Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2011-5 | March 29, 2011 | http://www.economics-ejournal.org/economics/discussionpapers/2011-5

On Some Problems in Discrete Wavelet Analysis of Bivariate Spectra with an Application to Business Cycle Synchronization in the Euro Zone

Joanna Bruzda
Nicolaus Copernicus University, Torun

Please cite the corresponding journal article:
http://dx.doi.org/10.5018/economics-ejournal.ja.2011-16

Abstract The paper considers some of the problems emerging from discrete wavelet analysis of popular bivariate spectral quantities like the coherence and phase spectra and the frequency-dependent time delay. The approach taken here, introduced by Whitcher and Craigmile (2004), is based on the maximal overlap discrete Hilbert wavelet transform (MODHWT). Firstly, we point at a deficiency in the implementation of the MODHWT and suggest using a modified implementation scheme resembling the one applied in the context of the dual-tree complex wavelet transform of Kingsbury (see Selesnick et al., 2005). Secondly, via a broad set of simulation experiments we examine small and large sample properties of two wavelet estimators of the scale-dependent time delay. The estimators are: the wavelet cross-correlator and the wavelet phase angle-based estimator. Our results provide some practical guidelines for empirical examination of short- and medium-term lead-lag relations for octave frequency bands. Besides, we show how the MODHWT-based wavelet quantities can serve to approximate the Fourier bivariate spectra and discuss certain issues connected with building confidence intervals for them. The discrete wavelet analysis of coherence and phase angle is illustrated with a scale-dependent examination of business cycle synchronization between 11 euro zone member countries. The study is supplemented with wavelet analysis of variance and covariance of the euro zone business cycles. The empirical examination underlines good localization properties and high computational efficiency of the wavelet transformations applied, and provides new arguments in favour of the endogeneity hypothesis of the optimum currency area criteria as well as a wavelet evidence on dating the Great Moderation in the euro zone.

JEL C19, E32, E58, O52

Keywords Hilbert wavelet pair; MODHWT; wavelet coherence; wavelet phase angle; business cycle synchronization

Correspondence Joanna Bruzda, Department of Logistics, Nicolaus Copernicus University, Torun, Poland; e-mail: bruzdaj@uni.torun.pl

1. Introduction

Wavelet analysis is a kind of frequency studies that enables to examine local signal properties efficiently. It is a relatively new mathematical concept with a broad range of applications in statistics, data compression and image processing. But this approach found also its place in a modern time series analysis as it makes it possible to analyze time series that are subject to structural breaks, local trends, changing cyclical patterns, outliers or show other transient characteristics. The distinguishing feature of this technique among other time-frequency methods is an endogenously varying time window, i.e. the ability to analyze short oscillations with narrow time windows and longer cycles with wider windows. Due to this, wavelet methodology is thought to constitute the next logical step in spectral analysis, one that elaborates on time localization properties of frequency methods. The methodology is known to have a significant impact in, e.g., geophysics, oceanography and medicine. However, with social sciences it is much less popular with business cycle studies becoming probably one of the most pronounced exceptions (see, among others, Jagrič, Ovin, 2004, Crowley, Lee, 2005, Raihan et al., 2005, Crowley et al., 2006, Gallegati, Gallegati, 2007, Yogo, 2008, Aguiar-Conraria, Soares, 2009, 2010).[1]

In the paper we try to answer some of the questions arising from discrete wavelet analysis of popular bivariate spectral quantities like the amplitude, phase and coherence spectra and the frequency-dependent time delay. The approach, introduced by Whitcher and Craigmile (2004), is based on a non-decimated version of the dual-tree wavelet transform of Kingsbury (1998, 2001). Following e.g. Percival and Walden (2000) we concentrate exclusively on the discrete wavelet transform (DWT) regarding it as a more natural way of handling discrete time series, especially in economics, where we often operate on frequency bands instead of a single frequency like for example in the case of a business cycle examination. In the theoretical considerations firstly, we point at a need to modify the implementation scheme of the MODHWT in a way similar to the implementation of the dual-tree complex wavelet transformation of Kingsbury (see Selesnick et al., 2005). Secondly, via a broad set of simulation experiments we examine small and large sample properties of two estimators of the wavelet time delay – a quantity measuring a causal distance between time series on a scale by scale basis. The estimators are: the wavelet cross-correlator and the wavelet phase angle-based estimator. Our results provide some practical guidelines for empirical examination of short-term lead-lag relations for octave frequency bands, pointing at a better small sample performance of the wavelet phase angle-based estimator for the first several decomposition stages and low signal-to-noise ratios. Further, we show how the wavelet quantities can serve to approximate the Fourier cross-spectrum and discuss issues connected with constructing confidence intervals for estimators of the wavelet bivariate

[1] Another promising area of applications arises in finance and includes examining comovements between financial time series, risk management as well as forecasting (see, e.g., Gençay et al., 2002, Wong et al., 2003, Fernandez, 2008, Rua, Nunes, 2009).

spectra. In our empirical analysis of business cycle variability and synchronization in the euro zone two 'continuous discrete' wavelet transformations are applied: the MODWT (the maximal overlap discrete wavelet transform) and MODHWT (the maximal overlap discrete Hilbert wavelet transform). The characteristic feature of the two transformations is that they are continuous in time and discrete in frequency (scales) in the sense that all time units and only octave frequency bands are considered in the analysis. From the point of view of an economist willing to study business cycles the MODWT and MODHWT offer the following:

- a model-free (nonparametric) approach to examining frequency characteristics of time series, i.e. short-, medium- and long-run features in the series; In particular, due to their nonparametric nature, wavelets enable to examine nonlinear processes without loss of information;
- good time-frequency resolution, and due to this, efficiency in terms of computations needed to extract the features; This enables precise examination of a time-varying frequency content of time series in an efficient way;
- decomposition of variance and covariance of stationary processes according to octave frequency bands[2]; In particular, the wavelet variance gives a simplified alternative to the spectral density function, which uses just one value per octave frequency band; The same is true for the wavelet co- and quadrature spectra, which give piecewise constant approximations to the appropriate Fourier cross-spectra on a scale by scale basis (see section 2.4);
- precise timing of shocks causing and influencing business cycles;
- low computational complexity[3];
- examination of trended, seasonal and integrated time series without prior transformations; In particular, we do not need to deseasonalize the data, as seasonal components are left automatically in lower decomposition levels, unless one is interested in examining very short cycles less than two years in length; Besides, there is no need of any prior elimination of deterministic and stochastic trends due to the fact that wavelet filtering usually embeds enough differencing operations;
- efficient estimation of short-term lead-lag relations for different frequency bands;
- global and local (short-term) measures of association for business cycle components like the wavelet correlations and cross-correlations, the wavelet coherence and the wavelet phase angle.

[2] See Percival (1995) for the variance. The covariance case is examined by Whitcher (1998) (see also Whitcher et al., 2000).

[3] The conventional DWT can be computed with an algorithm that is faster than the well known fast Fourier transform (FFT) – the Mallat's pyramid algorithm based on a mirror filters cascade and downsampling by 2, which requires only $O(N)$ multiplications. On the other hand, the computational complexity of the MODWT is $O(N\log_2 N)$ and is exactly the same as the FFT – see Percival, Walden (2000), p. 159 – while the MODHWT consumes twice more operations.

Recent studies on business cycle synchronization within the euro zone (see, e.g., de Haan et al., 2008, Gonçalves et al., 2009, and references therein) usually provide evidence in favour of the endogeneity hypothesis of the optimum currency area criteria as stated in Frankel and Rose (1998), according to which (intra-industry) trade intensification and monetary integration lead to more correlated business cycles. Our empirical examination covering 11 euro zone member countries tries to contribute to the debate by looking at synchronization patterns alone, decomposed on a scale by scale basis. The study documents a rise in synchronization between business cycles after the first steps towards European integration were taken in the second half of the 1980s. Besides, changes in business cycle variation are examined providing a new wavelet piece of evidence on dating the Great Moderation (comp. Aguiar-Conraria, Soares, 2010) and staying in agreement with the hypothesis about an early start of the process (see Blanchard, Simon, 2001).

The structure of the paper is as follows. In the next section we shortly introduce the wavelet transform in its conventional and non-decimated (maximal overlap) versions as well as the wavelet analysis of variance and covariance. Next, we present the maximal overlap discrete Hilbert wavelet transform based on the dual-tree wavelet transformation and discuss more deeply the bivariate wavelet spectral analysis, its connections with the Fourier analysis as well as implementation and statistical inference issues. Section 3 presents the results of simulation analysis comparing two wavelet methods of examining lead-lag relations for octave frequency bands, while Section 4 summarizes our empirical findings. Finally, the last section offers brief conclusions.

2. Wavelet analysis

Wavelet analysis consists in decomposing a signal into shifted and scaled versions of a basic function, $\psi(x)$, called the mother wavelet. There are different kinds of this decomposition depending on the wavelet transform applied. The continuous wavelet transform (CWT) enables to recognize local features in the data, especially in the case of signals that defined over the entire real axis, although it results in excessive redundancy of information. The discrete wavelet transform (DWT) provides a parsimonious representation of the data and is particularly useful in discrete time series processing, especially in noise reduction and information compression. The maximal overlap discrete wavelet transform (MODWT) removes certain deficiencies of the discrete transformation by considering all time units, while – similarly to the DWT – octave frequency bands are analyzed.

2.1. Conventional and maximal overlap discrete wavelet transforms[4]

The discrete wavelet transform of a real-valued function $f(x)$ is defined as follows:

[4] In this subsection we concentrate exclusively on real wavelets.

$$W_{j,t} = \int_{-\infty}^{\infty} f(x)\psi_{j,t}(x)\,dx, \qquad (1)$$

where $j = 1, 2, \ldots, J$, $t = 0, 1, \ldots, 2^{J-j} - 1$ and the wavelet daughters, $\psi_{j,t}(x)$, are shifted and scaled versions of the mother wavelet with dyadic shifts and scales, i.e.:

$$\psi_{j,t}(x) = 2^{-j/2}\psi\left(2^{-j}x - t\right). \qquad (2)$$

For certain functions $\psi(x)$ with good localization properties $\left\{\psi_{j,t}(x)\right\}$ is an orthonormal basis in $L^2(\Re)$. The function $\psi(x)$ is usually defined via another function (the scaling function or father wavelet), $\phi(x)$, that, applied to the signal after shifting and scaling analogously to (2), produces another set of coefficients in the form:

$$V_{j,t} = \int_{-\infty}^{\infty} f(x)\phi_{j,t}(x)\,dx \qquad (3)$$

known as scaling coefficients. For a given *j* the wavelet coefficients, $W_{j,t}$, are computed as differences of moving averages for the previous scale scaling coefficients and are associate with scale $\lambda_j = 2^{j-1}$, while their squares contribute to the decomposition of energy of the signal on the time-frequency plane. On the other hand, the level *j* scaling coefficients are moving averages of scale $\lambda_{j+1} = 2^j$. The two types of coefficients give the multiresolution decomposition of the original function in the form:

$$\begin{aligned} f(x) &= \sum_t V_{J,t}\phi_{J,t}(x) + \sum_t W_{J,t}\psi_{J,t}(x) + \sum_t W_{J-1,t}\psi_{J-1,t}(x) + \cdots + \sum_t W_{1,t}\psi_{1,t}(x) = \\ &= S_J(x) + D_J(x) + D_{J-1}(x) + \cdots + D_1(x). \end{aligned} \qquad (4)$$

The functions $S_j(x)$ and $D_j(x)$ are known as approximations (smooths) and details. The highest level approximation $S_J(x)$ represents smooth, low-frequency component of the signal, while the details $D_1(x)$, $D_2(x)$, …, $D_J(x)$ are associated with oscillations of length $2-4$, $4-8$, …, $2^J - 2^{J+1}$.

In filtering notation the discrete wavelet transform is defined via quadrature mirror filters: the low-pass (scaling) filter $\{g_l\}_{l=0,\ldots,L-1}$ and the high-pass (wavelet) filter $\{h_l\}_{l=0,\ldots,L-1}$, where the low-pass filter is obtained through the so-called two-scale relationship[5]. Consider a vector of length $N = 2^J$ in the form $\mathbf{x} = (x_0, x_1, \ldots, x_{N-1})'$. Then the highest possible decomposition level is *J* and the numbers of wavelet and scaling coefficients of the conventional DWT for each level are $N/2, N/4, \ldots, 1$. On the other hand, the maximal overlap discrete wavelet transform produces the same number of wavelet and scaling coefficients at each decomposition level ($\tilde{W}_{j,t}$ and $\tilde{V}_{j,t}$, accordingly) as it does not

[5] The two filters fulfill the quadrature mirror relationship $g_l = (-1)^{l+1}h_{L-1-l}$, have unit energy and are even-shift orthogonal; the wavelet filter integrates (sums) to zero, while the scaling filter – to $\sqrt{2}$.

use downsampling by 2. The coefficients are appropriately scaled in order to retain variance preservation. They are given as follows:

$$W_{j,t} = \sum_{l=0}^{L_j-1} h_{j,l}\, x_{[2^j(t+1)-1-l] \bmod N}\,, \quad t = 0,\ldots,2^{J-j}-1, \tag{5}$$

$$2^{j/2}\tilde{W}_{j,t} = \sum_{l=0}^{L_j-1} h_{j,l}\, x_{(t-l) \bmod N}, \quad t = 0,\ldots,N-1, \tag{6}$$

$$V_{j,t} = \sum_{l=0}^{L_j-1} g_{j,l}\, x_{[2^j(t+1)-1-l] \bmod N}\,, \quad t = 0,\ldots,2^{J-j}-1, \tag{7}$$

$$2^{j/2}\tilde{V}_{j,t} = \sum_{l=0}^{L_j-1} g_{j,l}\, x_{(t-l) \bmod N}, \quad t = 0,\ldots,N-1, \tag{8}$$

where $\{h_{j,l}\}$ and $\{g_{j,l}\}$ are the *j*-the level wavelet and scaling filters of length $L_j = (2^j - 1)(L-1)+1$ obtained by convolving together the following *j* filters (Percival, Walden, 2000, Chapter 4):

(A) for $\{h_{j,l}\}$:

$$\begin{aligned}
&\text{filter } 1: && g_0, g_1, \ldots, g_{L-2}, g_{L-1};\\
&\text{filter } 2: && g_0, 0, g_1, 0, \ldots, g_{L-2}, 0, g_{L-1};\\
&\text{filter } 3: && g_0, 0, 0, 0, g_1, 0, 0, 0, \ldots, g_{L-2}, 0, 0, 0, g_{L-1};\\
&\vdots\\
&\text{filter } j-1: && g_0, \underbrace{0,\ldots,0}_{2^{j-2}-1\,\text{zeros}}, g_1, \underbrace{0,\ldots,0}_{2^{j-2}-1\,\text{zeros}}, \ldots, g_{L-2}, \underbrace{0,\ldots,0}_{2^{j-2}-1\,\text{zeros}}, g_{L-1};\\
&\text{filter } j: && h_0, \underbrace{0,\ldots,0}_{2^{j-1}-1\,\text{zeros}}, h_1, \underbrace{0,\ldots,0}_{2^{j-1}-1\,\text{zeros}}, \ldots, h_{L-2}, \underbrace{0,\ldots,0}_{2^{j-1}-1\,\text{zeros}}, h_{L-1};
\end{aligned} \tag{9}$$

(B) for $\{g_{j,l}\}$:

Filters 1, …, *j*–1 as in (A)

$$\text{filter } j: \quad g_0, \underbrace{0,\ldots,0}_{2^{j-1}-1\,\text{zeros}}, g_1, \underbrace{0,\ldots,0}_{2^{j-1}-1\,\text{zeros}}, \ldots, g_{L-2}, \underbrace{0,\ldots,0}_{2^{j-1}-1\,\text{zeros}}, g_{L-1}. \tag{10}$$

$\{h_{j,l}\}$ is a bandpass filter with a nominal passband $1/2^{j+1} \le |f| \le 1/2^j$, while $\{g_{j,l}\}$ is a low-pass filter with cutoff frequency $1/2^{j+1}$. In the notation above we assume that $\{h_{1,l}\} \equiv \{h_l\}$ and $\{g_{1,l}\} \equiv \{g_l\}$. Further in the text we will also use $\tilde{h}_l = \frac{h_l}{\sqrt{2}}$, $\tilde{g}_l = \frac{g_l}{\sqrt{2}}$, $\tilde{h}_{j,l} = h_{j,l}/2^{j/2}$, $\tilde{g}_{j,l} = g_{j,l}/2^{j/2}$.

