VECM specification.

II2. NONLINEAR CAUSALITY (NONPARAMETRIC TEST)

Let $F\left(x_t \middle| \Theta_{t-1}\right)$ denote the conditional probability distribution of x_t given the information set Θ_{t-1}, which consists of an L_x-length lagged vector of x_t, $\mathbf{x}_{t-L_x}^{L_x} \equiv \left(x_{t-L_x}, x_{t-L_x+1}, ..., x_{t-1}\right)$ and an L_y-length lagged vector of y_t, $\mathbf{y}_{t-L_y}^{L_y} \equiv \left(y_{t-L_y}, y_{t-L_y+1}, ..., y_{t-1}\right)$. Hiemstra and Jones (1994) consider testing for a given pair of lags L_x and L_y the following null hypothesis

$$H_0 : F\left(x_t \middle| \Theta_{t-1}\right) = F\left(x_t \middle| \Theta_{t-1} - \mathbf{y}_{t-L_y}^{L_y}\right) \quad \text{(II.4)}$$

Denoting the m-length lead vector of $\mathbf{x}_t^m \equiv \left(x_t, x_{t+1}, ..., x_{t+m-1}\right)$, for $t \in \mathbf{Z}$, the claim made by Hiemstra and Jones (1994) is that the null hypothesis given in Eq. (II.4) implies for all $\varepsilon > 0$

$$\begin{aligned} P\left(\left\|\mathbf{x}_t^m - \mathbf{x}_s^m\right\| < \varepsilon \middle| \left\|\mathbf{x}_{t-L_x}^{l_x} - \mathbf{x}_{s-L_x}^{L_x}\right\| < \varepsilon, \left\|\mathbf{y}_{t-L_y}^{l_y} - \mathbf{y}_{s-L_y}^{L_y}\right\| < \varepsilon\right) \\ = P\left(\left\|\mathbf{x}_t^m - \mathbf{x}_s^m\right\| < \varepsilon \middle| \left\|\mathbf{x}_{t-L_x}^{l_x} - \mathbf{x}_{s-L_x}^{L_x}\right\| < \varepsilon\right) \end{aligned} \quad \text{(II.5)}$$

For the time series of realizations $\{x_t\}$ and $\{y_t\}$, $t = 1, ..., T$, the nonparametric test consists of choosing a value for ε typically in $[0.5,\ 1.5]$ after unit variance normalization, and testing Eq. (II.5) by expressing the conditional probabilities in terms of the corresponding ratios of joint probabilities

$$\begin{aligned} C_1\left(m + L_x, L_y, \varepsilon\right) &\equiv P\left(\left\|\mathbf{x}_{t-L_x}^{m+L_x} - \mathbf{x}_{s-L_x}^{m+L_x}\right\| < \varepsilon, \left\|\mathbf{y}_{t-L_y}^{L_y} - \mathbf{y}_{s-L_y}^{L_y}\right\| < \varepsilon\right) \\ C_2\left(L_x, L_y, \varepsilon\right) &\equiv P\left(\left\|\mathbf{x}_{t-L_x}^{L_x} - \mathbf{x}_{s-L_x}^{L_x}\right\| < \varepsilon, \left\|\mathbf{y}_{t-L_y}^{L_y} - \mathbf{y}_{s-L_y}^{L_y}\right\| < \varepsilon\right) \\ C_3\left(m + L_x, \varepsilon\right) &\equiv P\left(\left\|\mathbf{x}_{t-L_x}^{m+L_x} - \mathbf{x}_{s-L_x}^{m+L_x}\right\| < \varepsilon\right) \\ C_4\left(L_x, \varepsilon\right) &\equiv P\left(\left\|\mathbf{x}_{t-L_x}^{L_x} - \mathbf{x}_{s-L_x}^{L_x}\right\| < \varepsilon\right) \end{aligned} \quad \text{(II.6)}$$

Thus, Eq. (II.5) can be formulated as

$$\frac{C_1\left(m+L_x,L_y,\varepsilon\right)}{C_2\left(L_x,L_y,\varepsilon\right)}=\frac{C_3\left(m+L_x,\varepsilon\right)}{C_4\left(L_x,\varepsilon\right)} \qquad \text{(II.7)}$$

Using correlation-integral estimators and under the assumptions that $\{x_t\}$ and $\{y_t\}$ are strictly stationary, weakly dependent and satisfy the mixing conditions of Denker and Keller (1983), Hiemstra and Jones (1994) show that

$$\sqrt{n}\left(\frac{C_1\left(m+L_x,L_y,\varepsilon,n\right)}{C_2\left(L_x,L_y,\varepsilon,n\right)}-\frac{C_3\left(m+L_x,\varepsilon,n\right)}{C_4\left(L_x,\varepsilon,n\right)}\right)\sim N\left(0,\sigma^2\left(m,L_x,L_y,\varepsilon\right)\right) \qquad \text{(II.8)}$$

with $\sigma^2\left(m,L_x,L_y,\varepsilon\right)$ as given in their appendix. One-sided critical values are used based on this asymptotic result, rejecting when the observed value of the test statistic in Eq. (II.8) is too large.

