

Bruzda, Joanna

Article

Some aspects of the discrete wavelet analysis of bivariate spectra for business cycle synchronisation

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

Provided in Cooperation with:

Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Bruzda, Joanna (2011) : Some aspects of the discrete wavelet analysis of bivariate spectra for business cycle synchronisation, Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 5, Iss. 2011-16, pp. 1-46,
<https://doi.org/10.5018/economics-ejournal.ja.2011-16>

This Version is available at:

<https://hdl.handle.net/10419/50738>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/2.0/de/deed.en>

Some Aspects of the Discrete Wavelet Analysis of Bivariate Spectra for Business Cycle Synchronisation

Joanna Bruzda

Nicolaus Copernicus University, Torun

Abstract The paper considers some of the issues emerging from the discrete wavelet analysis of popular bivariate spectral quantities such as the coherence and phase spectra and the frequency-dependent time delay. The approach utilised here is based on the maximal overlap discrete Hilbert wavelet transform (MODHWT). Firstly, via a broad set of simulation experiments, we examine the 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 the empirical examination of short- and medium-term lead-lag relations for octave frequency bands. Further, we point out a deficiency in the implementation of the MODHWT and suggest using a modified implementation scheme, which was proposed earlier in the context of the dual-tree complex wavelet transform. In addition, we show how MODHWT-based wavelet quantities can serve to approximate the Fourier bivariate spectra and discuss 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 synchronisation between 11 euro zone countries. The study is supplemented by a wavelet analysis of the variance and covariance of the euro zone business cycles. The empirical examination underlines the good localisation 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 the 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 synchronisation; euro zone

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

Citation Joanna Bruzda (2011). Some Aspects of the Discrete Wavelet Analysis of Bivariate Spectra for Business Cycle Synchronisation. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 5, 2011-16. <http://dx.doi.org/10.5018/economics-ejournal.ja.2011-16>

1 Introduction

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

In the paper we consider some of the issues emerging from the discrete wavelet analysis of popular bivariate spectral quantities such as 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 complex wavelet transform of Kingsbury (1998, 2001) and Selesnick (2001), known as the maximal overlap discrete Hilbert wavelet transform (MODHWT) (see also Whitcher et al. 2005). Following Percival and Walden (2000), we concentrate exclusively on discrete wavelet analysis treating it as a natural way of handling discrete time series, especially in economics, where additionally we often operate on frequency bands instead of a single frequency such as in the case of a business cycle examination. The non-decimated (maximal overlap) version of the wavelet transform, known also as the continuous discrete wavelet transform (Antoniadis and Gijbels 2002), is defined

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

for all (discrete) time units and octave frequency bands (or dyadic scales). The MODHWT-based methodology encompasses the analysis performed with the usual maximal overlap discrete wavelet transform (MODWT) and goes beyond it producing quantities which directly correspond to the bivariate Fourier spectra. In our theoretical considerations we start with pointing out the need to modify the implementation scheme of the MODHWT in a way suggested in the context of the dual-tree complex wavelet transform of Kingsbury and Selesnick (see, e.g., Selesnick et al. 2005). Further, we show how the wavelet quantities can serve to approximate the Fourier cross-spectra and discuss issues connected with constructing confidence intervals (CIs) for them. Finally, via a broad set of simulation experiments we examine the small and large sample properties of two wavelet estimators of the scale-dependent time delay – a quantity measuring the causal distance between time series on a scale-by-scale basis. The estimators are: the wavelet cross-correlator (WCC) and the wavelet phase angle (WPA)-based estimator. Our results provide some practical guidelines for the empirical examination of short- and medium-term lead-lag relations for octave frequency bands, pointing out the better small sample performance of the WPA-based estimator, especially in the case of low signal-to-noise ratios (SNRs). In the empirical analysis of business cycle variability and synchronisation in the euro zone we utilize the two ‘continuous discrete’ wavelet transforms: the maximal overlap DWT and the MODHWT. From the point of view of an economist willing to study business cycles, the MODWT and MODHWT offer the following:

- A model-free (non-parametric) approach to examining the frequency characteristics of economic processes, i.e. short-, medium- and long-run features in the series. In particular, because of their non-parametric nature, wavelets enable us to examine non-linear processes without the loss of information.
- Good time-frequency resolution, and because of this, efficiency in terms of the computations needed to extract the features. This enables the precise examination of a time-varying frequency content of time series in an efficient way.

- The decomposition of the 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 the shocks causing and influencing business cycles.
- Low computational complexity.³
- The examination of trended, seasonal and integrated time series without prior transformations: In particular, we do not need to deseasonalise 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. In addition, there is no need for the prior elimination of deterministic and stochastic trends because wavelet filtering usually embeds enough differencing operations.
- The efficient estimation of short-term lead-lag relations for different frequency bands.
- Global and local (short-term) measures of association for business cycle components such as wavelet correlations and cross-correlations, wavelet coherence and wavelet phase angle.

Recent studies on business cycle synchronisation 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 synchronisation patterns

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

³ The conventional discrete wavelet transform (DWT) can be computed using an algorithm that is faster than is the well-known fast Fourier transform (FFT) – Mallat’s pyramid algorithm based on a mirror filters cascade and downsampling by two, which requires only $O(N)$ multiplications. By contrast, the computational complexity of the MODWT is $O(N\log_2 N)$, which is exactly the same as the FFT (Percival and Walden 2000: 159), while the MODHWT consumes twice as many operations.

alone, decomposed on a scale-by-scale basis. The study documents a rise in synchronisation between business cycles after the first steps towards European integration were taken in the second half of the 1980s. In addition, changes in business cycle variability are examined, thus providing new evidence on dating the Great Moderation and staying in agreement with the hypothesis of an early start of the process (Blanchard and Simon 2001).

The structure of the paper is as follows. In the next section, we briefly 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 MODHWT and discuss more deeply the wavelet bivariate spectral analysis, its connections with the Fourier analysis as well as certain implementation problems. Section 3 presents the results of the simulation analysis comparing two wavelet methods of examining lead-lag relations for octave frequency bands, while Section 4 summarises our empirical findings. Finally, the last section offers brief conclusions.

2 Wavelet Analysis

Wavelet transformation consists in decomposing a signal into the 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 us to recognise local features in the data, especially in the case of signals defined over the entire real axis, although it results in an excessive redundancy of information. The 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 MODWT removes certain deficiencies of the conventional discrete transformation by considering all time units, while – similarly to the DWT – octave frequency bands are analysed.

2.1 Conventional and Maximal Overlap DWTs

The conventional DWT of a real-valued function $f(x)$ is defined as follows:

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

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

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

For certain functions $\psi(x)$ with good localisation properties, $\{\psi_{j,t}(x)\}$ is an orthonormal basis in $L^2(\mathfrak{R})$. 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 analogous to (2), produces another set of coefficients in the form:

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

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

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

The functions $S_j(x)$ and $D_j(x)$ are known as approximations (smooths) and details. The highest level approximation $S_J(x)$ represents the 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 DWT is defined via quadrature mirror filters: the low-pass (scaling) filter $\{g_l\}_{l=0,\dots,L-1}$ and the high-pass (wavelet) filter $\{h_l\}_{l=0,\dots,L-1}$.⁴ When processing discrete signals we consider a vector of length $N = 2^J$ in the form $\mathbf{x} = (x_0, x_1, \dots, x_{N-1})'$. Then the highest possible decomposition level is J and the number of wavelet and scaling coefficients of the conventional DWT for level $j = 1, 2, \dots, J$ is $N/2, N/4, \dots, 1$, respectively. By contrast, the MODWT 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 use downsampling by two. The coefficients are appropriately scaled in order to retain variance preservation. They are given as follows:

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

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

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

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

where $\{h_{j,l}\}$ and $\{g_{j,l}\}$ are the j -th level wavelet and scaling filters of length $L_j = (2^j - 1)(L - 1) + 1$ obtained by convolving together the following j filters (Percival and Walden 2000: 95–96):

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

⁴ We concentrate on compactly supported orthonormal wavelets (see Daubechies 1992).

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

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

filters 1, ..., $j-1$ as in (A)

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

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

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

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

Then, we have (Percival and Walden 2000: 154):

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

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

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

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 $\tilde{H}(\cdot)$, $\tilde{G}(\cdot)$, $\tilde{H}_j(\cdot)$, $\tilde{G}_j(\cdot)$.