For further considerations, we provide relationships between transfer functions of the above filters with the following correspondence:

$\{h_l\} \leftrightarrow H(\cdot); \{h_{j,l}\} \leftrightarrow H_j(\cdot); \{g_l\} \leftrightarrow G(\cdot); \{g_{j,l}\} \leftrightarrow G_j(\cdot);$ and $H_1(f) \equiv H(f); G_1(f) \equiv G(f)$.

Then we have (Percival, Walden, 2000, p. 154):

$$H(f) = -e^{-i2\pi f(L-1)} G(\tfrac{1}{2} - f), \tag{11a}$$

$$H_j(f) = H(2^{j-1} f) \prod_{l=0}^{j-2} G(2^l f), \tag{11b}$$

$$G_j(f) = \prod_{l=0}^{j-1} G(2^l f). \tag{11c}$$

The same relationships also hold for the transfer functions of $\{\tilde{h}_l\}, \{\tilde{g}_l\}, \{\tilde{h}_{j,l}\}, \{\tilde{g}_{j,l}\}$, which we further denote as $\tilde{H}(\cdot), \tilde{G}(\cdot), \tilde{H}_j(\cdot), \tilde{G}_j(\cdot)$.

The reconstruction part of wavelet analysis utilizes the inverse wavelet transformation in its conventional or maximal overlap versions, what results in a sequence of details and smooths. Though the details and smooths form an additive decomposition of the signal, the lack of translation invariance of the DWT, on the one hand, and the lack of energy preservation of the MODWT details and smooths, on the other, make them somewhat less attractive in studies concerning business and growth cycle synchronization.

The distinguishing features of the conventional DWT and the MODWT can be summarized as follows (see Whitcher, 1998, Percival, Walden, 2000, Gençay et al., 2002):

- The MODWT can handle any sample size, while the *J*-th order partial DWT only multiplies of 2^J.
- The MODWT is translation invariant, which means that circularly shifting a time series is equivalent to analyzing its circularly shifted wavelet and scaling coefficients or details and smooths. This property is particularly useful in detecting singularities and examining lead-lag relations between frequency components of time series.
- Like the conventional DWT, the MODWT enables variance and covariance decomposition. But the MODWT provides a better estimator of the wavelet variance in terms of its efficiency and gives an estimator of the wavelet covariance whose variance does not depend on the true time lag between time series.
- The conventional DWT approximately decorrelates a broad range of stationary as well as nonstationary processes; Assuming that the decorrelation property of the DWT is effective, the maximal order of serial correlation of the level *j* MODWT wavelet coefficients is equal $2^j - 1$.
- The MODWT approximations and details from the multiresolution analysis are associated with zero phase filters, which makes it possible to align their features with those from the original series. Additionally, the result of the MODWT MRA is smoother and less dependent on the wavelet.
- An additive decomposition of the time series in terms of its details and approximations is valid for both the DWT an the MODWT. Contrary to the DWT, however, the MODWT details and approximations do not form an energy (and covariance) decomposition.

Among the most popular real wavelet and scaling filters are the compactly supported orthonormal Daubechies filters: the extremal phase (d*L*) and the least asymmetric (la*L*) filters. The two

families are characterized by the smallest filter length L for a given number of vanishing moments.[6] Besides, the extremal phase scaling filters have the fastest build-up of the energy sequence, while the least asymmetric filters are approximately linear phase.

2.2. Wavelet analysis of variance and covariance

For the stochastic process Y_t the time-dependent wavelet variance is defined as:

$$\sigma_t^2(\lambda_j) = \frac{1}{2\lambda_j}\mathrm{Var}(W_{j,t}) = \mathrm{Var}(\tilde{W}_{j,t}). \qquad (12)$$

Assuming that (12) does not depend on time[7], we arrive at variance decomposition according to different scales in the form:

$$\mathrm{Var}(Y_t) = \frac{1}{2}\sum_{j=1}^{\infty}\frac{1}{\lambda_j}\mathrm{Var}(W_{j,t}) = \sum_{j=1}^{\infty}\sigma^2(\lambda_j). \qquad (13)$$

The wavelet variance at level j corresponding to scale $\lambda_j = 2^{j-1}$, $\sigma^2(\lambda_j)$, informs about variation of oscillations of length approximately in the interval $2^j - 2^{j+1}$. Similarly, the wavelet covariance and wavelet correlation are introduced. For the stochastic processes X_t and Y_t, the time-varying wavelet covariance is defined as:

$$\gamma_t(\lambda_j) = \frac{1}{2\lambda_j}\mathrm{Cov}(W_{j,t}^X, W_{j,t}^Y) = \mathrm{Cov}(\tilde{W}_{j,t}^X, \tilde{W}_{j,t}^Y). \qquad (14)$$

As in the case of the variance decomposition (13), if the wavelet covariances do not depend on time, they produce decomposition of the covariance between X_t and Y_t according to different scales λ_j:

$$\mathrm{Cov}(X_t, Y_t) = \frac{1}{2}\sum_{j=1}^{\infty}\frac{1}{\lambda_j}\mathrm{Cov}(W_{j,t}^X, W_{j,t}^Y) = \sum_{j=1}^{\infty}\gamma(\lambda_j). \qquad (15)$$

Next, let us define the (time invariant) wavelet correlation coefficient for scale λ_j via:

$$\rho(\lambda_j) = \frac{\gamma(\lambda_j)}{\sigma_1(\lambda_j)\sigma_2(\lambda_j)}. \qquad (16)$$

---

[6] Roughly speaking, vanishing moments (VM) are responsible for eliminating artifacts due to the wavelet function itself as well as for the degree of approximation to an ideal bandpass filter and make it possible to interpret the filters as generalized differences of adjacent observations with the number of embedded difference operations equal the number of vanishing moments – see Daubechies (1992), §7.4, Mallat (1998), p. 166, Percival, Walden (2000), p. 483. The number of VMs for the Daubechies filters equals half the filter length.

[7] The assumption is also fulfilled for nonstationary processes provided that they are integrated of order d and the width of the wavelet filter, L, is sufficient to eliminate nonstationarity. In the case of the Daubechies wavelet filters the condition is: $L \geq 2d$ – see, e.g., Percival, Walden (2000), p. 304. Further we assume that $L > 2d$ in order to have $E\{\tilde{W}_{j,t}\} = 0$.

The quantity (16) measures the strength and direction of linear dependence between two processes for a given decomposition level j (scale λ_j). Finally, the wavelet cross-covariance and its normalized version are given as:

$$\gamma_\tau(\lambda_j) = \frac{1}{2\lambda_j}\mathrm{Cov}(W_{j,t}^X, W_{j,t+\tau}^Y) = \mathrm{Cov}(\tilde{W}_{j,t}^X, \tilde{W}_{j,t+\tau}^Y), \tag{17}$$

$$\rho_\tau(\lambda_j) = \frac{\gamma_\tau(\lambda_j)}{\sigma_1(\lambda_j)\sigma_2(\lambda_j)}. \tag{18}$$

As was mentioned in section 2.1, MODWT-based estimators have generally better statistical properties as compared to their DWT-based counterparts. Firstly, MODWT coefficients produce better estimates of the wavelet variance in terms of its efficiency; secondly, they give an estimator of the wavelet covariance whose variance does not depend on the true time lag between time series and, thirdly, decimation by 2 affects the lag-resolution of DWT-based estimators of the wavelet cross-covariances and cross-correlations, so they should not be used in practice (see Percival, Walden, 2000, 308–310, Gençay et al., 2002, p. 252–253). For these reasons, further we concentrate on estimation with the MODWT coefficients.

An unbiased estimator of the wavelet variance is defined as:

$$\hat{\sigma}^2(\lambda_j) = \frac{1}{\tilde{N}_j}\sum_{t=L_j-1}^{N-1}\tilde{W}_{j,t}^2, \tag{19}$$

where $\tilde{W}_{j,t}$ are the MODWT wavelet coefficients, $L_j = (2^j - 1)(L-1) + 1$ is the length of the wavelet filter for scale λ_j and $\tilde{N}_j = N - L_j + 1$ is the number of wavelet coefficients unaffected by the boundary.

Estimates of wavelet covariances and wavelet correlations are computed via the following formulas:

$$\hat{\gamma}(\lambda_j) = \frac{1}{\tilde{N}_j}\sum_{t=L_j-1}^{N-1}\tilde{W}_{j,t}^X\tilde{W}_{j,t}^Y, \tag{20}$$

$$\hat{\rho}(\lambda_j) = \frac{\tilde{\gamma}(\lambda_j)}{\tilde{\sigma}_1(\lambda_j)\tilde{\sigma}_2(\lambda_j)}, \tag{21}$$

while an unbiased estimate of the wavelet cross-covariance is obtained via:

$$\hat{\gamma}_\tau(\lambda_j) = \begin{cases} \dfrac{1}{\tilde{N}_j - \tau}\displaystyle\sum_{t=L_j-1}^{N-\tau-1}\tilde{W}_{j,t}^X\tilde{W}_{j,t+\tau}^Y & \text{for } \tau = 0, 1, \ldots, \tilde{N}_j - 1 \\ \dfrac{1}{\tilde{N}_j - \tau}\displaystyle\sum_{t=L_j-\tau-1}^{N-1}\tilde{W}_{j,t}^X\tilde{W}_{j,t+\tau}^Y & \text{for } \tau = -1, -2, \ldots, -(\tilde{N}_j - 1) \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Constructions of confidence intervals for the quantities described in this section are discussed by Percival (1995) and Whitcher (1998) (see also Percival, Walden, 2000, Serroukh et al., 2000, Whitcher et al., 2000, Gençay et al., 2002). All the above estimators may be based on only a portion of wavelet coefficients, which results in estimates of local versions of the wavelet quantities. A good time resolution is exactly what the non-decimated discrete wavelet transform offers and – together with certain simplifications in obtaining global spectral estimates – is the most important characteristic of the approach presented here. The same holds for the MODHWT-based quantities described in §2.4.

2.3. Maximal overlap discrete Hilbert wavelet transform

The maximal overlap discrete Hilbert wavelet transform (MODHWT) makes use of a recently introduced class of filters based on Hilbert wavelet pairs (HWP) and utilizes the non-decimated (maximal overlap) version of the dual-tree complex wavelet transform of Kingsbury.[8] The approach was advocated by Whitcher and Craigmile (2004) (see also Whitcher et al., 2005). The filters in a Hilbert wavelet pair are approximate Hilbert transforms of each other and, as in the case of the usual discrete wavelet transformation, form a basis for a collection of orthogonal bandpass filters. This time, however, the approximate analyticity of the filters enables to compute quantities with direct analogy to the appropriate bivariate Fourier spectra.[9]

Let $\{h_l^0\}$ and $\{g_l^0\}$ be conjugate quadrature mirror filters, i.e.

$$\sum_l h_l^0 = 0; \quad \sum_l (h_l^0)^2 = 1; \quad \sum_l h_l^0 h_{l+2n}^0 = 0, \quad n \neq 0; \quad g_l^0 = (-1)^{l+1} h_{L-1-l}^0 . \qquad (23)$$

The father and mother wavelets are obtained via:

$$\phi^0(t) = \sqrt{2}\sum_l g_l^0 \phi^0(2t-l); \quad \psi^0(t) = \sqrt{2}\sum_l h_l^0 \phi^0(2t-l). \qquad (24)$$

Now consider another pair of such filters: $\{h_l^1\}$ and $\{g_l^1\}$ that define another couple of father and mother wavelets: $\phi^1(t)$ and $\psi^1(t)$. We say that $\psi^1(t)$ is the Hilbert transform of $\psi^0(t)$ if:

$$\Psi^1(f) = \begin{cases} -i\Psi^0(f), & f > 0 \\ i\Psi^0(f), & f < 0 \end{cases}, \qquad (25)$$

where $\Psi^0(f)$ and $\Psi^1(f)$ are the Fourier transforms of $\psi^0(t)$ and $\psi^1(t)$, respectively. This means that the wavelets are $\frac{\pi}{2}$ out of phase with each other. The following theorem was proved by Selesnick (2001)[10]:

If transfer functions of two scaling filters fulfill the condition:

[8] For an introduction see Selesnick et al. (2005).
[9] Their continuous (in both time and scale) counterparts have been known for a longer time in applications of wavelet analysis – see, e.g., Torrence, Compo (1998).
[10] The converse of the theorem is also true, so the Selesnick's condition is both sufficient and necessary – see Selesnick et al. (2005) and references therein.

$$G^1(f) = G^0(f)e^{-i\theta(f)}, \quad |f| < 0{,}5, \qquad (26)$$

where $\theta(f) = \pi f$, then the corresponding wavelets are a Hilbert transform pair.

The condition (26) says that the digital filter $\{g_l^1\}$ should be a half-sample delayed version of $\{g_l^0\}$, i.e. $g_l^1 = g_{l-½}^0$. As a half-sample delay cannot be implemented with finite impulse response filters, only approximate solutions are available.