II3. SPECTRAL CAUSALITY (FREQUENCY DOMAIN TEST)

A two-dimensional vector of time series observed at $t=1,\ldots,T$ is denoted as $\left[x_t,y_t\right]^T$. It is assumed that it has a finite-order VAR representation as $\Theta\left(L\right)\cdot\left[x_t,y_t\right]^T=\varepsilon_t$ where $\Theta\left(L\right)=I-\Theta_1L-\cdots-\Theta_pL^p$ with $L^k\cdot\left[x_t,y_t\right]^T=\left[x_{t-k},y_{t-k}\right]^T$, and that $\mathrm{E}\left(\varepsilon_t\right)=0$ and $\mathrm{E}\left(\varepsilon_t\varepsilon_t'\right)=\Sigma$ with Σ positive definite. Let Q be the lower triangular matrix of the Cholesky decomposition $Q'Q=\Sigma^{-1}$ such that $\mathrm{E}\left(\eta_t\eta_t'\right)=I$ and $\eta_t=Q\varepsilon_t$. If the system is assumed to be stationary, its MA representation is the following

$$z_t=\Xi\left(L\right)\varepsilon_t=\begin{bmatrix}\Xi_{11}\left(L\right) & \Xi_{12}\left(L\right)\\ \Xi_{21}\left(L\right) & \Xi_{22}\left(L\right)\end{bmatrix}\begin{bmatrix}\varepsilon_{1t}\\ \varepsilon_{2t}\end{bmatrix}=\Psi\left(L\right)\eta_t=\begin{bmatrix}\Psi_{11}\left(L\right) & \Psi_{12}\left(L\right)\\ \Psi_{21}\left(L\right) & \Psi_{22}\left(L\right)\end{bmatrix}\begin{bmatrix}\eta_{1t}\\ \eta_{2t}\end{bmatrix} \qquad \text{(II.9)}$$

with $\Xi\left(L\right)=\Theta\left(L\right)^{-1}$ and $\Psi\left(L\right)=\Xi\left(L\right)Q^{-1}$. Thus, the spectral density of x_t can be expressed as $S_x\left(f\right)=\frac{1}{2\pi}\left\{\left|\Psi_{11}\left(e^{-if}\right)\right|^2+\left|\Psi_{12}\left(e^{-if}\right)\right|^2\right\}$. The measure of spectral causality suggested by Geweke (1982) and Hosoya (1991) is $M_{y\to x}\left(f\right)=\log\left[2\pi S_x\left(f\right)\Big/\left|\Psi_{11}\left(e^{-if}\right)\right|^2\right]$, or equally

$$M_{y\to x}\left(f\right)=\log\left[1+\left(\left|\Psi_{12}\left(e^{-if}\right)\right|^{2}\Big/\left|\Psi_{11}\left(e^{-if}\right)\right|^{2}\right)\right] \quad \text{(II.10)}$$

The null hypothesis that y does not cause x at frequency f is given by $H_0 : M_{y\to x}\left(f\right)=0$. Yao and Hosoya (2000) suggested estimating $M_{y\to x}\left(f\right)$ by replacing $\left|\Psi_{11}\left(e^{-if}\right)\right|$ and $\left|\Psi_{12}\left(e^{-if}\right)\right|$ in Eq. (II.10) with estimates obtained from a fitted VAR model. Considering that $\gamma=vec\left(\Theta_1,\ldots,\Theta_p,\Sigma\right)$ represents the vector of parameters, the estimated causality measure is $\widehat{M}_{y\to x}\left(f\right)=M_{y\to x}\left(f\right)+D_\gamma\left(\gamma\right)'\left(\hat{\gamma}-\gamma\right)+o_p\left(T^{-1/2}\right)$, where $D_\gamma\left(\gamma\right)$ denotes the vector of derivatives of $M_{y\to x}\left(f\right)$. Under suitable regularity conditions the asymptotic distribution of the Wald statistic for the null is given by $W=T\left[\widehat{M}_{y\to x}\left(f\right)\right]^2\Big/\left(D\left(\hat{\gamma}\right)'V\left(\hat{\gamma}\right)D_\gamma\left(\hat{\gamma}\right)\right)\xrightarrow{d}\chi_1^2$, where $V\left(\hat{\gamma}\right)$ is the asymptotic covariance matrix of $\hat{\gamma}$. However, the expression $\left|\Psi_{12}\left(e^{-if}\right)\right|$ is a complicated nonlinear function of the VAR parameters and $D_\gamma\left(\hat{\gamma}\right)$ is difficult to evaluate (Yao and Hosoya, 2000).