Among the most popular real wavelet and scaling filters are the compactly supported orthonormal Daubechies filters: the extremal phase (dL) and least asymmetric (laL) filters. The two families are characterised by the smallest filter length L for a given number of vanishing moments (VMs).⁵ In addition, 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:

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

Assuming that (12) does not depend on time,⁶ we arrive at the variance decomposition across scales in the form:

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

⁵ Roughly speaking, VMs are responsible for eliminating artefacts because of the wavelet function as well as for the degree of approximation to an ideal bandpass filter, and they make it possible to interpret the filters as generalised differences of adjacent observations with the number of embedded difference operations equal to the number of VMs (see Daubechies 1992: 153; Mallat 1998: 166; Percival and Walden 2000: 483). The number of VMs for the Daubechies filters equals half the filter length.

⁶ The assumption is also fulfilled for non-stationary processes if they are integrated of order d and the width of the wavelet filter, L , is sufficient to eliminate non-stationarity. In the case of Daubechies wavelet filters, the condition is: $L \geq 2d$ (see, e.g., Percival and Walden 2000: 304). In order to have $E\{\tilde{W}_{j,t}\} = 0$, we further assume that $L > 2d$.

The wavelet variance at level j , corresponding to scale $\lambda_j = 2^{j-1}$, $\sigma^2(\lambda_j)$, explains the variation of oscillations of a 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:

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

As in the case of the variance decomposition (13), if the wavelet covariance does not depend on time, it produces a decomposition of the covariance between X_t and Y_t across scales λ_j :

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

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

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

The quantity above 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 cross-correlation are given as:

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

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

As mentioned in section 2.1, MODWT-based estimators have generally better statistical properties compared with their DWT-based counterparts. Firstly, MODWT coefficients produce better estimates of the wavelet variance in terms of efficiency. Secondly, they give an estimator of the wavelet covariance whose

variance does not depend on the true time lag between time series. Furthermore, decimation by two affects the lag-resolution of the DWT-based estimators of the wavelet cross-covariance and cross-correlation, so they should not be used in practice (see Percival and Walden 2000: 308–310; Gençay et al. 2002: 252–253). For these reasons, we concentrate on estimation using the MODWT coefficients.

An unbiased estimator of the wavelet variance is defined as:

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

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 extrapolation method at the end of the sample.

Estimates of wavelet covariance and wavelet correlation are computed via the following formulas:

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

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

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

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

Methods of constructing CIs for the quantities described in this section are discussed by Percival (1995) and Whitcher (1998) (see also Percival and 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 provides estimates of the local versions of the wavelet quantities. A good time resolution is exactly what the non-decimated DWT 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 MODHWT

The MODHWT makes use of a recently introduced class of filters based on Hilbert wavelet pairs (HWPs) and utilises a non-decimated (maximal overlap) version of the dual-tree complex wavelet transform of Kingsbury (1998, 2001) and Selesnick (2001).⁷ The approach was advocated by Whitcher and Craigmile (2004) (see also Whitcher et al. 2005). The filters in a HWP are approximate Hilbert transforms of each other and, as in the case of the usual DWT, form a basis for the collection of orthogonal bandpass filters. This time, however, the approximate analyticity of the filters enables us to compute the quantities that directly correspond to the appropriate bivariate Fourier spectra.⁸

Let $\{h_l^0\}$ and $\{g_l^0\}$ be conjugate quadrature mirror filters. The father and mother wavelets are obtained via:

$$\phi^0(t) = \sqrt{2} \sum_l g_l^0 \phi^0(2t-l); \quad \psi^0(t) = \sqrt{2} \sum_l h_l^0 \phi^0(2t-l). \quad (23)$$

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:

$$\Psi^1(f) = \begin{cases} -i\Psi^0(f), & f > 0 \\ i\Psi^0(f), & f < 0 \end{cases} \quad (24)$$

⁷ For an introduction, see Selesnick et al. (2005).

⁸ Their continuous (in both time and scale) counterparts have long been known in applications of wavelet analysis – see, e.g., Torrence and Compo (1998).

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):⁹

If transfer functions of two scaling filters fulfil the condition:

$$G^1(f) = G^0(f)e^{-i\theta(f)}, \quad |f| < 0,5, \quad (25)$$

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

The condition (25) says that the digital filter $\{g_l^1\}$ should be a half-sample delayed version of $\{g_l^0\}$, i.e. $g_l^1 = g_{l-\frac{1}{2}}^0$. As a half-sample delay cannot be implemented with finite impulse response filters, only approximate solutions are available.

In following sections, 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 Selesnick's so-called 'common factor approach', firstly an all-pass filter with an approximately constant fractional group delay is constructed and then orthonormal filters are built via a linear system of equations and spectral factorisation. Under a specified degree of approximation (L) to the half-sample delay, the design procedure produces short filters with a given number of VMs (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) and HWP(4, 4) and denote them 'kKL'. An alternative approach introduced by Kingsbury (2001) produces the so-called Q-shift (quarter-shift) filters, which are approximately linear phase and whose 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 one VM and the 12-tap Q-shift filter of Tay et al. (2006) with five VMs and denote them 'kin' and 'tkp12', respectively.

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

⁹ The converse of the theorem is also true, so Selesnick's condition is both sufficient and necessary – see Selesnick et al. (2005) and references therein.

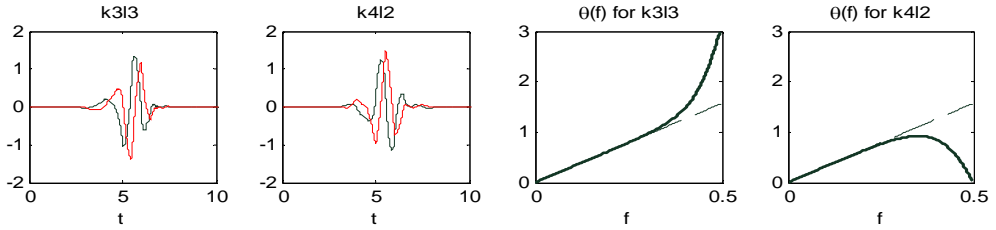


Figure 1. Example HWPs and phase differences for scaling filters

It is clear that the functions $\theta(f)$ fulfil the condition stated in 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 the 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 HWPs can be seen as short-term versions of the cosine and sine waves forming the classic Fourier transform.

The 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:

$$\begin{aligned}\tilde{h}_{j,l} &= \tilde{h}_{j,l}^0 + i\tilde{h}_{j,l}^1; \\ \tilde{g}_{j,l} &= \tilde{g}_{j,l}^0 + i\tilde{g}_{j,l}^1,\end{aligned}\tag{26}$$

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

$$\tilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_{l,j} X_{t-j} = \tilde{W}_{j,t}^0 + i\tilde{W}_{j,t}^1;\tag{27}$$

$$\tilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_{l,j} X_{t-j} = \tilde{V}_{j,t}^0 + i\tilde{V}_{j,t}^1.\tag{28}$$

As discussed in Selesnick et al. (2005), the simplest approach towards invertible analytic wavelet transform is via DWT post-processing. However, this results in a slightly higher computational complexity as in such a case we operate on two parallel complex wavelet transforms. Furthermore, performing the Hilbert transform as the first one is not recommended since then we lose the possibility of optimising it for all scales simultaneously. The dual-tree complex wavelet transform of Kingsbury and Selesnick is based on two real orthogonal wavelet filters with the Hilbert transform built into them. Thanks to this the Hilbert transform automatically adapts to the wavelet scales. This feature makes the approach particularly attractive compared with other time-frequency methods producing instantaneous amplitudes, phases and frequencies such as the classic demodulation method (see, e.g., Granger and Hatanaka 1964; Priestley 1981) or the modern Hilbert-Huang transform (see Huang and Shen 2005).