In what follows we use mainly the HWP filters introduced by Selesnick (2002). All the filters we apply below have the following property: the filters in the Hilbert pair are of the same length and have the same squared gain functions. In the Selesnick's so-called 'common factor approach' firstly an all-pass filter with approximately constant fractional group delay is constructed and then orthonormal filters are build via a solution to a linear system of equations and a spectral factorization. Under a specified degree of approximation (*L*) to the half-sample delay the design procedure produces short filters with a given number of vanishing moments (*K*). The length of each HWP(*K*, *L*) filter equals $2(K + L)$. In our study we apply mid-phase solutions for HWP(3, 3), HWP(4, 2), HWP(3, 5), HWP(4, 4) and denote them 'k*K*l*L*'. An alternative approach introduced by Kingsbury (2001) produces the so-called Q-shift (quarter-shift) filters which are approximately linear phase with the property that the wavelets in the Hilbert pair are mirror images of each other. For comparison purposes in our simulation analysis we also use the 6-tap Q-shift filter of Kingsbury (2001) with 1 VM and the 12-tap Q-shift filter of Tay et al. (2006) with 5 VM and denote them 'kin' and 'tkp12', respectively.

Figure 1 shows two examples of Selesnick's approximately analytic wavelets with their corresponding phase difference functions $\theta(f)$.

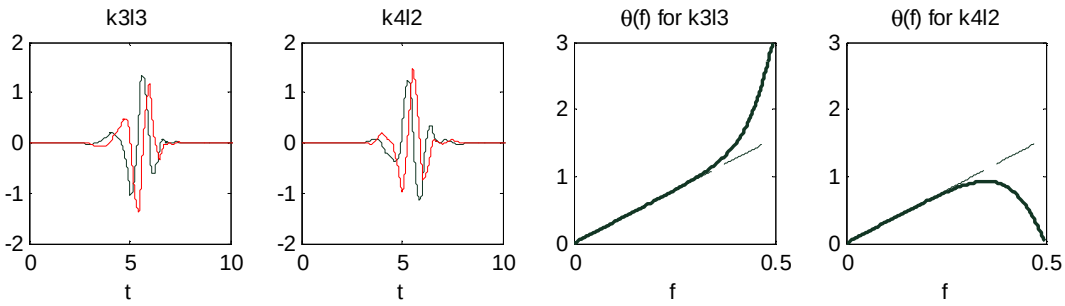

Figure 1. Example Hilbert wavelet pairs together with corresponding phase differences for scaling filters

It is clear that the functions $\theta(f)$ fulfill the condition stated in the Selesnick's theorem only for frequencies below 0.3–0.4. However, this is not of much worry as the condition concerns the approximately half-band low-pass scaling filters, while transfer functions of the wavelet filters are obtained via $H(f) = -e^{-i2\pi f(L-1)}G(\frac{1}{2} - f)$ and are not much affected either. As such, the Hilbert wavelet pairs can be seen as localized versions of cosine and sine waves forming the classic Fourier transformation.

The maximal overlap discrete Hilbert wavelet transform (MODHWT) consists in a simultaneous application of a pair of wavelet (and scaling) filters in their non-decimated (maximal overlap) forms. As a result, two sequences of coefficients are obtained, which are the real and imaginary parts of the final wavelet coefficients. In other words, the following filters are used:

$$\begin{aligned} \tilde{h}_{j,l} &= \tilde{h}^0_{j,l} + i\tilde{h}^1_{j,l}; \\ \tilde{g}_{j,l} &= \tilde{g}^0_{j,l} + i\tilde{g}^1_{j,l}, \end{aligned} \tag{27}$$

where $\tilde{h}^0_{j,l} = \frac{h^0_{j,l}}{\sqrt{2}}, \tilde{h}^1_{j,l} = \frac{h^1_{j,l}}{\sqrt{2}}, \tilde{g}^0_{j,l} = \frac{g^0_{j,l}}{\sqrt{2}}, \tilde{g}^1_{j,l} = \frac{g^1_{j,l}}{\sqrt{2}}$. These filters produce the complex wavelet and scaling coefficients in the following form:

$$\tilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_{l,j} X_{t-j} = \tilde{W}^0_{j,t} + i\tilde{W}^1_{j,t}; \tag{28}$$

$$\tilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_{l,j} X_{t-j} = \tilde{V}^0_{j,t} + i\tilde{V}^1_{j,t}. \tag{29}$$

To invert the MODHWT, its real and imaginary parts are inverted separately with the appropriate inverse maximal overlap discrete wavelet transformations. In this way two real signals are first obtained and then the signals are averaged.

As discussed in Selesnick et al. (2005), the simplest approach to combine wavelet and Hilbert transforms is via DWT post-processing. However, in such a case we end up with operations that are more computationally complex, as we operate on two parallel complex wavelet transformations. Also performing the Hilbert transform as the first one is not recommended, as then we loose the possibility to optimize it at all scales simultaneously. The dual-tree complex wavelet transform of Kingsbury is based on two real orthogonal wavelet filters with the Hilbert transform built into them. Thanks to this the Hilbert transformation automatically adapts to the wavelet scales. This feature makes the approach particularly attractive as compared to other time-frequency methods producing instantaneous amplitudes, phases and frequencies like the classic demodulation method (see, e.g., Granger, Hatanaka, 1964, Chapter 10, Priestley, 1981, §11.2.2) or the modern Hilbert-Huang transform (see Huang, Shen, 2005).

2.4. Wavelet analysis of coherence and phase angle

In this section we start with a detailed description of the relationships between the wavelet spectra obtained via the MODHWT and the appropriate Fourier spectra. The wavelet analogues of Fourier spectral characteristics of bivariate time series has been introduced by Whitcher and Craigmile (2004). Let $\tilde{W}^X_{j,t}$ and $\tilde{W}^Y_{j,t}$ be complex-valued wavelet coefficients obtained via filtering X_t and Y_t. Assuming that the wavelet filters applied have enough vanishing moments to eliminate deterministic

trend components of the series, the time-varying wavelet spectrum of (X_t, Y_t) for scale λ_j is defined as:

$$S_{XY}(\lambda_j,t) = E\left(\tilde{W}_{j,t}^{X}\,\overline{\tilde{W}_{j,t}^{Y}}\right) = E\left[\left(\tilde{W}_{j,t}^{X0} + i\tilde{W}_{j,t}^{X1}\right)\left(\tilde{W}_{j,t}^{Y0} - i\tilde{W}_{j,t}^{Y1}\right)\right] = \\ = E\left[\left(\tilde{W}_{j,t}^{X0}\tilde{W}_{j,t}^{Y0} + \tilde{W}_{j,t}^{X1}\tilde{W}_{j,t}^{Y1}\right) - i\left(\tilde{W}_{j,t}^{X0}\tilde{W}_{j,t}^{Y1} - \tilde{W}_{j,t}^{X1}\tilde{W}_{j,t}^{Y0}\right)\right] = C_{XY}(\lambda_j,t) - i\,Q_{XY}(\lambda_j,t), \quad (30)$$

where $C_{XY}(\lambda_j,t)$ and $Q_{XY}(\lambda_j,t)$ denote the time-varying wavelet cospectrum and quadrature spectrum (quad-spectrum), respectively. If the wavelet cospectrum and quadrature spectrum do not depend on time, it is possible to relate them to the appropriate Fourier quantities. Let the process (X_t, Y_t) be covariance stationary with absolute summable cross-covariance sequence. We will denote its cross-spectral density function as $S_{XY}(f) = C_{XY}(f) - i\,Q_{XY}(f)$.[11] The wavelet cospectrum for scale λ_j is then:

$$C_{XY}(\lambda_j) = E\left(\tilde{W}_{j,t}^{X0}\tilde{W}_{j,t}^{Y0} + \tilde{W}_{j,t}^{X1}\tilde{W}_{j,t}^{Y1}\right) = \int_{-1/2}^{1/2}\left|\tilde{H}_j^0(f)\right|^2 S_{XY}(f)df + \int_{-1/2}^{1/2}\left|\tilde{H}_j^1(f)\right|^2 S_{XY}(f)df\,, \quad (31)$$

where $\tilde{H}_j^0(f)$ and $\tilde{H}_j^1(f)$ are transfer functions of the scale λ_j MODWT wavelet filters $\{\tilde{h}_{j,l}^0\}$ and $\{\tilde{h}_{j,l}^1\}$. As in our case the two squared gain functions in (31) are identical, we obtain:

$$C_{XY}(\lambda_j) = 2\int_{-1/2}^{1/2}\left|\tilde{H}_j^0(f)\right|^2 S_{XY}(f)df = 2\int_{-1/2}^{1/2}\left|\tilde{H}_j^0(f)\right|^2 C_{XY}(f)df = 2\,\gamma_{XY}(\lambda_j), \quad (32)$$

where $\gamma_{XY}(\lambda_j)$ denotes the scale λ_j wavelet covariance computed with the filter $\{\tilde{h}_{j,l}^0\}$ (or, equivalently, $\{\tilde{h}_{j,l}^1\}$). Further, assuming that the wavelet filter is long enough to be considered a good approximation to an ideal bandpass filter, we have:

$$C_{XY}(\lambda_j) \approx 4\int_{1/2^{j+1}}^{1/2^j} C_{XY}(f)\,df\,. \quad (33)$$

As the quantity:

$$2^{j+1}\int_{1/2^{j+1}}^{1/2^j} C_{XY}(f)df \quad (34)$$

[11] For wavelet filters with enough vanishing moments the discussion concerning wavelet co- and quadrature spectra can be directly generalized to the case of nonstationary processes with stationary backward differences. To this end we consider two integrated processes: $X_t \sim I(d_X), Y_t \sim I(d_Y)$, whose differences of order d_X and d_Y, respectively, are jointly stationary. Then, following Whitcher and Craigmile (2004), we define $S_{XY}(f) = \dfrac{S_{WZ}(f)}{(1-e^{-i2\pi f})^{d_X}(1-e^{-i2\pi f})^{d_Y}}$, where $W_t = \Delta^{d_X}X_t$, $Z_t = \Delta^{d_Y}Y_t$. Note, however, that the generalization does not apply to the part of our analysis that utilizes the complex scaling coefficients $\tilde{V}_{j,t}$.

is the average value of $C_{XY}(f)$ in the interval $[\frac{1}{2^{j+1}}, \frac{1}{2^j}]$, we can interpret $C_j = \lambda_j C_{XY}(\lambda_j)$, $j = 1, \ldots, J$, as the average values of the Fourier cospectrum over the frequency bands $[-\frac{1}{2^j}, -\frac{1}{2^{j+1}}] \cup [\frac{1}{2^{j+1}}, \frac{1}{2^j}]$. If it is possible to assume that the Fourier cross-spectrum is piecewise constant over the octave frequency bands, estimators of C_j may serve to consistently estimate the Fourier cospectrum.[12] In any case however, the wavelet quantities discussed here will provide piecewise constant approximations to their Fourier counterparts and summarize the information included in the cross-spectrum in a way similar to the wavelet variance in the univariate spectral analysis (see Percival, 1995).

In order to obtain similar results for the wavelet quadrature spectrum, we recall the analyticity property of the Hilbert wavelet pair. The condition (25) is equivalent to:

$$\Psi^1(f) = \Psi^0(f) e^{i\xi(f)} = \begin{cases} \Psi^0(f) e^{-i\frac{\pi}{2}}, & f > 0 \\ \Psi^0(f) e^{i\frac{\pi}{2}}, & f < 0 \end{cases}. \qquad (25^*)$$

Making use of this and utilizing the approximations: $\tilde{H}_j^0(f) \approx \Psi^0(2^j f)$, $\tilde{H}_j^1(f) \approx \Psi^1(2^j f)$ (see Percival, Walden, 2000, p. 476), we obtain:

$$\begin{aligned} Q_{XY}(\lambda_j) &= E\left(\tilde{W}_{j,t}^{X0}\tilde{W}_{j,t}^{Y1} - \tilde{W}_{j,t}^{X1}\tilde{W}_{j,t}^{Y0}\right) = \int_{-1/2}^{1/2} \tilde{H}_j^0(f)\,\overline{\tilde{H}_j^1(f)}\,S_{XY}(f)df - \int_{-1/2}^{1/2} \tilde{H}_j^1(f)\,\overline{\tilde{H}_j^0(f)}\,S_{XY}(f)df \approx \\ &\approx \int_{-1/2}^{1/2} \Psi^0(2^j f)\,\overline{\Psi^1(2^j f)}\,S_{XY}(f)df - \int_{-1/2}^{1/2} \Psi^1(2^j f)\,\overline{\Psi^0(2^j f)}\,S_{XY}(f)df \approx \\ &\approx \int_{-1/2}^{1/2} \left|\Psi^0(2^j f)\right|^2 e^{-i\xi(f)} S_{XY}(f)df - \int_{-1/2}^{1/2} \left|\Psi^0(2^j f)\right|^2 e^{i\xi(f)} S_{XY}(f)df = \qquad (35) \\ &= \int_{-1/2}^{1/2} \left|\Psi^0(2^j f)\right|^2 S_{XY}(f)[-2i\sin\xi(f)]df = -2\int_{-1/2}^{1/2} \left|\Psi^0(2^j f)\right|^2 Q_{XY}(f)\sin\xi(f)df = \\ &= 2\int_0^{1/2} \left|\Psi^0(2^j f)\right|^2 Q_{XY}(f)df - 2\int_{-1/2}^{0} \left|\Psi^0(2^j f)\right|^2 Q_{XY}(f)df \approx 4\int_0^{1/2} \left|\tilde{H}_j^0(f)\right|^2 Q_{XY}(f)df. \end{aligned}$$

Unfortunately, it turns out that the first approximation assumed in (35) is of little use for the first several decomposition levels, as the approximately analytic mother wavelets $\psi^0(t)$ and $\psi^1(t)$ are most helpful in describing the behavior of the associated wavelet filters at level j as $j \to \infty$. In order to obtain more practical results, one is advised to apply different orthogonal quadrature mirror filters at the first stage of the Mallat's pyramid algorithm, namely filters approximately satisfying the

[12] Note that such an assumption is valid only for rather special kinds of relationships like the 'fixed angle lag' relationship (Granger, Hatanaka, 1964, p. 98) with constant amplitude spectrum over the frequency bands of interest or a linear regression without delay.

condition: $\tilde{g}_l^1 = \tilde{g}_{l-1}^0$ (see Selesnick et al., 2005). The condition is different (and easier to implement) than the half-sample delay requirement: $\tilde{g}_l^1 = \tilde{g}_{l-1/2}^0$.