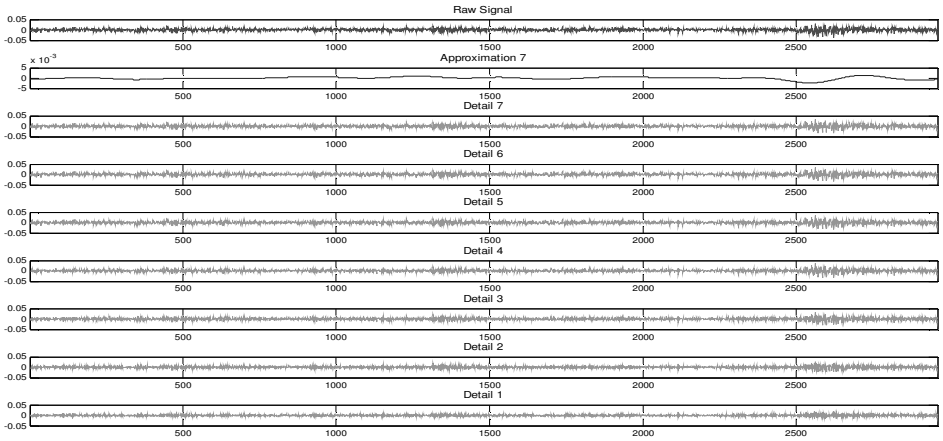
Recently, Breitung and Candelon (2006) proposed a simple approach to test the null hypothesis. Considering from Eq. (II.10) that $M_{y\to x}\left(f\right)=0$ if $\left|\Psi_{12}\left(e^{-if}\right)\right|=0$, then $\Psi\left(L\right)=\Theta\left(L\right)^{-1}Q^{-1}$ and specifically $\Psi_{12}\left(L\right)=-q^{22}\Theta_{12}\left(L\right)\Big/\left|\Theta\left(L\right)\right|$ where q_{22} is the (1,2)-element of Q^{-1} and $\left|\Theta\left(L\right)\right|$ the determinant of $\Theta\left(L\right)$. They proved that y does not cause x at frequency f if $\left|\Theta_{12}\left(e^{-if}\right)\right|=\left|\sum_{k=1}^{p}\theta_{12,k}\cos\left(kf\right)-\sum_{k=1}^{p}\theta_{12,k}\sin\left(kf\right)i\right|=0$, where $\theta_{12,k}$ is the (1,2)-element of Θ_k. Hence, a necessary and sufficient set of conditions for $\left|\Theta_{12}\left(e^{-if}\right)\right|=0$ is $\sum_{k=1}^{p}\theta_{12,k}\cos\left(kf\right)=0$ and $\sum_{k=1}^{p}\theta_{12,k}\sin\left(kf\right)=0$. By denoting $\kappa_j=\theta_{11,j}$ and $\mu_j=\theta_{12,j}$ the VAR equation for x_t is written as $x_t=\kappa_1x_{t-1}+\cdots+\kappa_px_{t-p}+\mu_1y_{t-1}+\cdots+\mu_py_{t-p}+\varepsilon_{1t}$ and thus the hypothesis $M_{y\to x}\left(f\right)=0$ is

equivalent to the following linear restriction $\mathrm{H}_0 : R(f)\mu = 0$, with $\mu = \left[\mu_1, \ldots, \mu_p\right]'$ and $R(f) = \begin{bmatrix} \cos(f) & \cos(2f) & \cdots & \cos(pf) \\ \sin(f) & \sin(2f) & \cdots & \sin(pf) \end{bmatrix}$. A cointegrating framework is also applicable if x_t is replaced by Δx_t. To study the local power they considered the simple model $x_t = b_f(L) y_{t-1} + u_t$ where the gain function of the filter $b_f(L) = \alpha\left[1 - 2\cos(f)L + L^2\right]$ is a Gegenbauer polynomial and $\{y_t\}$, $\{u_t\}$ are mutually independent white noise processes with $\mathrm{E}(y_t^2) = \sigma_y^2$ and $\mathrm{E}(u_t^2) = \sigma_u^2$. Breitung and Candelon (2006) proved that when the frequency being tested converges to the true frequency at a suitable rate under the sequence of local alternatives $f_T = f_0 + c/\sqrt{T}$, the Wald statistic for the null $\mathrm{H}_0 : R(f_0)\mu = 0$ is asymptotically distributed as non-central χ^2 with parameter $\lambda^2 = \sigma_y^2\left[2c\alpha\sin(f_0)\right]^2 \Big/ \left(\sigma_u^2\left[1 + 2\cos(f_0)^2\right]\right)$. This test is used to detect causality and to explore the short- and long-run relationships in a particular range of frequencies, which is determined by the input data frequency.

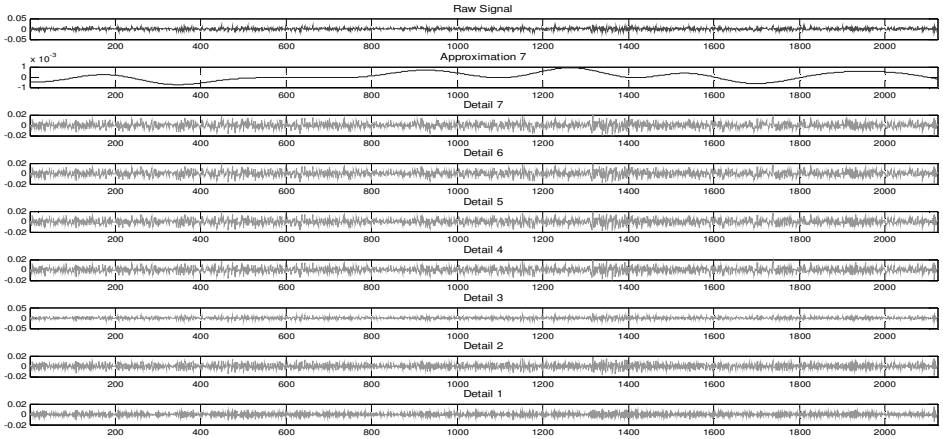
APPENDIX III: ADDITIONAL FIGURES

FIGURE III.1: WAVELET MULTI-SCALE ANALYSIS (GBP RETURNS)