2.4 Wavelet Analysis of Coherence and Phase Angle

This section includes our theoretical considerations and starts with a detailed description of the relationship between the MODHWT-based wavelet spectra and the appropriate Fourier quantities. The wavelet analogues of Fourier spectral characteristics of bivariate time series were introduced by Whitcher and Craigmile (2004). Let $\tilde{W}_{j,t}^X$ and $\tilde{W}_{j,t}^Y$ be the complex-valued wavelet coefficients obtained via filtering X_t and Y_t . Assuming that the wavelet filters applied have enough VMs to eliminate any deterministic trend components in the series, the time-varying wavelet spectrum of (X_t, Y_t) for scale λ_j is defined as:

$$\begin{aligned} S_{XY}(\lambda_j, t) &= E\left(\tilde{W}_{j,t}^X \overline{\tilde{W}_{j,t}^Y}\right) = E\left[\left(\tilde{W}_{j,t}^{X0} + i\tilde{W}_{j,t}^{X1}\right)\left(\tilde{W}_{j,t}^{Y0} - i\tilde{W}_{j,t}^{Y1}\right)\right] \\ &= E\left[\left(\tilde{W}_{j,t}^{X0}\tilde{W}_{j,t}^{Y0} + \tilde{W}_{j,t}^{X1}\tilde{W}_{j,t}^{Y1}\right) - i\left(\tilde{W}_{j,t}^{X0}\tilde{W}_{j,t}^{Y1} - \tilde{W}_{j,t}^{X1}\tilde{W}_{j,t}^{Y0}\right)\right] \\ &= C_{XY}(\lambda_j, t) - i Q_{XY}(\lambda_j, t), \end{aligned} \quad (29)$$

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 an absolute summable cross-covariance sequence. We will denote its cross-

spectral density function as $S_{XY}(f) = C_{XY}(f) - i Q_{XY}(f)$.¹⁰ The wavelet cospectrum for scale λ_j is then:

$$C_{XY}(\lambda_j) = E(\tilde{W}_{j,t}^{X0} \tilde{W}_{j,t}^{Y0} + \tilde{W}_{j,t}^{X1} \tilde{W}_{j,t}^{Y1}) = \int_{-1/2}^{1/2} |\tilde{H}_j^0(f)|^2 S_{XY}(f) df + \int_{-1/2}^{1/2} |\tilde{H}_j^1(f)|^2 S_{XY}(f) df, \quad (30)$$

where $\tilde{H}_j^0(f)$ and $\tilde{H}_j^1(f)$ are the transfer functions of 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 (30) are identical, we obtain:

$$C_{XY}(\lambda_j) = 2 \int_{-1/2}^{1/2} |\tilde{H}_j^0(f)|^2 S_{XY}(f) df = 2 \int_{-1/2}^{1/2} |\tilde{H}_j^0(f)|^2 C_{XY}(f) df = 2 \gamma_{XY}(\lambda_j), \quad (31)$$

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:

$$C_{XY}(\lambda_j) \approx 4 \int_{1/2^{j+1}}^{1/2^j} C_{XY}(f) df. \quad (32)$$

As the quantity:

¹⁰ For wavelet filters with enough VMs, the discussion concerning wavelet co- and quadrature spectra can be directly generalised to the case of non-stationary 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 $S_{XY}(f) = \frac{S_{WZ}(f)}{(1 - e^{-i2\pi f})^{d_X} (1 - e^{-i2\pi f})^{d_Y}}$, where

$W_t = \Delta^{d_X} X_t, Z_t = \Delta^{d_Y} Y_t$. Note, however, that the generalisation does not apply to the part of our analysis that utilises the complex scaling coefficients $\tilde{V}_{j,t}$.

$$2^{j+1} \int_{1/2^{j+1}}^{1/2^j} C_{XY}(f) df \quad (33)$$

is the average value of $C_{XY}(f)$ in the interval $[\frac{1}{2^{j+1}}, \frac{1}{2^j}]$, we can interpret $C_j = \lambda_j C_{XY}(\lambda_j)$, $j=1, \dots, J$, 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.¹¹ In any case however, the wavelet quantities discussed here provide piecewise constant approximations to their Fourier counterparts and summarise 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 HWP. The condition (24) is equivalent to:

$$\Psi^1(f) = \Psi^0(f) e^{i\xi(f)} = \begin{cases} \Psi^0(f) e^{-i\frac{\pi}{2}}, & f > 0 \\ \Psi^0(f) e^{i\frac{\pi}{2}}, & f < 0 \end{cases} \quad (24^*)$$

Making use of this and utilising 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 and Walden 2000: 476), we obtain:

¹¹ Note that such an assumption is valid only for rather special kinds of relationships such as the ‘fixed angle lag’ relationship (Granger and Hatanaka 1964: 98) with a constant amplitude spectrum over the frequency bands of interest or a linear regression without delay.

$$\begin{aligned}
 Q_{XY}(\lambda_j) &= E(\tilde{W}_{j,t}^{X0} \tilde{W}_{j,t}^{Y1} - \tilde{W}_{j,t}^{X1} \tilde{W}_{j,t}^{Y0}) = \int_{-1/2}^{1/2} \tilde{H}_j^0(f) \overline{\tilde{H}_j^1(f)} S_{XY}(f) df \\
 &\quad - \int_{-1/2}^{1/2} \tilde{H}_j^1(f) \overline{\tilde{H}_j^0(f)} S_{XY}(f) df \approx \int_{-1/2}^{1/2} \Psi^0(2^j f) \overline{\Psi^1(2^j f)} S_{XY}(f) df \\
 &\quad - \int_{-1/2}^{1/2} \Psi^1(2^j f) \overline{\Psi^0(2^j f)} S_{XY}(f) df \approx \int_{-1/2}^{1/2} |\Psi^0(2^j f)|^2 e^{-i\xi(f)} S_{XY}(f) df \\
 &\quad - \int_{-1/2}^{1/2} |\Psi^0(2^j f)|^2 e^{i\xi(f)} S_{XY}(f) df = \int_{-1/2}^{1/2} |\Psi^0(2^j f)|^2 S_{XY}(f) [-2i \sin \xi(f)] df \\
 &= -2 \int_{-1/2}^{1/2} |\Psi^0(2^j f)|^2 Q_{XY}(f) \sin \xi(f) df = 2 \int_0^{1/2} |\Psi^0(2^j f)|^2 Q_{XY}(f) df \\
 &\quad - 2 \int_{-1/2}^0 |\Psi^0(2^j f)|^2 Q_{XY}(f) df \approx 4 \int_0^{1/2} |\tilde{H}_j^0(f)|^2 Q_{XY}(f) df.
 \end{aligned} \tag{34}$$

Unfortunately, it turns out that the first approximation assumed in (34) 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 behaviour of the associated wavelet filters at level j as $j \rightarrow \infty$. In order to obtain more practical results, one is advised to apply different orthogonal quadrature mirror filters in the first stage of Mallat's pyramid algorithm, namely filters approximately satisfying the condition: $\tilde{g}_l^1 = \tilde{g}_{l-1}^0$ (see Selesnick et al. 2005). This condition is different (and easier to implement) than is 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. We will now discriminate between the level 1 and level j ($j = 2, 3, \dots$) filters. We maintain the previous notation for the transfer functions of the remaining filters, while the transfer functions of the first level 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 first level scaling filters, we have:

$${}^1\tilde{G}^1(f) = e^{-i2\pi f} {}^1\tilde{G}^0(f). \tag{35}$$

Using this and (11a), for the first level wavelet filters, we obtain:¹²

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

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

$$\tilde{H}_2^1(f) = \tilde{H}^1(2f) {}^1\tilde{G}^1(f) = -e^{-i2\pi f(2L-1)} \tilde{G}^1(1/2 - 2f) {}^1\tilde{G}^0(f). \quad (37)$$

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

$$\tilde{H}_2^1(f) = -e^{-i[4\pi f(L-1) + \frac{\pi}{2}]} \tilde{G}^0\left(\frac{1}{2} - 2f\right) {}^1\tilde{G}^0(f) = e^{-i\frac{\pi}{2}} \tilde{H}^0(2f) {}^1\tilde{G}^0(f) = e^{-i\frac{\pi}{2}} \tilde{H}_2^0(f), \quad (38)$$

while the appropriate relation for $f < 0$ is obtained via complex conjugation. Therefore, at the second decomposition level we do not need to substitute for $\tilde{H}_j^0(f)$ and $\tilde{H}_j^1(f)$ in (34) 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 by writing:

$$\tilde{H}_j^1(f) = \tilde{H}^1(2^{j-1}f) {}^1\tilde{G}^1(f) \tilde{G}^1(2f) \tilde{G}^1(4f) \dots \tilde{G}^1(2^{j-2}f). \quad (39)$$

For $f > 0$ the first factor in (39) is given as:

$$\begin{aligned} \tilde{H}^1(2^{j-1}f) &= -e^{-i2\pi 2^{j-1}f(L-1)} \tilde{G}^1(1/2 - 2^{j-1}f) = -e^{-i\pi(2^j f L - 3 \cdot 2^{j-1} + \frac{1}{2})} \tilde{G}^0(1/2 - 2^{j-1}f) \\ &= -e^{-i\pi(-2^{j-1}f + L - \frac{1}{2})} = -e^{-i\pi(-2^{j-1}f + L - \frac{1}{2})} \tilde{H}^0(2^{j-1}f) \\ &= -e^{-i\pi(-2^{j-1}f + L - \frac{1}{2})} \tilde{H}^0(2^{j-1}f) \end{aligned} \quad (40)$$

and the whole expression is then:

¹² Here we assume that the level one real and imaginary filters have 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.

$$\begin{aligned}\tilde{H}_j^1(f) &= \tilde{H}^1(2^{j-1}f)e^{-i2\pi f} \tilde{G}^0(f)e^{-i2\pi f} \tilde{G}^0(2f)e^{-i4\pi f} \tilde{G}^0(4f)\dots \\ &\dots e^{-i2^{j-2}\pi f} \tilde{G}^0(2^{j-2}f) = -e^{-i\pi(-2^{j-1}f+L-\frac{1}{2})} e^{-i2\pi f} e^{-i\pi f(2^{j-1}-2)} \tilde{H}_j^0(f) \quad (41) \\ &= e^{-i\frac{\pi}{2}} \tilde{H}_j^0(f).\end{aligned}$$

The first stage complex scaling filter is easily obtained via the 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. Because the 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, this 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 us to compute the average value of its Fourier analogue in the interval $[\frac{1}{2^{j+1}}, \frac{1}{2^j}]$, Q_j , via the following relationship:

$$Q_{XY}(\lambda_j) \approx 4 \int_{1/2^{j+1}}^{1/2^j} Q_{XY}(f) df = \frac{Q_j}{2^{j-1}} = \frac{Q_j}{\lambda_j}. \quad (42)$$

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

$$S_{XY}(f) \approx \lambda_j S_{XY}(\lambda_j) \quad \text{for } f \in [\frac{1}{2^{j+1}}, \frac{1}{2^j}]. \quad (43)$$

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

$$\begin{aligned}E(\tilde{V}_{j,t}^{X0} \tilde{V}_{j,t}^{Y0} + \tilde{V}_{j,t}^{X1} \tilde{V}_{j,t}^{Y1}) &= \int_{-1/2}^{1/2} |\tilde{G}_j^0(f)|^2 S_{XY}(f) df + \int_{-1/2}^{1/2} |\tilde{G}_j^1(f)|^2 S_{XY}(f) df \approx \\ &\approx 4 \int_0^{1/2^{J+1}} C_{XY}(f) df = \frac{C_{J+1}}{2^{J-1}} = \frac{C_{J+1}}{\lambda_J}, \quad (44)\end{aligned}$$

where C_{J+1} is the average value of the Fourier cospectrum in the interval $[-1/2^{J+1}, 1/2^{J+1}]$. For the imaginary part of the spectrum, firstly we notice that from (11c), (25) and (35) we have: $\tilde{G}_J^1(f) = \tilde{G}_J^0(f)e^{-i\pi f 2^J}$. Then, we obtain:

$$E(\tilde{V}_{J,t}^{X0}\tilde{V}_{J,t}^{Y1} - \tilde{V}_{J,t}^{X1}\tilde{V}_{J,t}^{Y0}) = \int_{-1/2}^{1/2} \tilde{G}_J^0(f)\overline{\tilde{G}_J^1(f)}S_{XY}(f)df - \int_{-1/2}^{1/2} \tilde{G}_J^1(f)\overline{\tilde{G}_J^0(f)}S_{XY}(f)df$$

$$\approx 4 \int_0^{1/2^{J+1}} Q_{XY}(f) \sin \pi f 2^J df = \frac{Q_{J+1}}{\pi 2^{J-2}},$$

(45)

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 to Q_{J+1} .

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

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

(46)

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

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

(47)

and the time-varying wavelet magnitude squared coherence:¹³

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

(48)

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

¹³ We refer to this as the wavelet coherence or wavelet coherence spectrum. Its square root is called the wavelet coherency.

Let us consider a simple example of a stationary bivariate process – a linear regression with delay in the form:

$$Y_t = \alpha X_{t-\tau} + \eta_t, \quad (49)$$

where X_t and η_t are stationary processes, uncorrelated with each other at all leads and lags. Then, the Fourier quantities take the following values:

$$\begin{aligned} C_{XY}(f) &= \alpha \cos(2\pi f\tau) S_X(f), \quad Q_{XY}(f) = \alpha \sin(2\pi f\tau) S_X(f), \\ A_{XY}(f) &= \alpha S_X(f), \quad K_{XY}(f) = \alpha^2 \frac{S_X(f)}{S_Y(f)}, \quad \theta_{XY}(f) = -2\pi f\tau. \end{aligned}$$

Furthermore, we can use also the time delay: $\tau_{XY}(f) = -\frac{\theta_{XY}(f)}{2\pi f} = \tau$ as well as the gain: $G_{XY}(f) = \frac{A_{XY}(f)}{S_X(f)} = \alpha$. Assuming that the wavelet transform produces a bandpass white noise, the wavelet co-, quad- and amplitude spectra are the following:

$$\begin{aligned} C_{XY}(\lambda_j) &= 2 \int_{-1/2}^{1/2} \alpha \left| \tilde{H}_j^0(f) \right|^2 \cos(2\pi f\tau) S_X(f) df \approx \alpha 2^{j+2} \sigma_X^2(\lambda_j) \int_{1/2^{j+1}}^{1/2^j} \cos(2\pi f\tau) df \\ &= \frac{\alpha 2^{j+2} \sigma_X^2(\lambda_j)}{\pi\tau} \cos \frac{3\pi\tau}{2^{j+1}} \sin \frac{\pi\tau}{2^{j+1}}, \end{aligned} \quad (50)$$

$$\begin{aligned} Q_{XY}(\lambda_j) &= 2 \int_{-1/2}^{1/2} \alpha \left| H_j^0(f) \right|^2 \sin(2\pi f\tau) S_X(f) df \approx \alpha 2^{j+2} \sigma_X^2(\lambda_j) \int_{1/2^{j+1}}^{1/2^j} \sin(2\pi f\tau) df = \\ &= \frac{\alpha 2^{j+2} \sigma_X^2(\lambda_j)}{\pi\tau} \sin \frac{3\pi\tau}{2^{j+1}} \sin \frac{\pi\tau}{2^{j+1}}, \end{aligned} \quad (51)$$

$$A_{XY}(\lambda_j) \approx \frac{\alpha 2^{j+2} \sigma_X^2(\lambda_j)}{\pi\tau} \sin \frac{\pi\tau}{2^{j+1}} = 2\alpha \sigma_X^2(\lambda_j) \frac{\sin \frac{\pi\tau}{2^{j+1}}}{\frac{\pi\tau}{2^{j+1}}}, \quad (52)$$

where the last approximation takes place for high enough decomposition levels j .¹⁴ The wavelet coherence is given as:

$$K_{XY}(\lambda_j) \approx \frac{\alpha^2 \sigma_X^2(\lambda_j) \sin^2 \frac{\pi\tau}{2^{j+1}}}{\sigma_Y^2(\lambda_j) \left(\frac{\pi\tau}{2^{j+1}} \right)^2}, \quad (53)$$

and for higher decomposition levels this can be approximated via:

$$K_{XY}(\lambda_j) \approx \frac{\alpha^2 \sigma_X^2(\lambda_j)}{\sigma_Y^2(\lambda_j)}. \quad (54)$$

Assuming that the individual Fourier spectra are approximately piecewise constant over octave frequency bands, for high enough scales the wavelet coherence should provide an acceptable approximation to the appropriate Fourier quantity. The wavelet phase spectrum is:

$$\theta_{XY}(\lambda_j) = \text{atan} \left[\frac{-Q_{XY}(\lambda_j)}{C_{XY}(\lambda_j)} \right] \approx -\frac{3}{2^{j+1}} \pi\tau. \quad (55)$$

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

$$\tau_{XY}(\lambda_j) \stackrel{\text{df}}{=} -\frac{\theta_{XY}(\lambda_j)}{2\pi f_{j0}}, \quad (56)$$

where f_{j0} is the centre frequency of the octave band, computed as the arithmetic mean of the upper and lower cut-off frequencies, i.e. $f_{j0} = \frac{3}{2^{j+2}}$. Then, in our example we have: $\tau_{XY}(\lambda_j) \approx \tau$. Similarly, the wavelet gain is defined as:

¹⁴ Assuming that $\sin x \approx x$ for $x \leq \frac{\pi}{8}$, the last approximation will work for $\tau \leq 2^{j-2}$ ($j = 2, 3, \dots$)

– comp. Percival and Walden (2000: 344).

$$G_{XY}(\lambda_j) \stackrel{\text{df}}{=} \frac{A_{XY}(\lambda_j)}{S_X(\lambda_j)} \quad (57)$$

and in the example above for $\tau \leq 2^{j-2}$ we get: $G_{XY}(\lambda_j) \approx \alpha$.