To see that it solves the approximation problem let us consider in more detail the implementation of the MODHWT. Now we are going to discriminate between level 1 and level j (j = 2, 3, …) filters. We maintain the previous notation for the transfer functions of the remaining filters, while transfer functions of the level 1 scaling and wavelet filters will be denoted as: ${}^1\tilde{G}^0(f)$, ${}^1\tilde{G}^1(f)$, ${}^1\tilde{H}^0(f)$, ${}^1\tilde{H}^1(f)$, respectively. For the level 1 scaling filters we have:

$${}^1\tilde{G}^1(f) = e^{-i2\pi f}\ {}^1\tilde{G}^0(f). \qquad (36)$$

Using this and (11a) for the level 1 wavelet filters we obtain:[13]

$${}^1\tilde{H}^1(f) = -e^{i2\pi f}\ {}^1\tilde{H}^0(f). \qquad (37)$$

The imaginary part of the second stage wavelet filter is given as:

$$\tilde{H}_2^1(f) = \tilde{H}^1(2f)\ {}^1\tilde{G}^1(f) = -e^{-i2\pi f(2L-1)}\ \tilde{G}^1(1/2 - 2f)\ {}^1\tilde{G}^0(f). \qquad (38)$$

If the level 2 scaling filters approximately satisfy the half-sample delay condition, we further obtain the following approximation for $f > 0$:

$$\tilde{H}_2^1(f) = -e^{-i\left[4\pi f(L-1)+\frac{\pi}{2}\right]}\tilde{G}^0(1/2 - 2f)\ {}^1\tilde{G}^0(f) = e^{-i\frac{\pi}{2}}\ \tilde{H}^0(2f)\ {}^1\tilde{G}^0(f) = e^{-i\frac{\pi}{2}}\tilde{H}_2^0(f), \qquad (39)$$

while the appropriate relation for $f < 0$ is obtained via complex conjugation. So, for the second decomposition level we do not need to substitute for $\tilde{H}_j^0(f)$ and $\tilde{H}_j^1(f)$ in (35) and the approximation is valid. A similar relationship holds for all the subsequent stages. To see this let us consider the j-th decomposition level and the frequencies satisfying $|f| < \frac{1}{2^{j-1}}$. We start with writing:

$$\tilde{H}_j^1(f) = \tilde{H}^1(2^{j-1}f)\ {}^1\tilde{G}^1(f)\tilde{G}^1(2f)\tilde{G}^1(4f)\ldots\tilde{G}^1(2^{j-2}f). \qquad (40)$$

For $f > 0$ the first factor in (40) is given as:

$$\tilde{H}^1(2^{j-1}f) = -e^{-i2\pi 2^{j-1}f(L-1)}\tilde{G}^1(1/2 - 2^{j-1}f) = -e^{-i\pi\left(2^j fL - 3\cdot 2^{j-1} + \frac{1}{2}\right)}\tilde{G}^0(1/2 - 2^{j-1}f) = -e^{-i\pi\left(-2^{j-1}f + L - \frac{1}{2}\right)}$$
$$= -e^{-i\pi\left(-2^{j-1}f + L - \frac{1}{2}\right)}\tilde{H}^0(2^{j-1}f) \qquad (41)$$

and the whole expression is then:

$$\tilde{H}_j^1(f) = \tilde{H}^1(2^{j-1}f)e^{-i2\pi f}\ {}^1\tilde{G}^0(f)e^{-i2\pi f}\tilde{G}^0(2f)e^{-i4\pi f}\tilde{G}^0(4f)\ldots e^{-i2^{j-2}\pi f}\tilde{G}^0(2^{j-2}f) =$$
$$= -e^{-i\pi\left(-2^{j-1}f + L - \frac{1}{2}\right)}e^{-i2\pi f}e^{-i\pi f(2^{j-1}-2)}\tilde{H}_j^0(f) = e^{-i\frac{\pi}{2}}\tilde{H}_j^0(f). \qquad (42)$$

The first stage complex scaling filter is easily obtained via translation of any real scaling filter by one sample and using it as the imaginary part of the resulting filter. Then the complex wavelet filter is computed via the quadrature mirror relationship applied separately to these two parts. As the

[13] Here we assume that the level 1 real and imaginary filters are of the same (even) length. We change this assumption further in our computations by considering the imaginary filters to be of length L + 1, where L is even and equals the length of the real parts of the filters. This, however, does not change the result that follows.

transfer functions at the first stage of the wavelet decomposition obviously do not satisfy the analyticity property, all the quantities computed with the help of the wavelet quad-spectrum should basically be interpreted starting from the second level. However, it does not cause a problem for business cycle studies which are typically based on monthly or quarterly data.

As in the case of the cospectrum, the wavelet quad-spectrum enables to compute the average value of its Fourier analog in the interval $[\frac{1}{2^{j+1}}, \frac{1}{2^j}]$, Q_j, via the following relationship:

$$Q_{XY}(\lambda_j) \approx 4 \int_{1/2^{j+1}}^{1/2^j} Q_{XY}(f) df \overset{\text{df}}{=} \frac{Q_j}{2^{j-1}} = \frac{Q_j}{\lambda_j}. \tag{43}$$

Finally, we arrive at the following approximation for the Fourier cross-spectrum:

$$S_{XY}(f) \approx \lambda_j S_{XY}(\lambda_j) \quad \text{for} \quad f \in [\tfrac{1}{2^{j+1}}, \tfrac{1}{2^j}]. \tag{44}$$

To approximate the Fourier cross-spectrum in the interval $\left[-1/2^{J+1}, 1/2^{J+1}\right]$ we may use the complex scaling coefficients $\tilde{V}_{J,t}$ instead of the wavelet coefficients $\tilde{W}_{j,t}$. Similar computations to the given above lead to:

$$\begin{aligned} E\left(\tilde{V}_{J,t}^{X0}\tilde{V}_{J,t}^{Y0} + \tilde{V}_{J,t}^{X1}\tilde{V}_{J,t}^{Y1}\right) &= \int_{-1/2}^{1/2} \left|\tilde{G}_J^0(f)\right|^2 S_{XY}(f) df + \int_{-1/2}^{1/2} \left|\tilde{G}_J^1(f)\right|^2 S_{XY}(f) df \approx \\ &\approx 4 \int_0^{1/2^{J+1}} C_{XY}(f) df \overset{\text{df}}{=} \frac{C_{J+1}}{2^{J-1}} = \frac{C_{J+1}}{\lambda_J}, \end{aligned} \tag{45}$$

where C_{J+1} is the average value of the Fourier cospectrum in the interval $\left[-1/2^{J+1}, 1/2^{J+1}\right]$. For the imaginary part of the spectrum firstly we notice that from (11c), (26) and (36) we have: $\tilde{G}_J^1(f) = \tilde{G}_J^0(f) e^{-i\pi f 2^J}$. Then we obtain:

$$\begin{aligned} E\left(\tilde{V}_{J,t}^{X0}\tilde{V}_{J,t}^{Y1} - \tilde{V}_{J,t}^{X1}\tilde{V}_{J,t}^{Y0}\right) &= \int_{-1/2}^{1/2} \tilde{G}_J^0(f)\overline{\tilde{G}_J^1(f)} S_{XY}(f) df - \int_{-1/2}^{1/2} \tilde{G}_J^1(f)\overline{\tilde{G}_J^0(f)} S_{XY}(f) df \approx \\ &\approx 4 \int_0^{1/2^{J+1}} Q_{XY}(f) \sin \pi f 2^J df = \frac{Q_{J+1}}{\pi 2^{J-2}}, \end{aligned} \tag{46}$$

where in the last equality it is assumed that the value of the Fourier quad-spectrum in the interval $[0, \frac{1}{2^{J+1}}]$ is constant and equal Q_{J+1}.

Next, as in Whitcher and Craigmile (2004), we consider the time-varying wavelet cross-amplitude spectrum:

$$A_{XY}(\lambda_j, t) = \left|S_{XY}(\lambda_j, t)\right| = \left[C_{XY}^2(\lambda_j, t) + Q_{XY}^2(\lambda_j, t)\right]^{1/2}, \tag{47}$$

the time-varying wavelet phase spectrum (wavelet phase angle):

$$\theta_{XY}(\lambda_j, t) = \operatorname{atan}\left[\frac{-Q_{XY}(\lambda_j, t)}{C_{XY}(\lambda_j, t)}\right], \tag{48}$$

and the time-varying wavelet magnitude squared coherence (MSC)[14]:

$$K_{XY}(\lambda_j,t)=\frac{A_{XY}^2(\lambda_j,t)}{S_X(\lambda_j,t)S_Y(\lambda_j,t)}=\frac{C_{XY}^2(\lambda_j,t)+Q_{XY}^2(\lambda_j,t)}{S_X(\lambda_j,t)S_Y(\lambda_j,t)}, \qquad (49)$$

where $S_X(\lambda_j,t)=E\left|\tilde{W}_{j,t}^X\right|^2$, $S_Y(\lambda_j,t)=E\left|\tilde{W}_{j,t}^Y\right|^2$ denote the time-varying wavelet spectra equal two times the wavelet variance, $\sigma_t^2(\lambda_j)$. The Schwartz inequality for complex random variables guaranties that $0\le K_{XY}(\lambda_j,t)\le 1$.

Let us consider some simple examples of stationary bivariate processes. We start with a linear regression without delay in the form:

$$Y_t=\alpha X_t+\eta_t, \qquad (50)$$

where X_t and η_t are stationary processes, uncorrelated with each other at all leads and lags. Then the appropriate Fourier quantities are as follows (comp., for example, Priestley, 1981, p. 663-664):

$$C_{XY}(f)=A_{XY}(f)=\alpha S_X(f),\ Q_{XY}(f)=\theta_{XY}(f)=0,\ K_{XY}(f)=\frac{C_{XY}^2(f)+Q_{XY}^2(f)}{S_X(f)S_Y(f)}=\alpha^2\frac{S_X(f)}{S_Y(f)}.$$

The corresponding wavelet quantities are the following:

$$C_{XY}(\lambda_j)=2\int_{-1/2}^{1/2}\alpha\left|H_j^0(f)\right|^2 S_X(f)df=2\alpha\sigma_X^2(\lambda_j), \qquad (51)$$

$$Q_{XY}(\lambda_j)=\theta_{XY}(\lambda_j)=0,\ A_{XY}(\lambda_j)=2\alpha\sigma_X^2(\lambda_j), \qquad (52)$$

$$K_{XY}(\lambda_j)=\frac{C_{XY}^2(\lambda_j)+Q_{XY}^2(\lambda_j)}{S_X(\lambda_j)S_Y(\lambda_j)}=\frac{\left[2\alpha\sigma_X^2(\lambda_j)\right]^2}{\left[2\sigma_X^2(\lambda_j)\right]\left[2\sigma_Y^2(\lambda_j)\right]}=\alpha^2\frac{\sigma_X^2(\lambda_j)}{\sigma_Y^2(\lambda_j)}. \qquad (53)$$

As we can see, assuming that the individual Fourier spectra are approximately piecewise constant over octave frequency bands, the wavelet coherence will provide a good approximation to the appropriate Fourier quantity.

In the next example we consider a linear regression with delay in the form:

$$Y_t=\alpha X_{t-\tau}+\eta_t, \qquad (54)$$

where, as previously, X_t and η_t are stationary and uncorrelated. In this case we have:

$$C_{XY}(f)=\alpha\cos(2\pi f\tau)S_X(f),\ Q_{XY}(f)=\alpha\sin(2\pi f\tau)S_X(f),\ A_{XY}(f)=\alpha S_X(f),$$

$$K_{XY}(f)=\alpha^2\frac{S_X(f)}{S_Y(f)},\ \theta_{XY}(f)=-2\pi f\tau$$

[14] We will refer to it as wavelet coherence or wavelet coherence spectrum. Its square root is called the wavelet coherency.

and the time delay defined as: $\tau_{XY}(f) \overset{\text{df}}{=} -\frac{\theta_{XY}(f)}{2\pi f}$ equal $\tau_{XY}(f) = \tau$. Then, assuming that the wavelet transformation produces a bandpass white noise, the wavelet co-, quad- and amplitude spectra are as follows:

$$C_{XY}(\lambda_j) = 2\int_{-1/2}^{1/2} \alpha \left|\tilde{H}_j^0(f)\right|^2 \cos(2\pi f\tau) S_X(f) df \approx \alpha 2^{j+2} \sigma_X^2(\lambda_j) \int_{1/2^{j+1}}^{1/2^j} \cos(2\pi f\tau) df = \\ = \frac{\alpha 2^{j+2} \sigma_X^2(\lambda_j)}{\pi\tau} \cos\frac{3\pi\tau}{2^{j+1}} \sin\frac{\pi\tau}{2^{j+1}}, \tag{55}$$

$$Q_{XY}(\lambda_j) = 2\int_{-1/2}^{1/2} \alpha \left|H_j^0(f)\right|^2 \sin(2\pi f\tau) S_X(f) df \approx \alpha 2^{j+2} \sigma_X^2(\lambda_j) \int_{1/2^{j+1}}^{1/2^j} \sin(2\pi f\tau) df = \\ = \frac{\alpha 2^{j+2} \sigma_X^2(\lambda_j)}{\pi\tau} \sin\frac{3\pi\tau}{2^{j+1}} \sin\frac{\pi\tau}{2^{j+1}}, \tag{56}$$

$$A_{XY}(\lambda_j) \approx \frac{\alpha 2^{j+2} \sigma_X^2(\lambda_j)}{\pi\tau} \sin\frac{\pi\tau}{2^{j+1}} = 2\alpha\sigma_X^2(\lambda_j) \frac{\sin\frac{\pi\tau}{2^{j+1}}}{\frac{\pi\tau}{2^{j+1}}} \approx 2\alpha\sigma_X^2(\lambda_j), \tag{57}$$

where the last approximation takes place for high enough decomposition levels j[15]. The wavelet coherence is given as:

$$K_{XY}(\lambda_j) \approx \frac{\alpha^2 \sigma_X^2(\lambda_j)}{\sigma_Y^2(\lambda_j)} \frac{\sin^2\frac{\pi\tau}{2^{j+1}}}{\left(\frac{\pi\tau}{2^{j+1}}\right)^2} \tag{58}$$

and, as previously, for higher decomposition levels can be approximated via:

$$K_{XY}(\lambda_j) \approx \frac{\alpha^2 \sigma_X^2(\lambda_j)}{\sigma_Y^2(\lambda_j)}. \tag{59}$$

The wavelet phase spectrum is:

$$\theta_{XY}(\lambda_j) = \operatorname{atan}\left[\frac{-Q_{XY}(\lambda_j)}{C_{XY}(\lambda_j)}\right] \approx -\frac{3}{2^{j+1}}\pi\tau. \tag{60}$$

In order to obtain a wavelet estimator of the parameter τ, we also introduce a quantity called the wavelet time delay. It is defined as:

$$\tau_{XY}(\lambda_j) \overset{\text{df}}{=} -\frac{\theta_{XY}(\lambda_j)}{2\pi f_{j0}}, \tag{61}$$

[15] Assuming that $\sin x \approx x$ for $x \le \frac{\pi}{8}$ the last approximation will work for $\tau \le 2^{j-2}$ (j = 2, 3, …) – comp. Percival, Walden (2000), p. 344.

where f_{j0} is the center frequency of the octave band, computed as the arithmetic mean of the upper and lower cutoff frequencies, i.e. $f_{j0} = \frac{3}{2^{J+2}}$. Then in our example we have: $\tau_{XY}(\lambda_j) \approx \tau$.