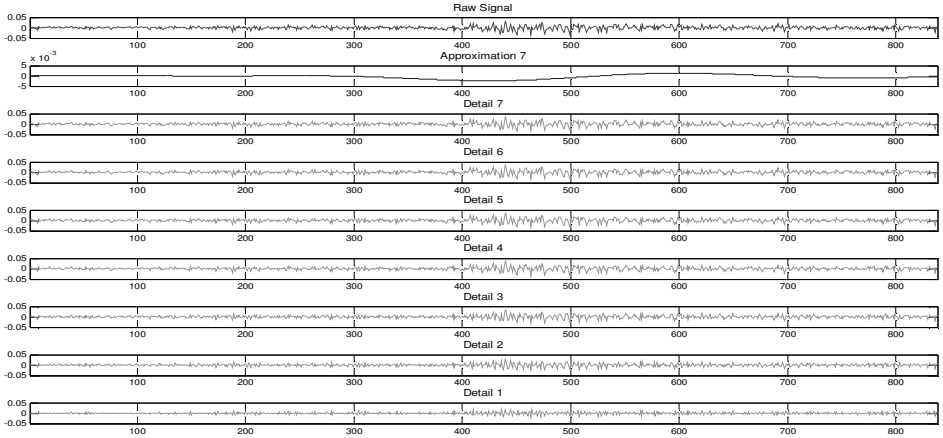
P_{Total}



P_1



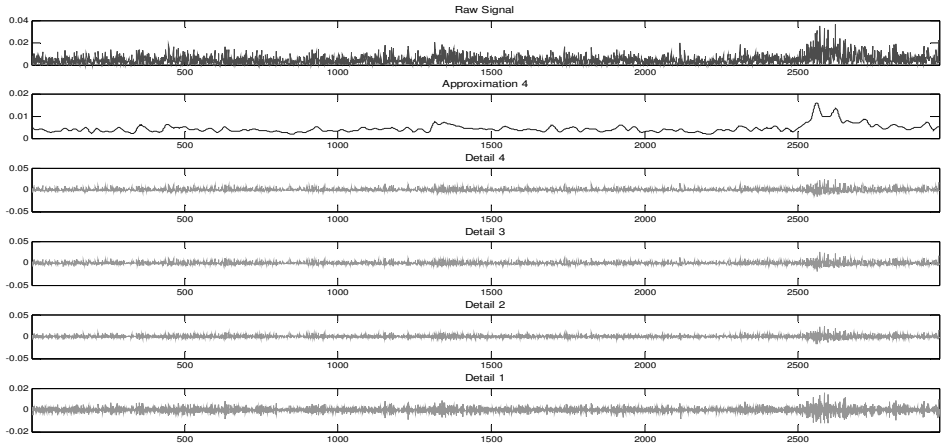
P_2



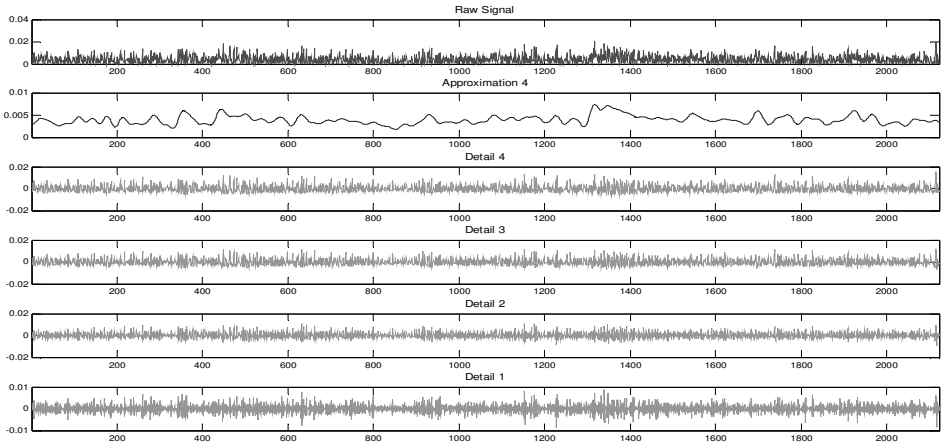
Notation: The results of SIDWT (db8) multiresolution wavelet analysis include the $D_1 - D_7$ wavelet details and the 7-th level approximation A_7. The raw signal is also displayed.