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

$$\begin{aligned} X_t &= S_t + \eta_{Xt}, \\ Y_t &= \alpha S_{t-\tau} + \eta_{Yt}, \end{aligned} \quad (58)$$

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

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

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 $\tau_{XY}(\lambda_j) \approx \tau$.

The estimators of (46)–(48) and (56) 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. This smoothing is particularly necessary for the 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 present the 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 (49) with $\tau = 1$. Figure 2 illustrates the case, when the first stage filters are different and fulfil 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 two, while the simplified method introduces more bias in both the coherence and time delay estimations, especially at the second decomposition level. Another observation concerns the small sample bias of the wavelet coherence estimator, which clearly increases with the scale.

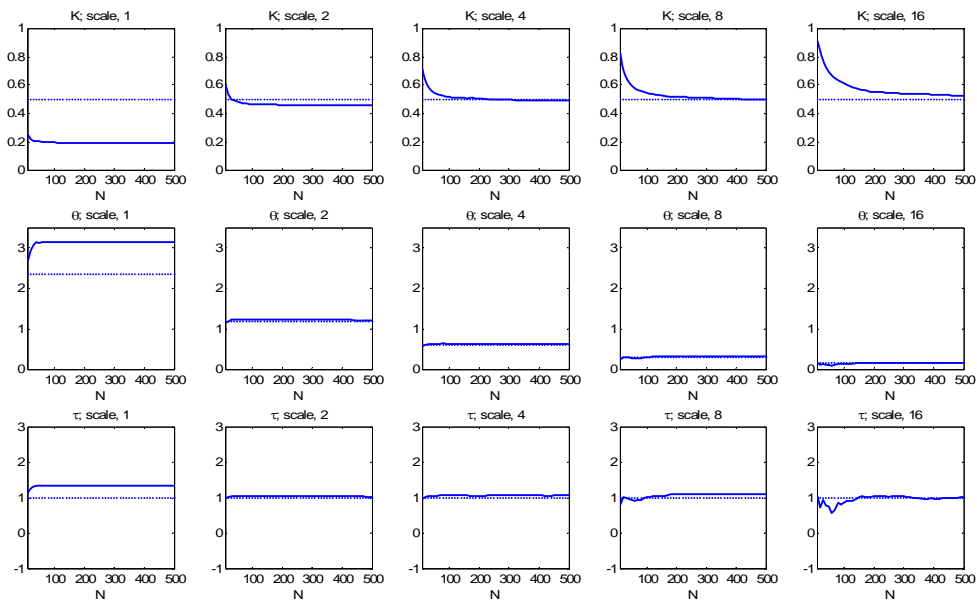


Figure 2. Estimates of the wavelet coherence, wavelet phase spectrum and wavelet time delay with the modified method; the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k412; figures present the theoretical Fourier quantities (thin dotted 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 $\tau = 1$, $\alpha = 1$ and X_t and η_t being two independent AR(1) processes with autoregressive parameters 0.8 and unit error variances (thick solid lines).

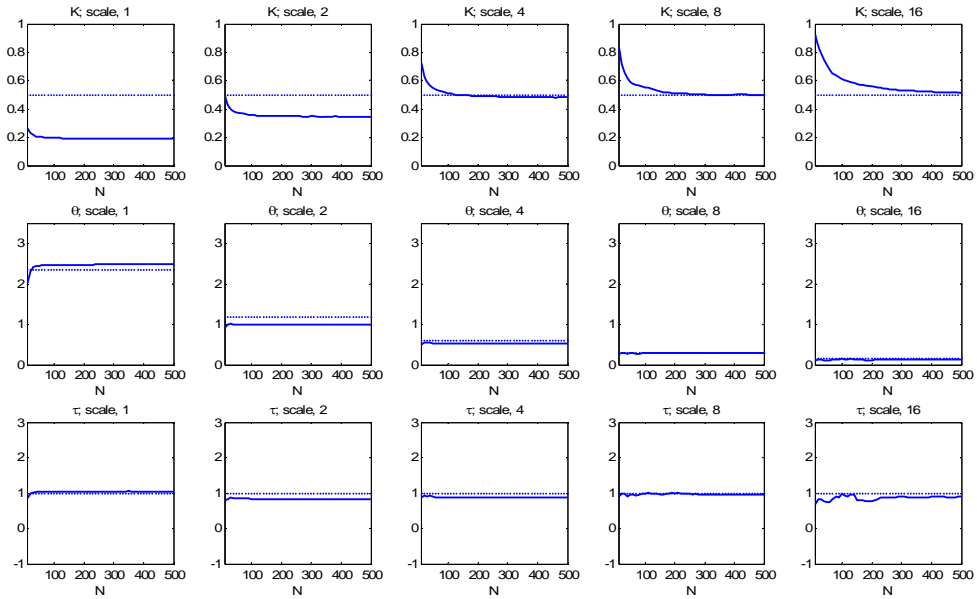


Figure 3. Estimates of the wavelet coherence, wavelet phase spectrum and wavelet time delay with the simplified method; k412 is applied at each decomposition level; see Figure 2 for more details.

To construct CIs for the wavelet coherence the multivariate process in the form:

$$\mathbf{P}_{j,t} = \left[\left| \tilde{W}_{j,t}^X \right|^2, \left| \tilde{W}_{j,t}^Y \right|^2, \Re\left(\tilde{W}_{j,t}^X \overline{\tilde{W}_{j,t}^Y}\right), \Im\left(\tilde{W}_{j,t}^X \overline{\tilde{W}_{j,t}^Y}\right) \right]^T \quad (59)$$

is considered together with the function:

$$g([a, b, c, d]^T) = \frac{c^2 + d^2}{ab}. \quad (60)$$

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

$$\sqrt{\tilde{N}_j} \left(\hat{K}_{XY}(\lambda_j) - K_{XY}(\lambda_j) \right) \sim AN(0, R_{abc,j}(0)), \quad (61)$$

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 $\nabla g([a,b,c,d]^T)$ denotes the vector of partial derivatives. As suggested by Whitcher and Craigmile (2004), an estimate of the large sample variance $R_{abc,j}(0)/\tilde{N}_j$ of $\hat{K}_{XY}(\lambda_j)$ can 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 utilising the sample cross-covariance of elements of the vector process $\mathbf{P}_{j,t}$. Then, an approximate large sample CI for the wavelet coherence has the form:

$$\hat{K}_{XY}(\lambda_j) \pm \varsigma_{\frac{\alpha}{2}} \left(\frac{\hat{R}_{abcd,j}(0)}{\tilde{N}_j} \right)^{0.5}, \quad (62)$$

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

To construct confidence intervals for the WPA we assume that $C_{XY}(\lambda_j) \neq 0$ and take:

$$g([a,b,c,d]^T) = \tan^{-1}\left(\frac{d}{c}\right). \quad (63)$$

Then, applying the delta method, we arrive at:

$$\sqrt{\tilde{N}_j} (\hat{\theta}_{XY}(\lambda_j) - \theta_{XY}(\lambda_j)) \sim AN(0, R_{abc,j}(0)), \quad (64)$$

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

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

Then, the confidence intervals are given in the form similar to (62). Multiplying them by a constant will produce approximate CIs for the wavelet time delay. Finally, assuming that the wavelet gain (57) is positive, similar reasoning to the presented above will provide approximate CIs for this quantity.