A slightly more general situation arises in the case of the so-called time delay estimation (TDE) problem described as follows. Let us assume that X_t and Y_t are two spatially separated sensor measurements of an unobserved signal S_t that satisfy:

$$\begin{aligned} X_t &= S_t + \eta_{Xt}, \\ Y_t &= \alpha S_{t-\tau} + \eta_{Yt}, \end{aligned} \tag{62}$$

where S_t, η_{Xt} and η_{Yt} are stationary and mutually uncorrelated at all leads and lags. Then:

$$\mathrm{Cov}(X_t, Y_{t+k}) = \alpha \,\mathrm{Cov}(S_t, S_{t+k-\tau}) = \alpha \,\mathrm{Cov}_S(k-\tau)$$

and

$$S_{XY}(f) = \alpha e^{-i2\pi f\tau} S_S(f).$$

From this we obtain that the phase spectrum is as in the previous example, i.e. $\theta_{XY}(f) = -2\pi f\tau$. Similarly, $\theta_{XY}(\lambda_j) \approx -\frac{3}{2^{j+1}}\pi\tau$ and the wavelet time delay is $\tau_{XY}(\lambda_j) \approx \tau$.

Estimators of (47)–(49) and (61) are obtained by replacing the wavelet cospectrum and quad-spectrum as well as the wavelet individual spectra with their estimates computed via smoothing in time. The smoothing is necessary, in particular in estimation of the wavelet coherence. In Whitcher and Craigmile (2004) a simple two-sided moving average is suggested and this is the approach taken here as well.

Figures 2 and 3 below present mean estimates of the wavelet coherence, phase spectrum and time delay for samples ranging from 10 to 500 wavelet coefficients not affected by the boundary, obtained with the k4l2 HWP filter for the linear regression with delay (54) with $\tau = 1$. Figure 2 illustrates the case, when the first stage filters are different and fulfill the one-sample delay condition. The la12 Daubechies filter was applied in the real part of the first stage complex filter. For comparison purposes the appropriate results obtained without this modification are also presented (see Figure 3). As we can see, the modified procedure gives acceptable results starting from scale 2, while the simplified method introduces a more bias in both the coherence and time delay estimation, especially at the second decomposition level. Another observation concerns the small sample bias of the wavelet coherence estimator. The bias clearly increases with the scale.

To construct confidence intervals for the wavelet coherence the multivariate process in the form:

$$\mathbf{P}_{j,t} = \left[\left|\tilde{W}^X_{j,t}\right|^2, \left|\tilde{W}^Y_{j,t}\right|^2, \Re\left(\tilde{W}^X_{j,t}\,\overline{\tilde{W}^Y_{j,t}}\right), \Im\left(\tilde{W}^X_{j,t}\,\overline{\tilde{W}^Y_{j,t}}\right) \right]^T \tag{63}$$

is considered together with the function:

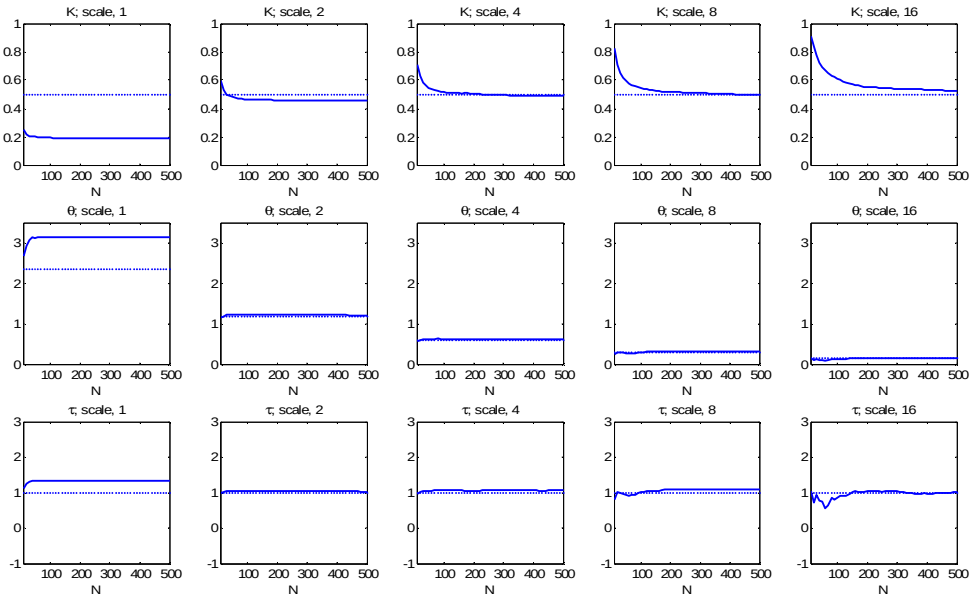

Figure 2. Estimates of the wavelet coherence, wavelet phase spectrum and wavelet time delay with the modified method; the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k4l2; figures present the theoretical Fourier quantities (thin dotted lines) together with the mean estimates of the corresponding wavelet quantities obtained with 500 replications for samples consisting of 10, 20, …, 500 wavelet coefficients unaffected by the boundary for the linear regression model with delay τ = 1, α = 1 and X_t and η_t being two independent AR(1) processes with autoregressive parameters 0.8 and unit error variances (thick solid lines).

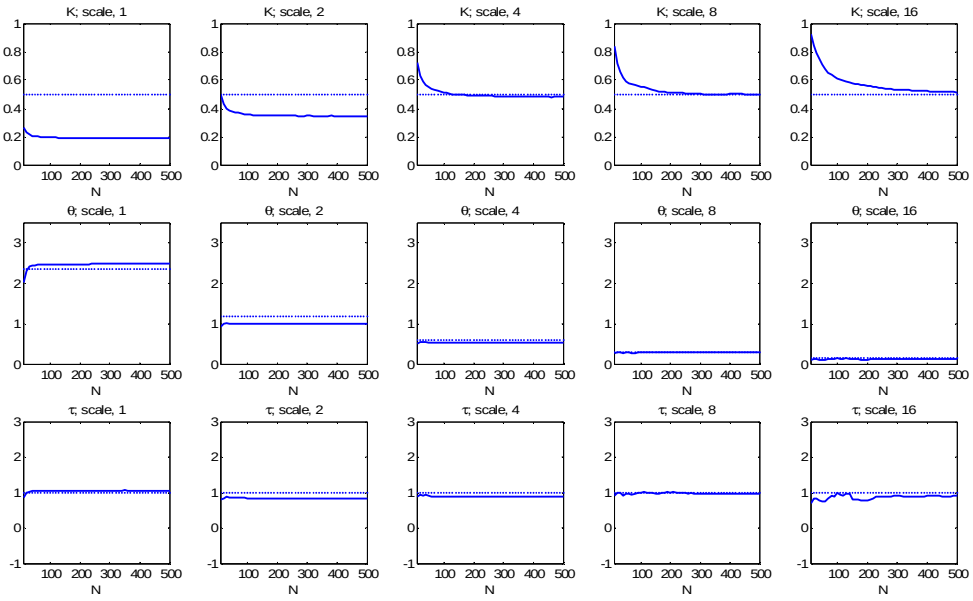

Figure 3. Estimates of the wavelet coherence, wavelet phase spectrum and wavelet time delay with the simplified method; k4l2 HWP filter is applied at each decomposition level; see detailed description below Figure 2.

$$g\left([a,b,c,d]^T\right)=\frac{c^2+d^2}{ab}. \tag{64}$$

Then, assuming that $K_{XY}(\lambda_j)>0$ and applying the delta method, the following result holds (Whitcher, Craigmile, 2004):

$$\sqrt{\widetilde{N}_j}\left(\hat{K}_{XY}(\lambda_j)-K_{XY}(\lambda_j)\right)\sim AN(0,R_{abc,j}(0)), \tag{65}$$

where $R_{abc,j}(0)=\nabla g(\mathbf{P}_{j,t})^T\cdot S_{abcd,j}(0)\cdot\nabla g(\mathbf{P}_{j,t})$, $S_{abcd,j}(\cdot)$ is the 4×4 spectral matrix for $\mathbf{P}_{j,t}$ and:

$$\nabla g\left([a,b,c,d]^T\right)=\left[-\frac{c^2+d^2}{a^2b},-\frac{c^2+d^2}{ab^2},\frac{2c}{ab},\frac{2d}{ab}\right]^T.$$

As suggested by Whitcher and Craigmile (2004), an estimate of the large sample variance $R_{abc,j}(0)/\widetilde{N}_j$ of $\hat{K}_{XY}(\lambda_j)$ may be obtained by replacing $\nabla g(\mathbf{P}_{j,t})$ with $\nabla g(\hat{\mathbf{P}}_{j,t})$ and $S_{abcd,j}(0)$ with an estimate utilizing sample covariances of elements of the vector process $\mathbf{P}_{j,t}$. Then, an approximate large sample CI for the wavelet coherence is:

$$\hat{K}_{XY}(\lambda_j)\pm\varsigma_{\frac{\alpha}{2}}\left(\frac{\hat{R}_{abcd,j}(0)}{\widetilde{N}_j}\right)^{0,5}, \tag{66}$$

where $\varsigma_{\frac{\alpha}{2}}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

Due to the analytic complexity, when testing for significance of the wavelet coherence bootstrap methods are usually applied. Both parametric and nonparametric bootstrap techniques are recommended. For example, Whitcher et al. (2005) use the block bootstrap method, while Aguiar-Conraria and Soares (2010) suggest bootstrapping based on ARMA models either in the parametric or nonparametric setting, i.e. assuming or not a particular distribution for residuals. It is worth noticing that asymptotic results for significance tests were also derived, though they concern only certain specific wavelet families – see Ge (2008), Cohen, Walden (2010a), (2010b).

To construct confidence intervals for the wavelet phase angle we assume that $C_{XY}(\lambda_j)\neq0$ and take:

$$g\left([a,b,c,d]^T\right)=\tan^{-1}\left(\frac{d}{c}\right). \tag{67}$$

Then, applying the delta method, we arrive at:

$$\sqrt{\widetilde{N}_j}\left(\hat{\theta}_{XY}(\lambda_j)-\theta_{XY}(\lambda_j)\right)\sim AN(0,R_{abc,j}(0)), \tag{68}$$

where the large sample variance, $R_{abc,j}(0)$, is computed as previously utilizing an appropriate vector of partial derivatives equal:

$$\nabla g\left([a,b,c,d]^T\right)=\left[0,0,-\frac{d}{c^2+d^2},\frac{c}{c^2+d^2}\right]^T.$$

Then, the confidence intervals are given in the form similar to (66). Multiplying them by a constant will produce approximate CIs for the wavelet time delay.

Estimates of the large sample variances of the estimators of the wavelet spectra can be obtained via nonparametric kernel methods. Below we examine properties of two kernel estimators: one based on the popular Bartlett kernel and the other based on the truncated (rectangular) kernel that is recommended when the order of serial correlation is known[16]. Figures 4 and 5 present an example investigation conducted for the same linear regression with delay as in the case of the mean wavelet estimates in Figures 2 and 3. In terms of unbiasedness, both methods seem to produce quite satisfactory results for the variances of the wavelet phase spectrum and time delay estimators, even for relatively small sample sizes, when carefully chosen truncation parameters are used. However, their performance in the case of the wavelet coherence estimator is much worse, depending on the particular value of the delay parameter and the scale of the analysis.[17]

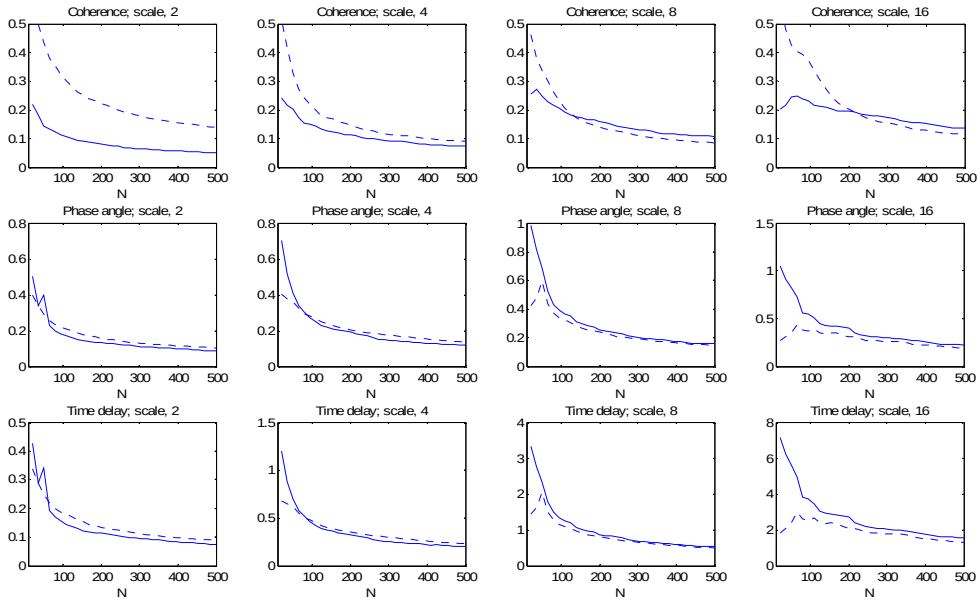

Figure 4. Sample standard deviations of wavelet estimators and their kernel estimates – the case of the Bartlett kernel; truncation parameter was set to twice the scale ($M = 2\lambda_j$); solid lines are the sample SDs and dashed lines present the kernel estimates; results based on 200 simulations of the linear regression model with delay $\tau = 1$, $\alpha = 1$ and X_t and η_t being two independent AR(1) processes with autoregressive parameters 0.8 and unit error variances; samples consist of 20, 35, 50 …, 500 wavelet coefficients unaffected by the boundary; the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k4l2.

[16] See Ogaki et al. (2009), Chapter 6.
[17] Results of a detailed examination are available upon request.