FIGURE III.2: WAVELET MULTI-SCALE ANALYSIS (GBP VOLATILITY SERIES)

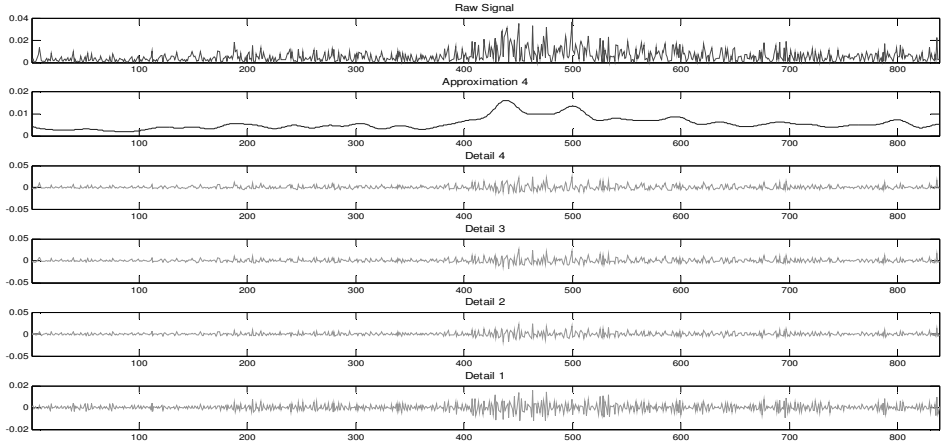
P_{Total}



P_1



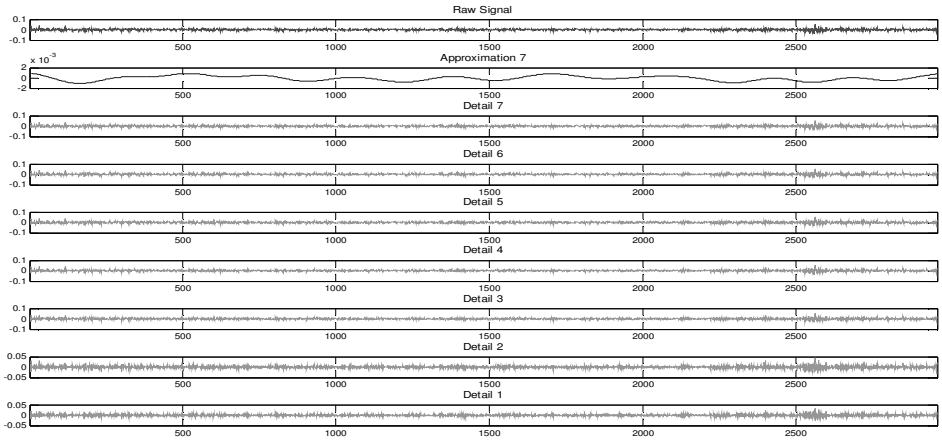
P_2



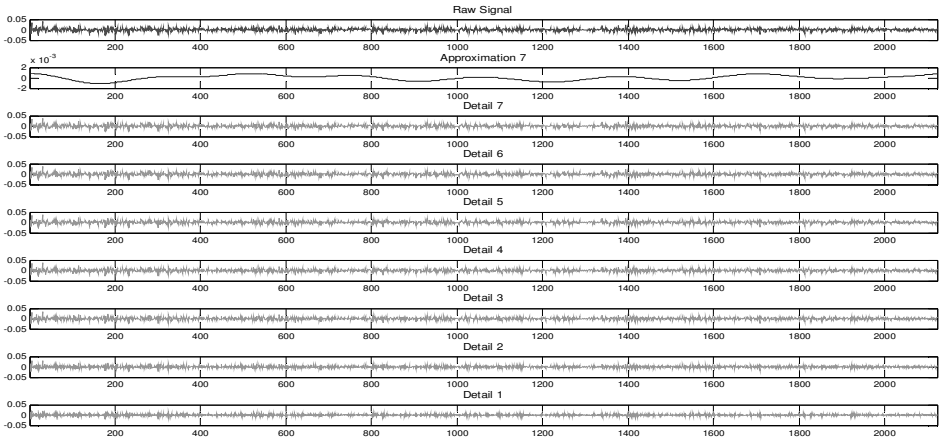
Notation: The results of SIDWT (db8) multiresolution wavelet analysis include the $D_1 - D_4$ wavelet details and the 4-th level approximation A_4. The raw signal is also displayed.

FIGURE III.3: WAVELET MULTI-SCALE ANALYSIS (JPY RETURNS)