Estimates of the large sample variance of the wavelet spectra estimators can be obtained via nonparametric kernel methods. We examined the properties of two kernel estimators: one based on the popular Bartlett kernel and the other based on

the truncated (rectangular) kernel, which however can provide negative values of the estimates. The speed of convergence of the kernel estimators to the experimentally driven true values depends on the kind of the wavelet spectrum, the type of the bivariate data-generating process and the scale of the analysis. We concentrated on a linear regression with delay, for which different stationary AR(1) models for X_t and η_t were considered. For lower scales – $\lambda = 2$ (4) – good results in terms of unbiasedness were obtained, even in samples of length as small as $N = 50$ (75) wavelet coefficients, while for $\lambda = 8$ acceptable approximations started with $N = 200$, and even more data were necessary for $\lambda = 16$. The experimentally determined best values of the truncation parameters M are both sample size- and scale-dependent.¹⁵

2.5 Wavelet Estimation of the Time Delay – a Simulation Study

In this section, we summarise the results of our simulation experiments examining the statistical properties of two wavelet estimators of the time delay parameter. We concentrate on model (49) and analyse mainly the small sample performance of the estimators in order to recommend a method for examining short- and medium-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 analysing changing patterns of business and growth cycle synchronisation. The estimators compared are: the wavelet cross-correlator (WCC) that is based on maximising the values of the cross-covariance estimates, i.e.:

$$\hat{\tau}_j^{WCC} = \arg \max_k \text{Cov}_{\tilde{W}_j^X \tilde{W}_j^Y}(k) = \arg \max_k \text{Cov}(\tilde{W}_{j,t}^X, \tilde{W}_{j,t+k}^Y), \quad (65)$$

and the estimator of the wavelet time delay (56), i.e.:

¹⁵ For exactly the same DGP and data lengths as in the case of the estimates in Figures 2 and 3, acceptable results in larger samples were obtained with $M = 1,5 \cdot 2^j$ for the rectangular kernel and $M = 2,5 \cdot 2^j$ for the Bartlett kernel, where $j = 2, 3, \dots$ denotes the decomposition level. More details of this examination are available upon request. Owing to a programming error, the conclusions from this part of our simulations differ slightly compared with those reported in a previous version of the paper.

$$\hat{\tau}_j^{WPA} \stackrel{\text{df}}{=} -\frac{\hat{\theta}_{XY}(\lambda_j)}{2\pi f_{j_0}}, \quad (66)$$

which we call further the WPA-based delay estimator. It is worth stressing that we do not maximise the absolute value of the cross-covariance in (65) since a large negative covariance would be interpreted as an anti-phase relationship.

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

$$\min \sigma^2 = \frac{1}{N \int_{-0,5}^{0,5} (2\pi f)^2 \frac{K(f)}{1-K(f)} df}, \quad (67)$$

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 (67) predicts that the variance of an optimal estimator decreases with the value of coherence, the signal bandwidth and the centre frequency. For the WCC, it is known that for jointly stationary processes and large enough data samples the Cramér-Rao lower bound 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, the 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 SNRs (see the simulation results in Scarbrough et al. 1981, Carter 1987). Regarding the 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 the properties of the estimators (65) and (66), it is worth underlining that the WPA-based method enables us to estimate delays that are not integer multiples of the sampling period. In addition, it guarantees a maximal time localisation limited only by the length of the applied filter. By contrast, the WCC makes it possible to estimate delays that are longer than half of the centre period, $1/f_{j_0}$, it is asymptotically unbiased, even for the first stage of analysis, and it can be based on shorter filters, as the approximately analytic complex wavelet filters

are usually longer than are the 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 was used:

$$\begin{aligned} Y_t &= \alpha X_{t-\tau} + \eta_t, \\ X_t &= \beta X_{t-1} + \varepsilon_t, \\ \eta_t &= \gamma \eta_{t-1} + \xi_t, \end{aligned} \tag{68}$$

$$\begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right).$$

Figures 4–6 present the comparisons of small sample biases, root mean square errors and large sample standard errors of the two estimators for the cases: $\alpha = 1$; $\tau = 1$; $\beta = \gamma = 0,8$; $SNR = \frac{\sigma_1}{\sigma_2} = \frac{1}{2}, 1, 2, 3, 4$, and the following wavelet filters: la12 for the WCC and la12 (first stage) + k4l2 (higher stages) for the WPA-based estimator.¹⁶ The search range for the WCC was $\pm 10 + \tau$. In each case, 1000 replications were run.¹⁷ 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 summarised below:¹⁸

- For wide ranges of SNRs and scales, the WPA-based estimator is better than the WCC in small samples (see Figure 5). For the majority of outcomes, the relative efficiency of the two methods defined as $\frac{RMSE(WPA)}{RMSE(WCC)}$

¹⁶ These wavelet filters were chosen because of their popularity and also to guarantee maximal similarity in the 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 and real part of k4l2 + k4l2. Selesnick's HWP(K, L) filters outperformed the two Q-shift filters that were also considered, i.e. kin and tkp12.

¹⁷ All computations, including the empirical part, were executed in Matlab. Numerical codes are available by e-mailing the author.

¹⁸ We also comment shortly on other experiments we performed. More detailed results are available upon request.

increases with the scale, the sample size and partially also with the SNR. In large samples (see Figure 6), the WCC dominates the WPA-based estimator or their performance is similar. In small samples, the relative efficiency of these methods depends largely on the search range for the WCC, although for similar ranges of delays for both methods the WPA-based estimator gives better results.

- In larger samples, the root mean square errors for both estimators increase with the scale. It is generally advisable to assume smoothing windows with a length proportional to the scale (comp. Cohen and Walden 2010). In addition, we can see that even in large samples in the case of high decomposition levels and low SNRs both estimates can be very imprecise.

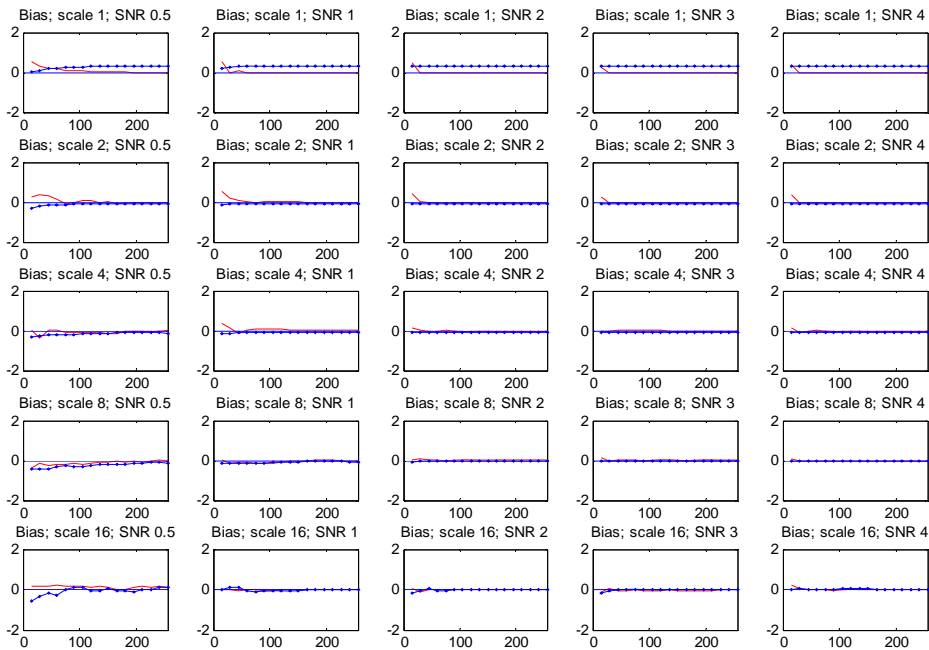


Figure 4. Small sample bias of time delay estimators; lines with and without markers correspond to the WPA and the WCC methods, respectively; samples consist of 15, 30, 45, ..., 255 wavelet coefficients unaffected by the boundary.

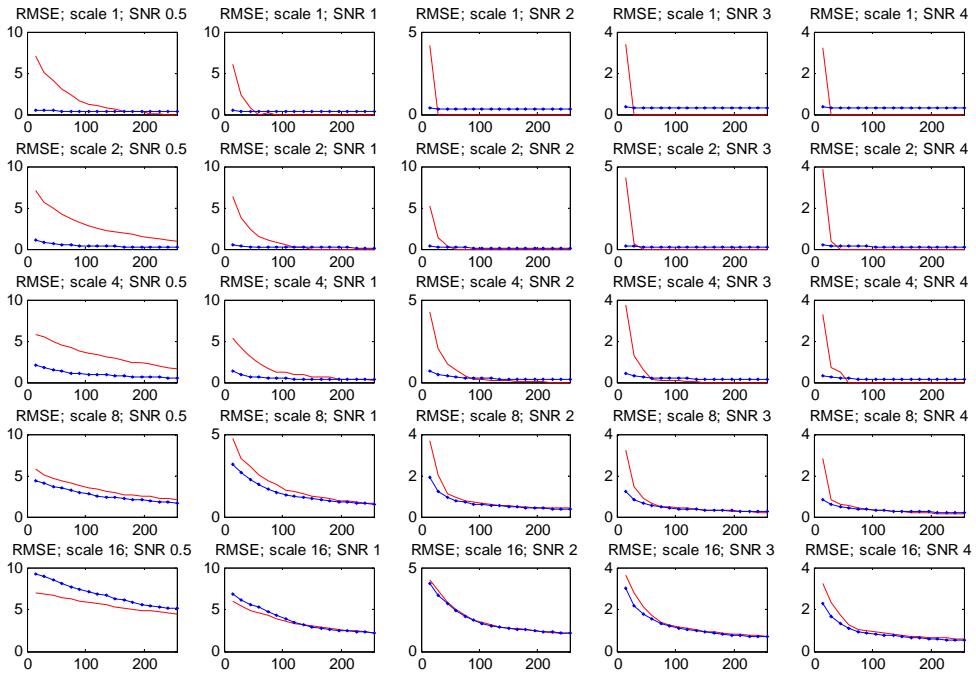


Figure 5. Root mean square errors of time delay estimators in small samples; lines with and without markers correspond to the WPA and the WCC methods, respectively; samples consist of 15, 30, 45, ..., 255 wavelet coefficients unaffected by the boundary.