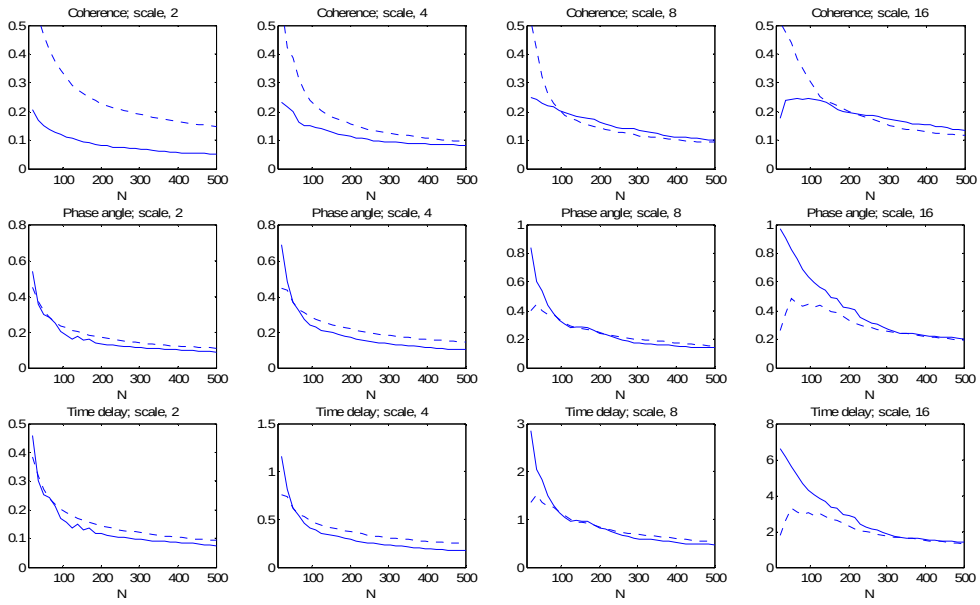

Figure 5. Sample standard deviations of wavelet estimators and their kernel estimates – the case of the truncated kernel; truncation parameter was set to the scale ($M = \lambda_j$); see detailed description below Figure 4.

3. Wavelet time delay estimation – a simulation study

In this section we summarize results of simulation experiments examining statistical properties of two wavelet estimators of the time delay parameter. We concentrate on the model (55) and analyze mainly a small sample performance of the estimators in order to recommend a method of examining short-term lead-lag relations for octave frequency bands. In particular, such a method might be of interest in business cycle studies, as it should be useful when analyzing changing patterns of business and growth cycle synchronization. The estimators compared are: the wavelet cross-correlator (WCC) that is based on maximizing values of the cross-covariance estimates, i.e.:

$$\hat{\tau}_j^{WCC} = \arg\max_k \mathrm{Cov}_{\tilde{W}_j^X \tilde{W}_j^Y}(k) = \arg\max_k \mathrm{Cov}(\tilde{W}_{j,t}^X, \tilde{W}_{j,t+k}^Y), \qquad (71)$$

and the estimator of the wavelet time delay (62), i.e.:

$$\hat{\tau}_j^{WPA} \overset{\mathrm{df}}{=} -\frac{\hat{\theta}_{XY}(\lambda_j)}{2\pi f_{j0}}, \qquad (72)$$

which we call further the wavelet phase angle (WPA) delay estimator. It is worth stressing that we do not maximize the absolute value of the cross-covariance in (71), as is often done in the time delay literature, due to the fact that a large negative covariance will be interpreted as an anti-phase relationship.

The Cramér-Rao lower bound (CRLB) on the variance of any unbiased time delay estimator was derived to be (see Carter, 1987):

$$\min \sigma^2 = \frac{1}{N \int_{-0,5}^{0,5} (2\pi f)^2 \frac{K(f)}{1-K(f)} df}, \qquad (73)$$

where $K(f)$ is the Fourier coherence of the processes under study (in our case – the wavelet coefficients $\tilde{W}_j^X$ and $\tilde{W}_j^Y$). Formula (73) predicts that the variance of an optimal estimator decreases with the value of coherence, the signal bandwidth and the center frequency. For the wavelet cross-correlator it is known that for jointly stationary processes and large enough data samples the CRLB is automatically achieved in the case of the signal and noise processes with spectra that are flat over the same range of frequencies and zero outside this range – see Scarbrough et al. (1981), Carter (1987). If the spectra of X_t and η_t are relatively featureless within octave frequency bands, MODWT wavelet coefficients of X_t and Y_t will be approximately bandpass white noises and the WCC becomes efficient asymptotically. However, the actual performance of the asymptotically efficient estimator can be much worse, especially for low signal-to-noise ratios, SNR (see simulation results in Scarbrough et al., 1981, Carter, 1987). As for estimators based on the Fourier phase angle, it has been proven that they are fully consistent with other asymptotically optimal methods after regression analysis is applied to the phase data – see Piersol (1981).

When discussing properties of the estimators (71) and (72), it is worth underlining that the WPA method enables to estimate delays that are not an integer multiple of the sampling period and guarantees a maximal time localization limited only by the length of the applied filter. On the other hand, the wavelet cross-correlator makes it possible to estimate delays that are longer than half of the center period, $1/f_{j0}$, is asymptotically unbiased, also for the first stage of analysis, and can be based on shorter filters, as the approximately analytic complex wavelet filters are usually longer than real filters with similar squared gain functions. The length of the wavelet filter plays a crucial role in empirical examination, since it directly influences the number of wavelet coefficients that are unaffected by the extrapolation method at the ends of the sample and therefore determines the maximal number of decomposition levels as well as the precision of estimation.

In our simulations the following data generating process is used:

$$\begin{aligned} Y_t &= \alpha X_{t-\tau} + \eta_t, \\ X_t &= \beta X_{t-1} + \varepsilon_t, \\ \eta_t &= \gamma \eta_{t-1} + \xi_t, \\ \begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix} &\sim N\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right). \end{aligned} \qquad (74)$$

Figures 6–8 present a comparison of small sample biases, root mean square errors (RMSE) and large sample standard errors of the two estimators for the case: $\alpha = 1$; $\tau = 1$; $\beta = \gamma = 0{,}8$; $SNR = \frac{\sigma_1}{\sigma_2} = \frac{1}{3}, \frac{1}{2}, 1, 2, 3$, and the following wavelet filters: la12 for WCC and la12 (first stage) + k4l2

(higher stages) for WPA.[18] The search range for the WCC was $\pm 10 + \tau$. In each case 1000 replications were run.[19] In the presentation we also include the outcomes for the first decomposition level, largely because to some extent they are comparable to the other stages. The findings resulting from the experiments are summarized below:[20]

- For wide ranges of signal-to-noise ratios and scales, the WPA estimator is better than WCC in small samples (see Figure 7). For the majority of outcomes, the relative efficiency of the two methods defined as $\frac{\text{RMSE}(WPA)}{\text{RMSE}(WCC)}$ increases with the SNR, the scale and the sample size. In large samples (see Figure 8) the WCC dominates the WPA estimator or their performance is similar. In small samples the relative efficiency of the methods depends largely on the search range for the WCC, although for similar ranges of delays for both methods and for lower signal-to-noise ratios, the WPA method performs better.
- In larger samples the RMSE for both estimators increase with the scale. For this reason, it is generally advisable to assume smoothing windows with length proportional to the scale.[21]
- For the lowest SNR case, both estimators show small sample bias, although with opposite signs: the WPA estimator towards 0, while the cross-correlator in the opposite direction (see Figure 6). This suggests that the usual biased estimator of the cross-covariance might be preferred in very small samples, when an estimate of the time delay parameter is needed. Other our experiments also demonstrate that the small sample bias of the WCC largely depends on whether the search range for the WCC is symmetric around the true value of the delay.
- Other experiments not reported here indicate that for jointly stationary processes including observations affected by circularity increases an effective sample size and improves the overall performance of both estimators, especially for highest scales, where the number of affected coefficients is large. This, however, does not take place for nonstationary processes.
- All the above observations are unchanged across different wavelets, although the outcomes obtained with the WPA method depend on the analytic properties of the HWP filters. Good analytic wavelets, however, produce almost identical results (for example, k4l2, k3l3). Different values of β and γ (including the nonstationary case) do not change the conclusions either.

[18] These wavelet filters where chosen due to their popularity and also to guarantee maximal similarity in implementation of the WCC and WPA methods: la12 and k4l2 are of the same length (L = 12) and have similar squared gain functions. However, basically the same results were obtained for, e.g., la8 + k3l3, la12 + k4l4, real part of k4l2 + k4l2. The Selesnick's HWP(K, L) filters outperformed the two Q-shift filters that were also considered, i.e. kin and tkp12.

[19] All computations, including the empirical part, were executed in Matlab. Numerical codes are available via e-mailing the author.

[20] We also comment shortly on other experiments we performed. More detailed results are available upon request.

[21] Comp. Cohen, Walden (2010b).

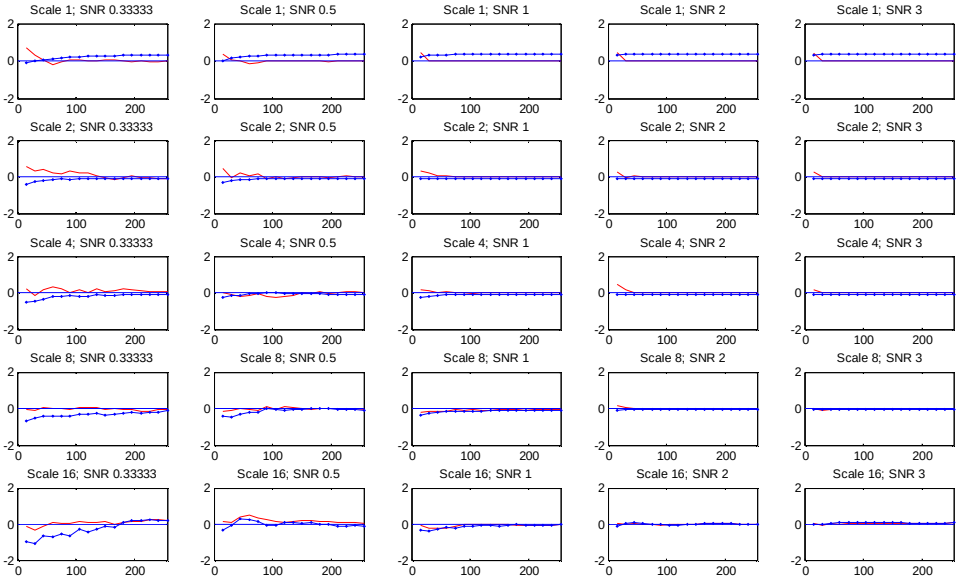

Figure 6. Small sample bias of time delay estimators; lines with and without markers correspond to the WPA and the WCC methods, respectively; samples consist of 15, 30, 45, …, 255 wavelet coefficients unaffected by the boundary.

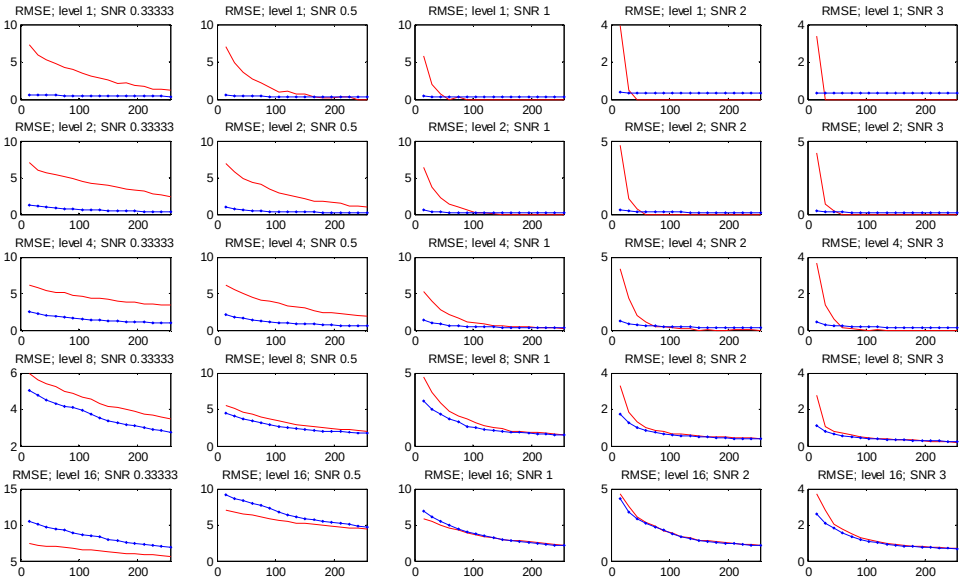

Figure 7. Root mean square errors of time delay estimators in small samples; lines with and without markers correspond to the WPA and the WCC methods, respectively; samples consist of 15, 30, 45, …, 255 wavelet coefficients unaffected by the boundary.

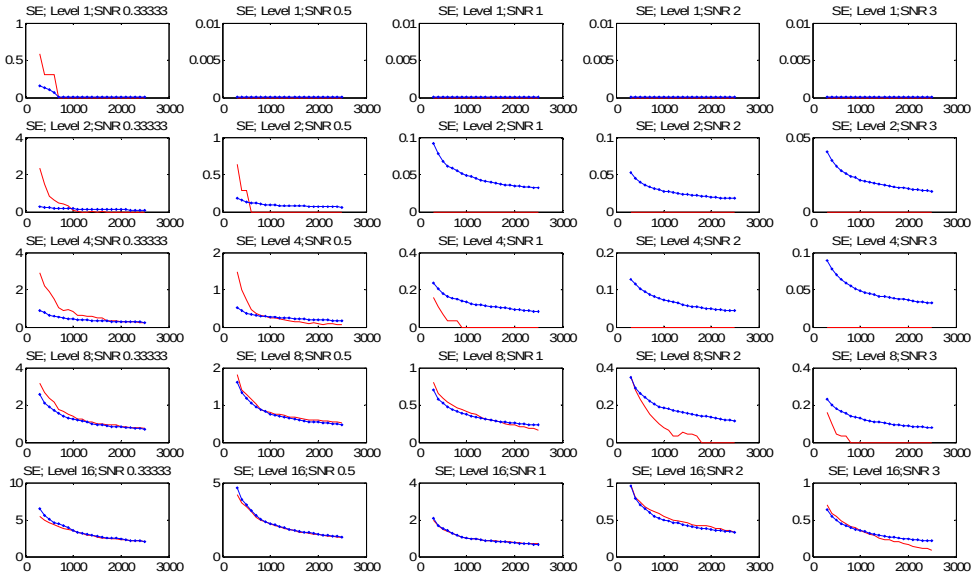

Figure 8. Standard errors of time delay estimators in large samples; lines with and without markers correspond to the WPA and the WCC methods, respectively; samples consist of 300, 400, , …, 2500 wavelet coefficients unaffected by the boundary.

The main finding of this section is that in business cycle studies, in which one often deals with low signal-to-noise ratios and which are typically based on relatively short time series, especially when countries undergoing transitions are examined, the wavelet phase angle methodology seems to be particularly attractive and might be used at least as a supplementary method. It is worth stressing that, besides its good localization properties, the WPA estimator is also simple and efficient computationally. For these reasons we believe it might be recommended for empirical analysis on business cycle synchronization.