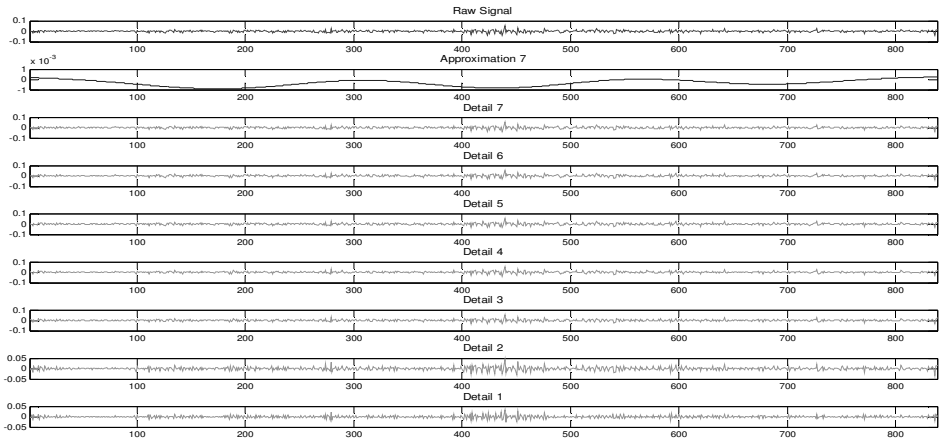
P_{Total}



P_1



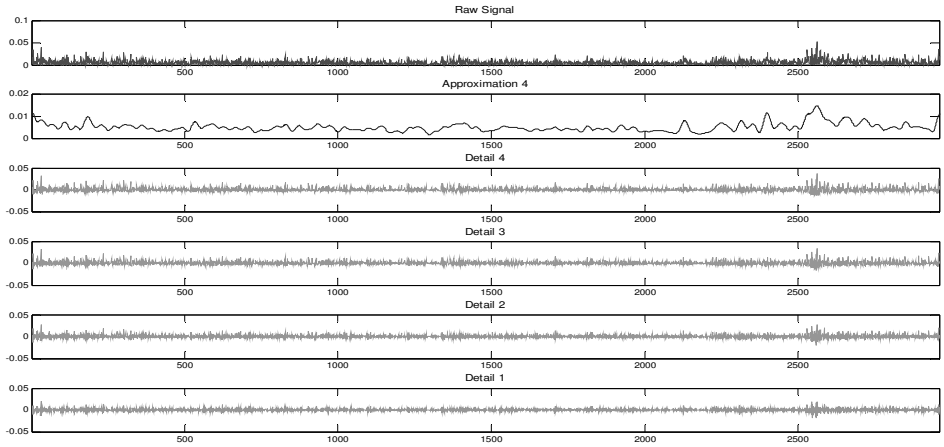
P_2



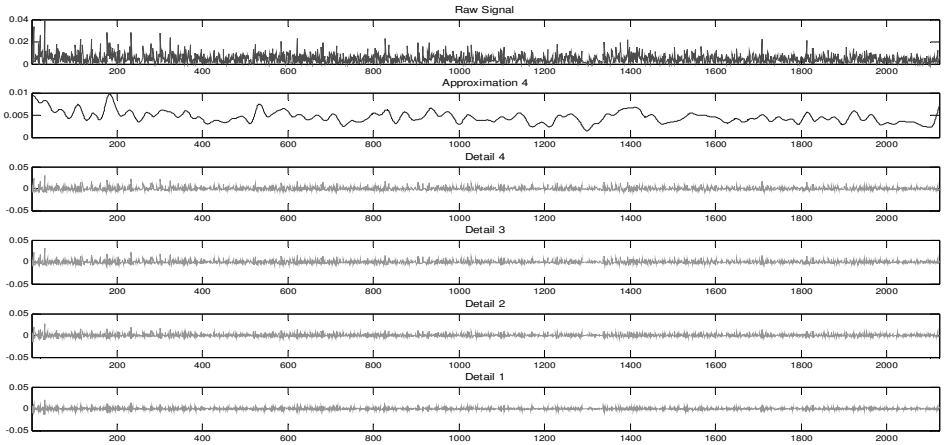
Notation: As in Figure III.1

FIGURE III.4: WAVELET MULTI-SCALE ANALYSIS (JPY VOLATILITY SERIES)