- For low SNR, both estimators show small sample bias, although with opposite signs: the WPA-based estimator – towards 0, while the cross-correlator – in the opposite direction (see Figure 4). This suggests that when an estimate of the time delay parameter is needed, the usual biased estimator of the cross-covariance might be preferred in small samples. Our other 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.

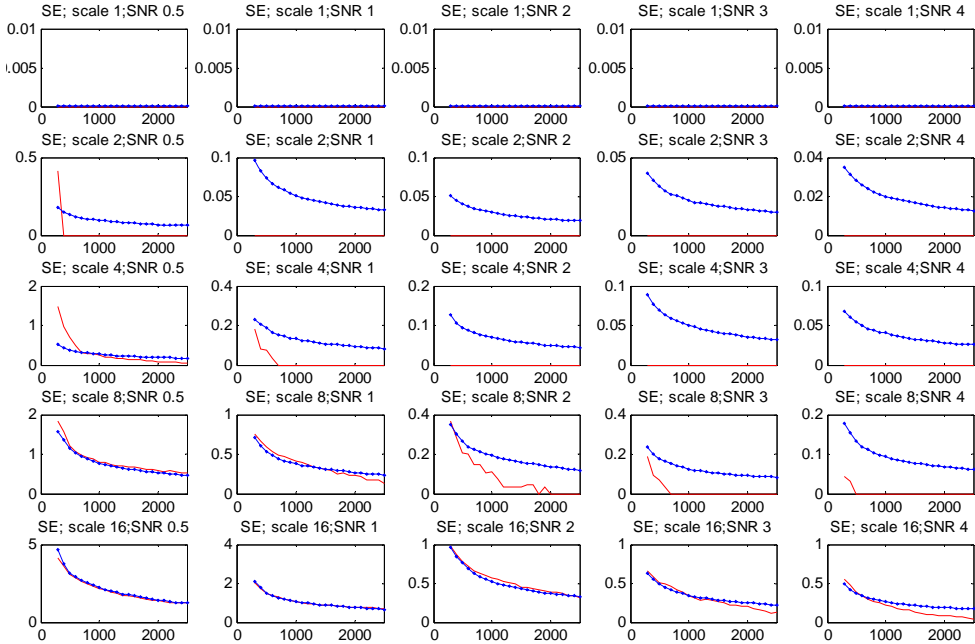


Figure 6. Standard errors of time delay estimators in large samples; lines with and without markers correspond to the WPA and the WCC methods, respectively; samples consist of 300, 400, ..., 2500 wavelet coefficients unaffected by the boundary.

- In a different set of simulation experiments, in which the WPA-based estimator (66) is replaced with its integer-valued counterpart:

$$\hat{\tau}_j^{WPA*} \stackrel{\text{df}}{=} \text{round} \left(-\frac{\hat{\theta}_{XY}(\lambda_j)}{2\pi f_{j0}} \right), \quad (66^*)$$

the outcomes in small and moderate samples are even more in favour of the WPA-based estimator as (66*) removes the bias caused by the centre frequency imprecision.

- Other experiments not presented here indicate that if the bivariate data generating process is jointly stationary, including observations affected by the periodic extension at the end of the sample can give better small sample results with both methods at higher scales, where the number of affected coefficients is large. This, however, does not take place for non-stationary processes with a stationary delay, which exhibit a slowly varying local mean. In such a case, a periodic extension with level adjustment should provide better results.
- All the above observations are unchanged across different wavelets, although the outcomes obtained with the WPA-based method depend on the analytic properties of the HWP filters. Good analytic wavelets, however, produce almost identical results (for example, k412, k313). Different values of β and γ (including the non-stationary case) do not change the conclusions either.

The main finding of this section is that in business cycle studies, which are typically based on relatively short time series, the WPA-based methodology seems to be particularly attractive and can be used at least as a supplementary method. It is worth stressing that, in addition to its good localisation properties, the WPA-based estimator is also simple and efficient computationally. For these reasons, we believe it can be recommended for empirical analysis on business cycle synchronisation.

3 Empirical Examination

In the empirical study we use 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 and Spain. In addition, 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, the local wavelet variance analysis is performed, while the second deals with the local and global wavelet analysis of synchronisation.

3.1 Business Cycle Variability

Our examination of business cycle variability was performed with the help of the d4 Daubechies wavelet filter of length 4, which guarantees very good localisation properties. The results are presented in Figures 7 and 8.

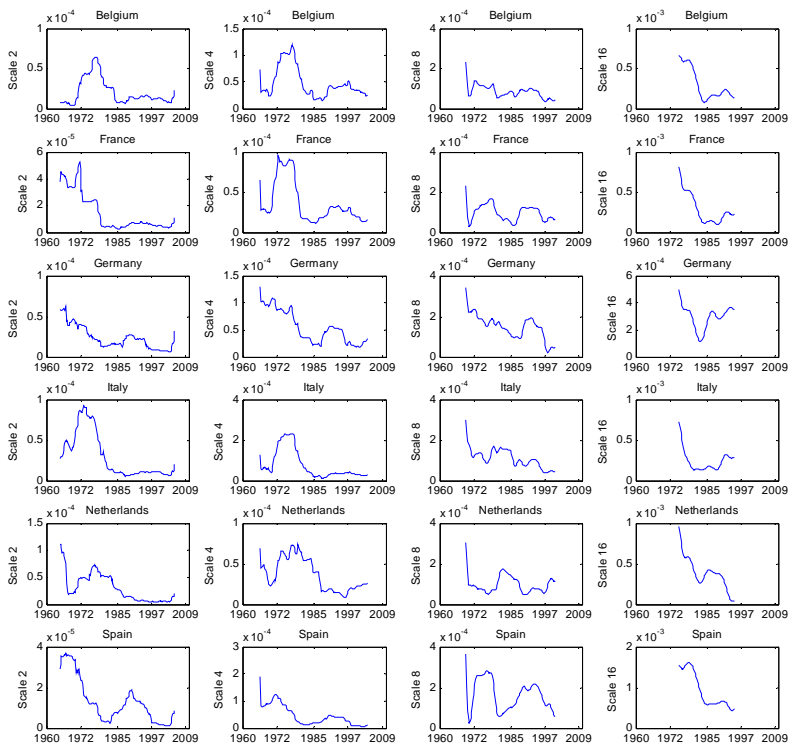


Figure 7. Running wavelet variance for scales 2, 4, 8 and 16 corresponding to oscillations with period lengths 4–8 (1–2 years), 8–16 (2–4 years), 16–32 (4–8 years) and 32–64 (8–16 years); results obtained with d4 Daubechies filter of length 4 and windows of 30 wavelet coefficients unaffected by circularity after aligning them to the observations in the sample; step = 1.