4. Empirical examination

The data used in the empirical study are quarterly GDP volume estimates from the OECD Quarterly National Accounts (measure: VOBARSA) covering the period from the first quarter of 1960 till the second quarter of 2010 (202 observations) for the following 11 countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain. Besides, the OECD GDP volume for the euro area (16 countries) is used, which covers the shorter period from 1995 till the end of the sample (62 observations). The examination is divided into two parts. In the first part local wavelet variance analysis is performed, while the second deals with local and global wavelet analysis of synchronization.

4.1. Business cycle variability

Our examination of business cycle variability has been performed with the help of the d4 Daubechies wavelet filter of length 4 that guarantees very good localization properties. The estimates of the wavelet variance have been computed in windows consisting of 30 wavelet coefficients unaffected by the boundary. The results are presented in Figures 9 and 10.

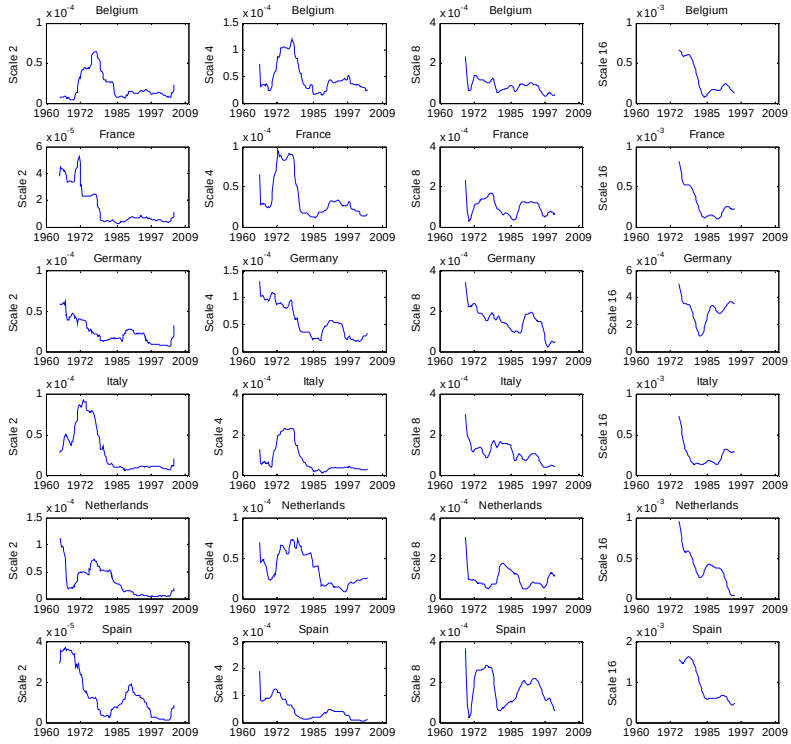

Figure 9. Running wavelet variance for scales 2, 4, 8 and 16 corresponding to oscillations with period lengths 4–8 (1–2 years), 8–16 (2–4 years), 16–32 (4–8 years), 32–64 (8–16 years); results obtained with d4 Daubechies wavelet filter of length 4 and windows of 30 wavelet coefficients unaffected by circularity after aligning them to the observations in the sample; step = 1.

Firstly, we notice quite similar patters of volatility changes across countries in our sample, except for Finland (scale 4 and 8) and Ireland (all scales). Also contributions of different scales to the total variance as well as estimates of the wavelet variance alone seem not to vary much across the economies. For some countries we observe a systematic decline in the variance for all decomposition levels, which started at the beginning of our sample, as is the case for Germany (except for the more volatile period about the reunification as well as for the highest scale) and Spain. It is seen that the oil price shocks of 1973 and 1979 have been captured almost entirely by the shortest components of business cycle fluctuations. Thanks to this the scale 8 wavelet variance provides a more clear view of the Great Moderation, revealing that the process might have started well before the mid-1980s (comp. Aguiar-Conraria, Soares, 2010, for similar evidence for the United States obtained with the continuous

wavelet transformation). Also the most recent perturbations (the financial crisis of 2007–2009) are becoming apparent at the lowest decomposition level, as seen for example in the case of the euro area GDP.

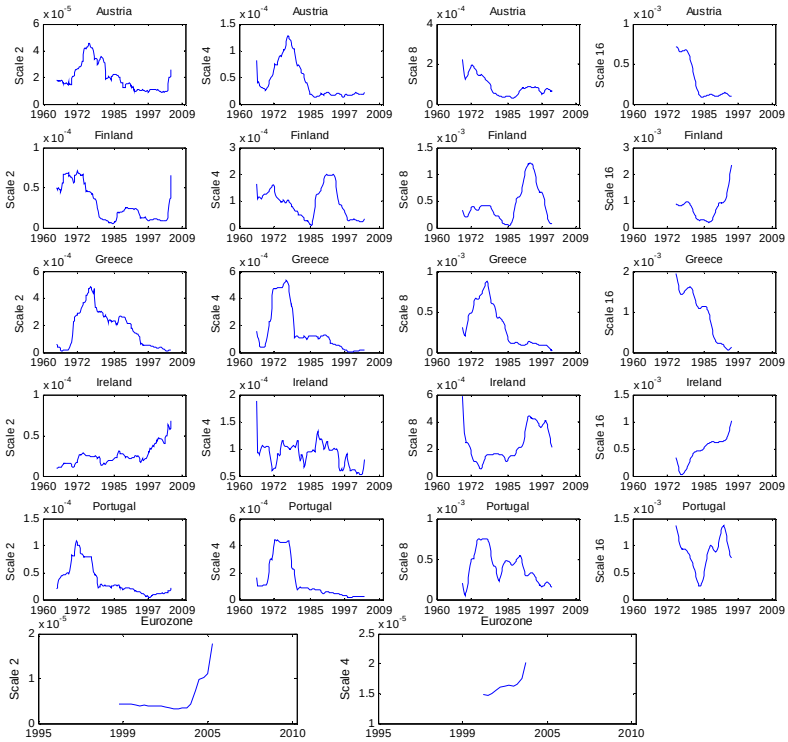

Figure 10. Running wavelet variance for scales 2, 4, 8 and 16 corresponding to oscillations with period lengths 4–8 (1–2 years), 8–16 (2–4 years), 16–32 (4–8 years), 32–64 (8–16 years), except for the euro zone data, where scales 2 and 4 are only considered; results obtained with d4 Daubechies wavelet filter of length 4 and windows of 30 wavelet coefficients unaffected by circularity after aligning them to the observations in the sample; step = 1.

4.2. Business cycle synchronization

Business cycle synchronization in the euro area was examined with the help of the local wavelet correlations, the global and local wavelet coherences and the global and local wavelet time delays. Figure 11 presents running wavelet correlations for scales 4, 8 and 16 computed in windows consisting of 40 non-boundary wavelet coefficients after the MODWT based on d4 wavelet has been applied to the observations. We decided to treat Germany as the reference country due to the high correlation between scale 4 wavelet coefficients for Germany and the euro zone as compared to e.g. France and the euro zone (see Figure 11). The most interesting finding resulting from the analysis of the wavelet correlations is that for the majority of countries in the sample we observe a systematic increase in the strength of the instantaneous relationships between business cycles of the examined

countries starting from the second half of the 1980s, especially for scales 4 and 8 corresponding to oscillations with period lengths below 4 and 8 years, respectively. The change in the dynamical correlation patterns agrees with the introduction of the Single European Act, which was signed in 1986 and came into effect in 1987. It may also be observed that for the highest scale considered (cycles of length 8 year and above) there often is an opposite tendency in the instantaneous dependencies.

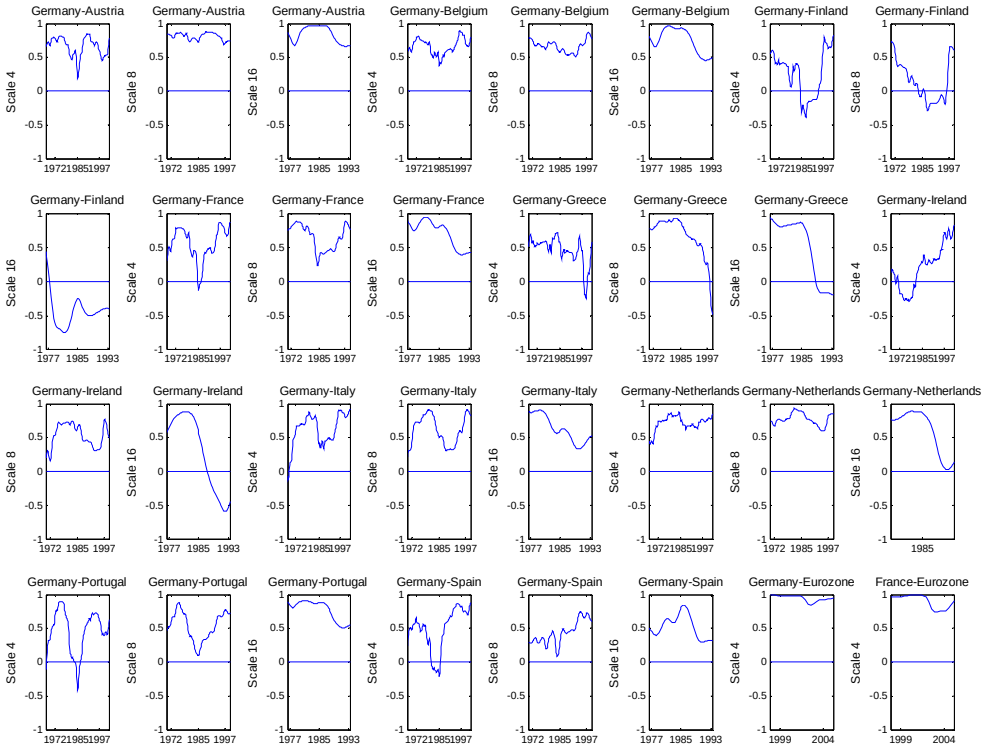

Figure 11. Running wavelet correlations for scales 4, 8 and 16 corresponding to oscillations with period lengths 8–16 (2–4 years), 16–32 (4–8 years), 32–64 (8–16 years), except for the euro zone data, where scale 4 is only considered; results obtained with d4 Daubechies wavelet filter of length 4 and windows of 40 wavelet coefficients unaffected by circularity after aligning them to the observations in the sample; step = 1.

Figures 12–13 present the results of the complex discrete wavelet analysis performed with the modified method described in section 2.4 based on the k4l2 wavelet filter for levels 2–4 and the la12 Daubechies filter together with its one-sample shifted variant at the first stage of the analysis. The global examination shows high dependencies at all leads and lags between Germany and the other countries in the sample at the third decomposition level (for shorter cycles) and for countries like Austria and the Netherlands also at the fourth level. Positive delays mean that the German cycle is behind the other countries' cycles, as is, for example, in the case of the long French cycle and the short Irish cycle. Instantaneous dependencies with the German cycles take place for countries like Belgium, the Netherlands, Greece and Italy. The local analysis in Figure 13 reveals that shorter cycles are becoming more synchronized starting from the middle of our sample, while in the case of the longer cycles there are certain patterns of lead-lag relations that seem to be quite stable over time (partially

because only a couple of coefficient windows were available), though the wavelet coherences are uprising.

The overall conclusion from both the real as well as complex wavelet analysis is that the synchronization between euro zone business cycles started to rise after the first important steps toward European integration were taken. This stays in line with the endogeneity hypothesis of the optimum currency area criteria as stated by Frankel and Rose (1998). Finally, Figure 14 presents comparison of local wavelet time delay estimates obtained with the WCC and WPA methods. It turns out that to a large extent both methods produce similar results.

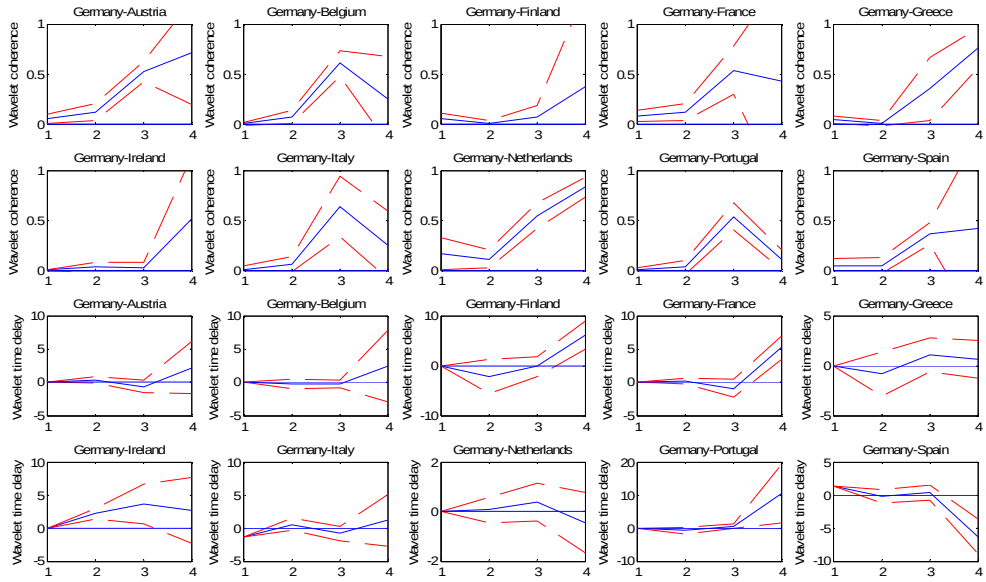

Figure 12. Wavelet coherence and wavelet time delay for scales 1, 2, 4 and 8 (decomposition levels 1–4) corresponding to oscillations with period lengths 2–4 (below one year), 4–8 (1–2 years), 8–16 (2–4 years) and 16–32 (4–8 years) together with large sample 90% confidence intervals; in the wavelet time delay estimation with the WPA estimator the reference country (Germany) is X and the other country – Y – see equation (55); the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k4l2; only wavelet coefficients unaffected by circularity are considered; the truncated kernel is used in variance estimation with truncation parameters for the four decomposition levels equal: 1, 2, 2, 2 (for the wavelet coherence) and 1, 2, 4, 8 (for the wavelet time delay).