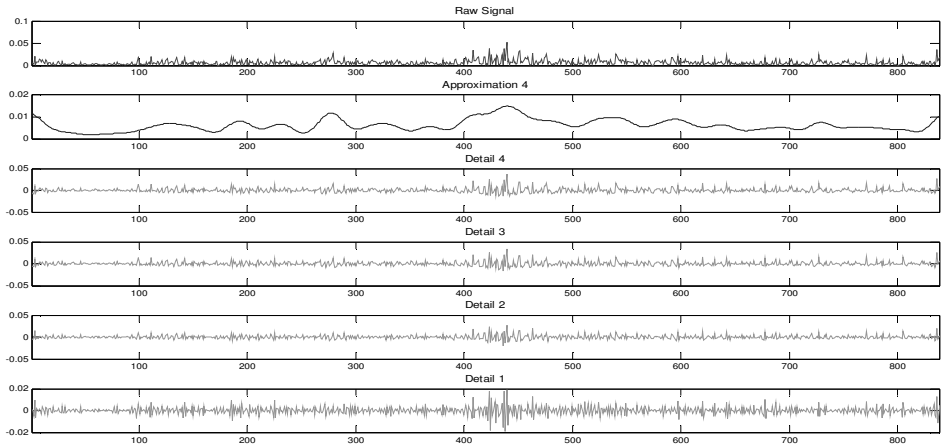
P_{Total}



P_1

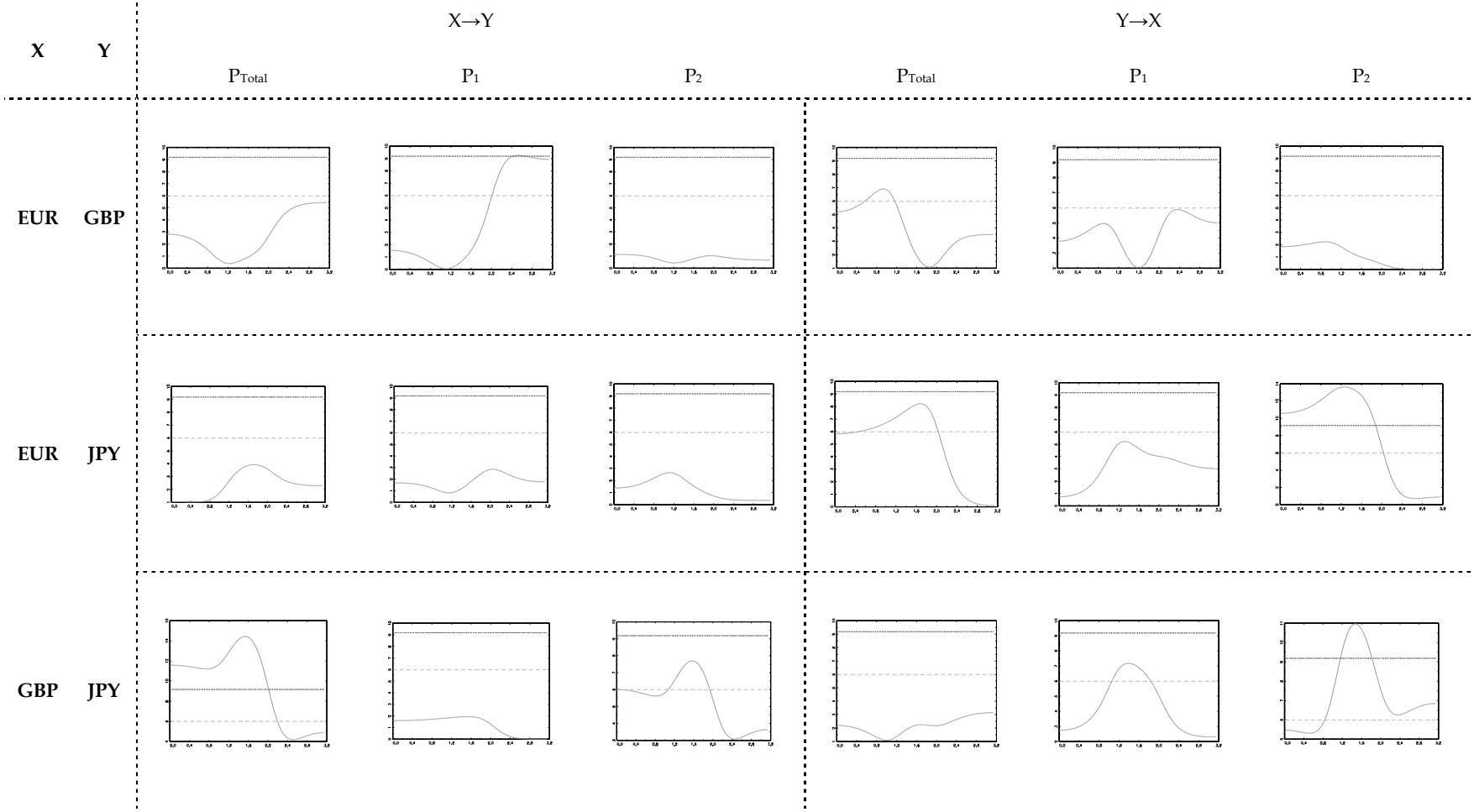


P_2



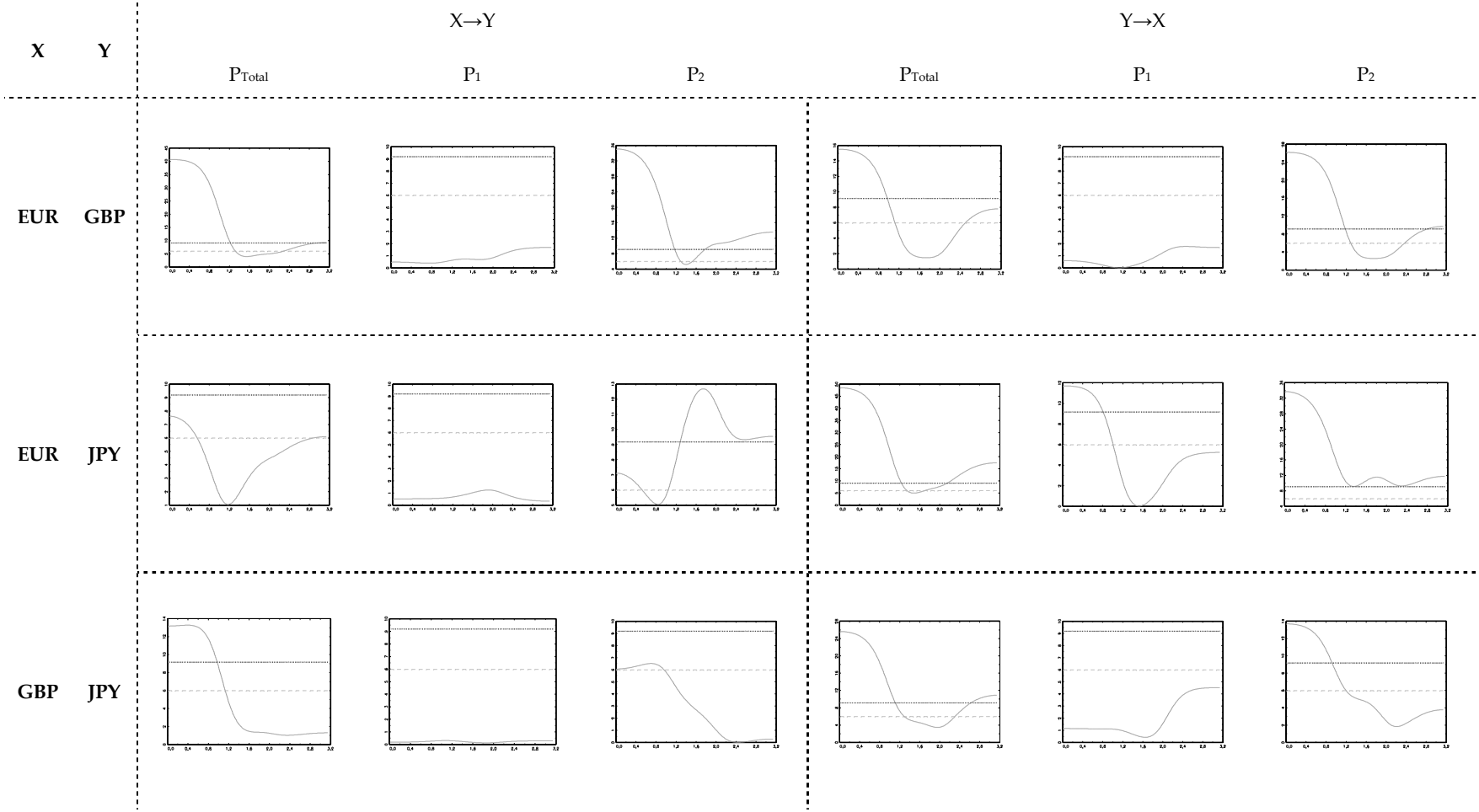
Notation: As in Figure III.2

FIGURE III.5: SPECTRAL CAUSALITY (RETURN SERIES)



Notation: The blue (dotted) line corresponds to 95% confidence level for the Breitung and Candelon (2006) frequency domain causality test, while the red (solid) to 99%.

FIGURE III.6: SPECTRAL CAUSALITY (VOLATILITY SERIES)



Notation: As in Figure III.5

REFERENCES
(APPENDIX I AND II)

Almasri, G. and A. Shukur, 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.

Baek, E. and W. Brock, 1992, A general test for non-linear Granger causality: bivariate model. *Working paper*, Iowa State University and University of Wisconsin, Madison, WI.

Breitung, J. and B. Candelon, 2006, Testing for short- and long-run causality: A frequency-domain approach. *Journal of Econometrics* 132, 363-378.

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

Davidson, R., W.C. Labys and J-B. Lesourd, 1998, Wavelet analysis of commodity price behaviour. *Computational Economics* 11(1–2), 103-128.

Denker, M., and G. Keller, 1983, On U-statistics and von-Mises statistics for weakly dependent processes. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete* 64, 505-522.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55(2), 251–276.

Fan, Y. and R. Gençay, 2010, Unit Root Tests with Wavelets. *Econometric Theory* 26, 1305-1331

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., B. Whitcher and F. Selçuk, 2002, *An introduction to wavelets and other filtering methods in finance and economics*. Academic Press, San Diego.

Geweke, J., 1982, Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association* 77, 304–324.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3), 424-438.

Greenblatt, S.A., 1998, Atomic decomposition of financial data. *Computational Economics* 12(3), 275-293.

Hiemstra, C. and J.D. Jones, 1994, Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance* 49(5), 1639-1664.