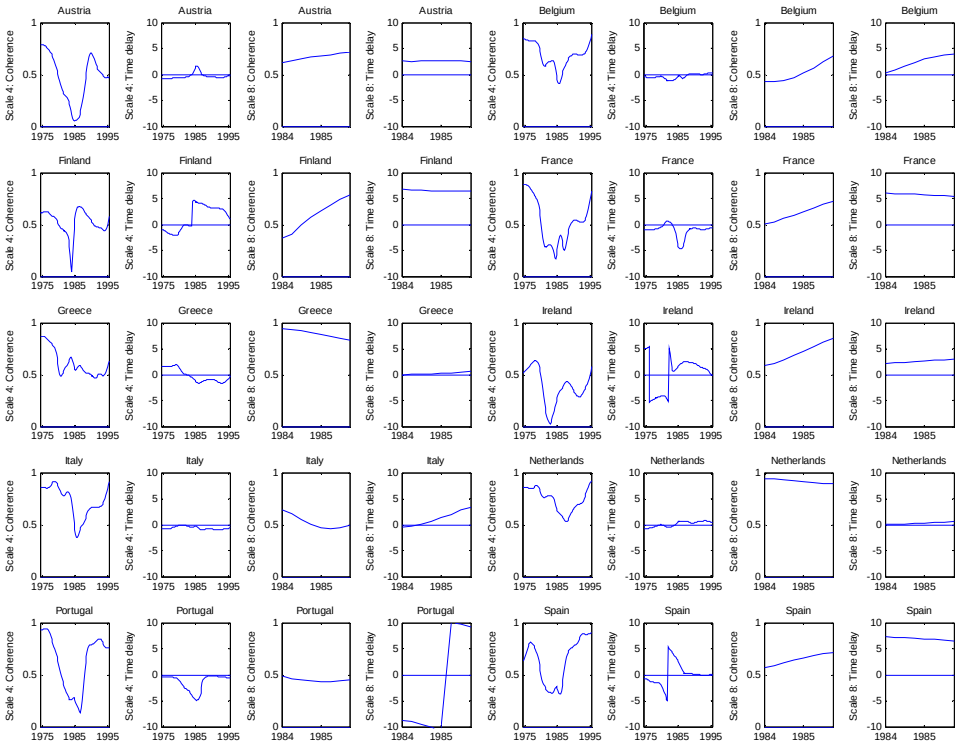

Figure 13. Running wavelet coherence and wavelet time delay for scales 4 and 8 corresponding to oscillations with period lengths 8–16 (2–4 years) and 16–32 (4–8 years); in the wavelet time delay estimation with the WPA estimator the reference country (Germany) is X and the other country – Y – see equation (55); the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k4l2; data windows consist of 30 wavelet coefficients unaffected by circularity for scale 4 and 40 – for scale 8, circularly shifted to align them to the real data; step = 1.

5. Conclusions

The discrete wavelet analysis provides a summary of evolutionary spectral and cross-spectral properties of processes under scrutiny with high computational efficiency, good localization properties and without an excessive redundancy of information that takes place in the case of the continuous wavelet methodology. These features together with a certain specific fresh look at an old problem seem to be the main reasons why the approach might be worth considering in business cycle examination.

The paper has discussed some of the questions arising in discrete wavelet analysis of popular bivariate spectral quantities like the amplitude, coherence and phase spectra and the frequency-dependent time delay. In particular, we show how the wavelet bivariate spectra can serve to approximate the corresponding Fourier quantities, discuss certain implementation issues and statistical inference problems. Our simulation study of properties of two wavelet estimators of the time delay parameter points at a practical relevance of the wavelet phase angle-based estimator suggested here,

which can be used at least as a supplementary method of examining short- and medium-term lead-lag relations for octave frequency bands.

The complex discrete wavelet methodology has been illustrated with an examination of business cycle synchronization in the euro zone. The study has also been supplemented with wavelet analysis of variance and covariance of European business cycles. The empirical examination gives some new arguments in favour of the endogeneity hypothesis of the optimum currency area criteria as well as an early start of the Great Moderation in Europe.

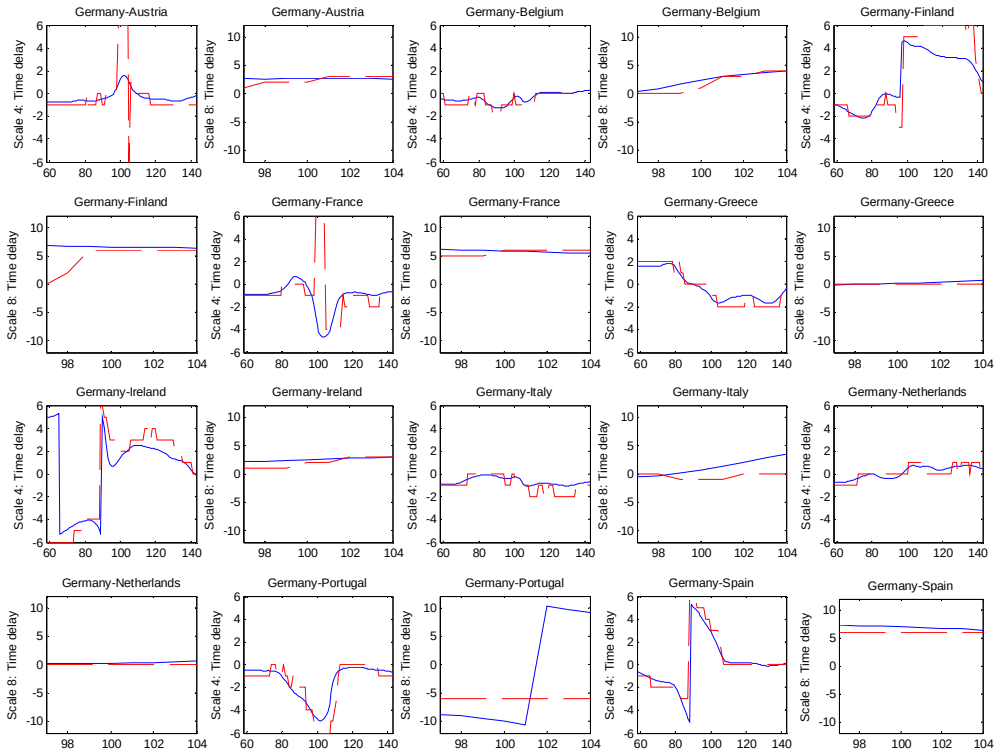

Figure 14. Running wavelet estimates of time delay for scales 4 and 8 corresponding to oscillations with period lengths 8–16 (2–4 years) and 16–32 (4–8 years); the solid blue line is the result obtained with the WPA estimator and the dashed red line – with the WCC; in the WPA method the first stage filters are la12 and its one-sample shifted version and the higher level filter is k4l2; the WCC is based on la12; data windows consist of 30 wavelet coefficients unaffected by circularity for scale 4 and 40 – for scale 8, circularly shifted to align them to the real data; step = 1; the numbers on the horizontal axis are the mid-points of the subsamples.

Nicolaus Copernicus University, Faculty of Economic Sciences and Management, Gagarina 13a, 87-100 Toruń, Poland, tel. 56 611 46 11, e-mail Joanna.Bruzda@umk.pl. *Suggestions and comments are welcome. The author acknowledges the financial support from the Polish Ministry of Science and Higher Education under the grant „Applications of wavelet decompositions to economic time series – exploratory analysis, business cycle synchronization and forecasting”.*

References

Aguiar-Conraria L., Soares M. J. (2009), Business Cycle Synchronization Across the Euro Area: a Wavelet Analysis, *NIPE Working Papers*, 8/2009, University of Minho, http://www3.eeg.uminho.pt/economia/nipe/docs/2009/NIPE_WP_8_2009.pdf

Aguiar-Conraria L., Soares M. J. (2010), The Continuous Wavelet Transform: A Primer, *NIPE Working Papers*, 23/2010, University of Minho, http://www3.eeg.uminho.pt/economia/nipe/docs/2010/NIPE_WP_23_2010.pdf

Blanchard O., Simon J. (2001), The Long and Large Decline in U.S. Output Volatility, *Brookings Papers on Economic Activity*, 2001-1, 135-164.

Carter G. C. (1987), Coherence and Time Delay Estimation, *Proceedings of the IEEE*, 75 (2), 236-255.

Cohen E. A. K., Walden A. T. (2010a), A Statistical Analysis of Morse Wavelet Coherence, *IEEE Transactions of Signal Processing*, 58 (3), 980-989.

Cohen E. A. K., Walden A. T. (2010b), A Statistical Study of Temporally Smoothed Wavelet Coherence, *IEEE Transactions of Signal Processing*, 58 (6), 2964-2973.

Crowley P. M., Lee J. (2005), Decomposing the Co-movement of the Business Cycle: A Time-Frequency Analysis of Growth Cycles In the Euro Area, *Bank of Finland Research Discussion Papers*,12, http://www.suomenpankki.fi/en/julkaisut/tutkimukset/keskustelualoitteet/Documents/0512netti.pdf

Crowley P. M., Maraun D., Mayes D. (2006), How Hard Is the Euro Area Core? An Evaluation of Growth Cycles Using Wavelet Analysis, *Bank of Finland Research Discussion Papers*, 18. http://www.suomenpankki.fi/en/julkaisut/tutkimukset/keskustelualoitteet/Documents/0618netti.pdf

Daubechies I. (1992), *Ten Lectures on Wavelets*, SIAM, Philadelphia.

de Haan J., Inklaar R., Jong-A-Pin R. (2008), Will Business Cycles in the Euro Area Converge? A Critical Survey of Empirical Research, *Journal of Economic Surveys*, 22 (2), 234-273.

Fernandez V. (2008), Multi-period Hedge Ratios for a Multi-asset Portfolio When Accounting for Returns Co-movement, *Journal of Futures Markets*, 28 (2), 182-207.

Frankel J., Rose A. (1998), The Endogeneity of the Optimum Currency Area Criteria, *The Economic Journal*, 108 (449), 1009–1025.

Gallegati M., Gallegati M. (2007), Wavelet Variance Analysis of Output in G-7 Countries, *Studies in Nonlinear Dynamics and Econometrics*, 11(3), art. 6.

Ge Z. (2008), Significance Tests for the Wavelet Cross Spectrum and Wavelet Linear Coherence, *Annales Geophysicae*, 26, 3819–3829.

Gençay R. F., Selçuk F., Whitcher B. (2002), An Introduction to Wavelets and Other Filtering *Methods in Finance and Economics*, Academic Press, San Diego.

Gonçalves E. S., Rodrigues M., Soares T. (2009), Correlation of Business Cycles in the Euro Zone, *Economics Letters*, 102 (1), 56-58.

Granger C. W. J., Hatanaka M. (1964), *Spectral Analysis of Economic Time Series*, Princeton University Press, Princeton.

Huang N. E., Shen S. S. P. (eds.) (2005), *Hilbert-Huang Transform and Its Applications*, World Scientific, Singapure.

Jagrič T., Ovin R. (2004), Method of Analyzing Business Cycles in a Transition Economy: The Case of Slovenia, *The Developing Economies*, 42 (1), 42-62.

Kingsbury N. G. (1998), The Dual-tree Complex Wavelet Transform: A New Technique for Shift Invariance and Directional Filters, *Proc. 8th IEEE DSP Workshop*, Utah, Aug. 9-12, 1998, Paper no. 86.

Kingsbury N. G. (2001), Complex Wavelets for Shift Invariant Analysis and Filtering of Signals, *Journal of Applied and Computational Harmonic Analysis*, 10 (3), 234-253.

Knapp C., Carter G. (1976), The Generalized Correlation Method for Estimation of Time Delay, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 24 (4), 320-327.

Mallat S. (1998), *A Wavelet Tour of Signal Processing*, Academic Press, San Diego.

Ogaki M, Jang K., Lim H.-S., Bae Y., Imutra Y. (2009), *Structural Macroeconometrics*, in preparation, http://www.econ.ohio-state.edu/ogaki/GCOE2009/ch06.pdf

Percival D. B. (1995), On Estimation of the Wavelet Variance, *Biometrika*, 82 (3), 619-631.

Percival D. B., Walden A. T. (2000), *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge.

Piersol A. G. (1981), Time Delay Estimation Using Phase Data, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29 (3), 471-477.

Priestley M. B. (1981), *Spectral Analysis and Time Series*, Academic Press, London.

Raihan S. Md., Wen Y., Zeng B. (2005), Wavelet: A New Tool for Business Cycle Analysis, *Federal Reserve Bank of St. Louis Working Paper Series*, 2005-050A,
http://research.stlouisfed.org/wp/more/2005-050/

Rua A., Nunes L. C. (2009), International Comovement of Stock Market Returns: A Wavelet Analysis, *Journal of Empirical Finance*, 16 (4), 632-639.

Scarbrough K., Ahmed N., Carter C. (1981), On the Simulation of a Class of Time Delay Estimation Algorithms, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29 (3), 534-539.

Selesnick I. W. (2001), Hilbert Transform Pairs of Wavelet Bases, *IEEE Signal Processing Letters*, 8 (6), 170-173.

Selesnick I. W. (2002), The Design of Approximate Hilbert Transform Pairs of Wavelet Bases, *IEEE Transactions on Signal Processing*, 50 (5), 1144-1152.

Selesnick I. W., Baraniuk R. G., Kingsbury N. G. (2005), The Dual-tree Complex Wavelet Transform. A Coherent Framework for Multiscale Signal and Image Processing, *IEEE Signal Processing Magazine*, 22 (6), 123-149.

Serroukh A., Walden A. T., Percival D. B. (2000), Statistical Properties and Uses of the Wavelet Variance Estimator for the Scale Analysis of Time Series, *Journal of the American Statistical Association*, 95 (449), 184-196.

Tay D. B. H., Kingsbury N. G., Palaniswami M. (2006), Orthonormal Hilbert-pair of Wavelets with (Almost) Maximum Vanishing Moments, *IEEE Signal Processing Letters*, 13 (9), 533-536.

Torrence C., Compo G. P. (1998), A Practical Guide to Wavelet Analysis, *Bulletin of the American Meteorological Society*, 79 (1), 61-78.

Whitcher B. J. (1998), Assessing Nonstationary Time Series Using Wavelets, Ph.D. thesis, University of Washington, www.martinsewell.com/wavelets/Whit98.pdf

Whitcher B. J., Craigmile P. F. (2004), Multivariate Spectral Analysis Using Hilbert Wavelet Pairs, *International Journal of Wavelets, Multiresolution and Information Processing*, 2 (4), 567-587.

Whitcher, B. J., Craigmile P. F., Brown P. (2005). Time-varying Spectral Analysis in Neurophysiological Time Series Using Hilbert Wavelet Pairs, *Signal Processing*, 85 (11), 2065-2081.

Whitcher B. J., Guttorp P., Percival D. B. (2000), Wavelet Analysis of Covariance with Application to Atmospheric Time Series, *Journal of Geophysical Research*, 105 (D11), 941-962.

Wong H., Ip W.-C., Xie Z., Lui X. (2003), Modelling and Forecasting by Wavelets, and the Application to Exchange Rates, *Journal of Applied Statistics*, 30 (5), 537-553.

Yogo M. (2008), Measuring Business Cycles: A Wavelet Analysis of Economic Time Series, *Economics Letters*, 100 (2), 208-212.