Hong, Y. and C. Kao, 2004, Wavelet-based testing for serial correlation of unknown form in panel models. *Econometrica* 72(5), 1519–1563.

Hosoya, Y., 1991, The decomposition and measurement of the interdependency between second-order stationary processes. *Probability Theory and Related Fields* 88(4), 429–444.

Jensen, M.J., 2000, An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. *Journal of Economic Dynamics and Control* 24(3), 361–387.

Johansen, S., 1988, Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12(2-3), 231–254.

Johansen, S. and K. Juselius, 1990, Maximum likelihood estimation and inference on cointegration with application to the demand for money. *Oxford Bulletin of Economics and Statistics* 52, 169–209.

Lee J. and Y. Hong, 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 Intelligence* 11(7), 674–693.

McCoy, E. J. and A.T. Walden, 1996, Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics* 5, 26-56.

Pan, X and Z. Wang, 2000, A wavelet-based nonparametric estimator of the variance function. *Computational Economics* 15(1–2), 79–87.

Percival, D.B. and A.T. Walden, 2000, *Wavelet methods for time series analysis*. Cambridge University Press, Cambridge, UK.

Priestley, M.B., 1996, Wavelets and time-dependent spectral analysis. *Journal of Time Series Analysis* 17(1), 85–103.

Ramsey, J.B. and C. Lampart, 1998a, The decomposition of economic relationships by time scale using wavelets: Money and income. *Macroeconomic Dynamics* 2, 49–71.

_____, 1998b, The decomposition of economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics & Econometrics* 3(1), 23-42.

Stengos, T. and Y. Sun, 2001, A consistent model specification test for a regression function based on nonparametric wavelet estimation. *Econometric Reviews* 20(1), 41–60.

Wong, H., I. Wai-Cheung, X. Zhongjie and X. Lui, 2003, Modelling and forecasting by wavelets, and their application to exchange rates. *Journal of Applied Statistics* 30(5), 537–553.

Yao, F. and Y. Hosoya, 2000, Inference on one-way effect and evidence in Japanese macroeconomic data. *Journal of Econometrics* 98(2), 225–255.


European
University
Institute
European University Press