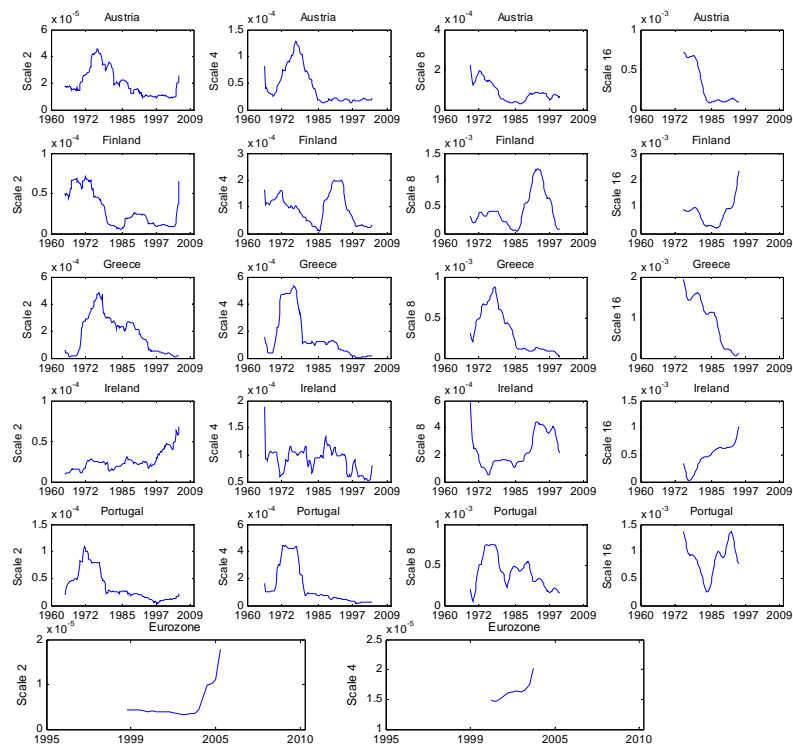


Figure 8. Running wavelet variance for scales 2, 4, 8 and 16 corresponding to oscillations with period lengths 4–8 (1–2 years), 8–16 (2–4 years), 16–32 (4–8 years) and 32–64 (8–16 years), except for the euro zone data, where scales 2 and 4 are only considered; results obtained with d4 Daubechies filter 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 very similar patterns of volatility changes across countries in our sample, except for Finland (scales 4 and 8) and Ireland (all scales). Furthermore, the contributions of different scales to the total variance as well as the estimates of the wavelet variance alone vary little across the economies. For some countries, we observe a systematic decline in the variance at all decomposition levels, which started at the beginning of our sample, as is the case

for Germany (except for the more volatile period around the reunification as well as for the highest scale) and Spain. It is seen that the oil price shocks of 1973 and 1979 were captured almost entirely by the shortest components of business cycle fluctuations. Thanks to this, the scale 8 wavelet variance provides a clearer view of the Great Moderation, revealing that the process might have started well before the mid-1980s (comp. Aguiar-Conraria and Soares 2010 for similar evidence for the United States obtained with the continuous wavelet methodology). Moreover, the most recent perturbations (the financial crisis of 2007–2009) are becoming apparent at the lowest decomposition level as can be seen in the case of the euro area GDP.

3.2 Business Cycle Synchronisation

Business cycle synchronisation in the euro area was examined with the help of local wavelet correlations, global and local wavelet coherences and global and local wavelet time delays. Figure 9 presents the running wavelet correlations for scales 4, 8 and 16 computed after the MODWT based on d4 wavelet had been applied to the observations. We decided to treat Germany as the reference country because of the high correlation between its scale 4 wavelet coefficients and those for the euro zone compared with the appropriate correlation for France and the euro zone (see Figure 9). The most interesting finding is that for the majority of countries in the sample we observe a systematic increase in the strength of the instantaneous relationships between the 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 patterns of the dynamic correlations 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 8 years and above) there often is the opposite tendency, i.e. a cyclical divergence. For Greece, this seems to take place also at scale 8.

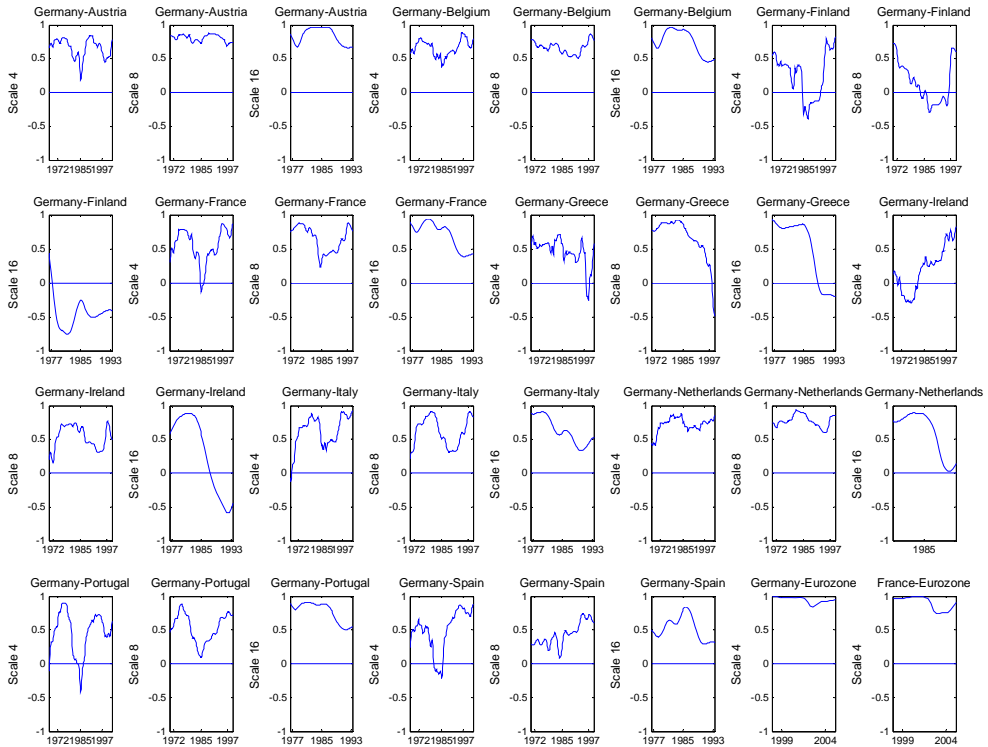


Figure 9. Running wavelet correlation for scales 4, 8 and 16 corresponding to oscillations with period lengths 8–16 (2–4 years), 16–32 (4–8 years) and 32–64 (8–16 years), except for the euro zone data, where scale 4 is only considered; results obtained with d4 Daubechies wavelet filter of length 4 and windows of 40 wavelet coefficients unaffected by circularity after aligning them to the observations in the sample; step = 1.

Figures 10 and 11 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 in the first stage of the analysis. The global examination shows moderate to high dependencies at all leads and lags between Germany and

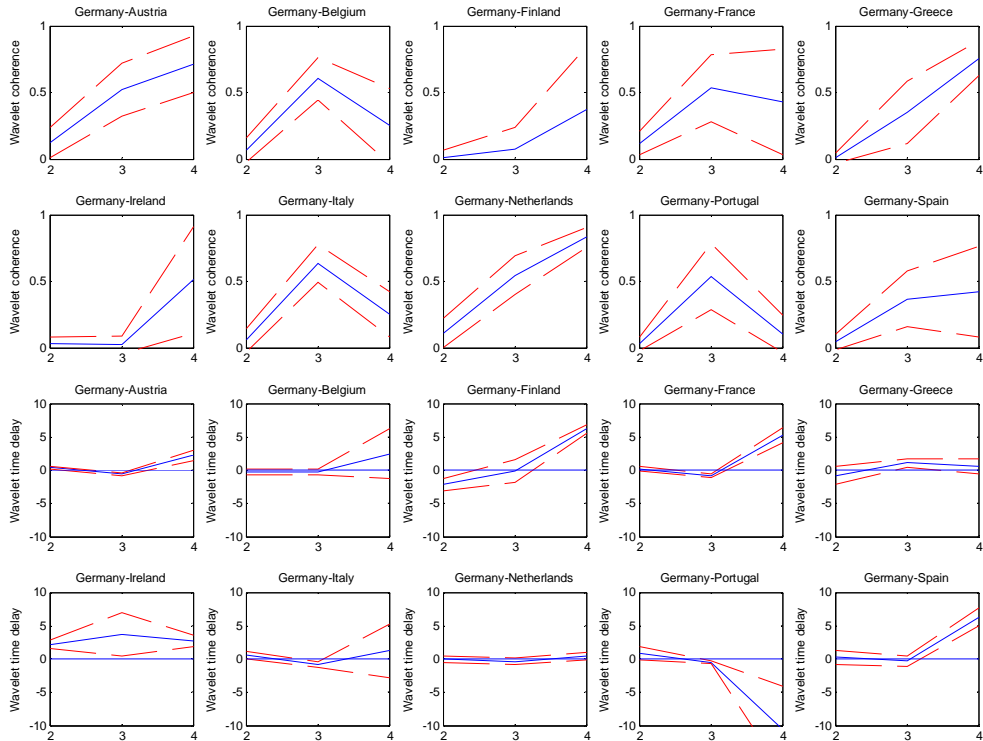


Figure 10. Wavelet coherence and wavelet time delay for scales 2, 4 and 8 (decomposition levels 2, 3 and 4) corresponding to oscillations with period lengths 4–8 (1–2 years), 8–16 (2–4 years) and 16–32 (4–8 years) together with large sample 90% CIs; the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k4l2; only wavelet coefficients unaffected by circularity are considered; the Bartlett kernel is used in variance estimation with the truncation parameter $M = 10$.

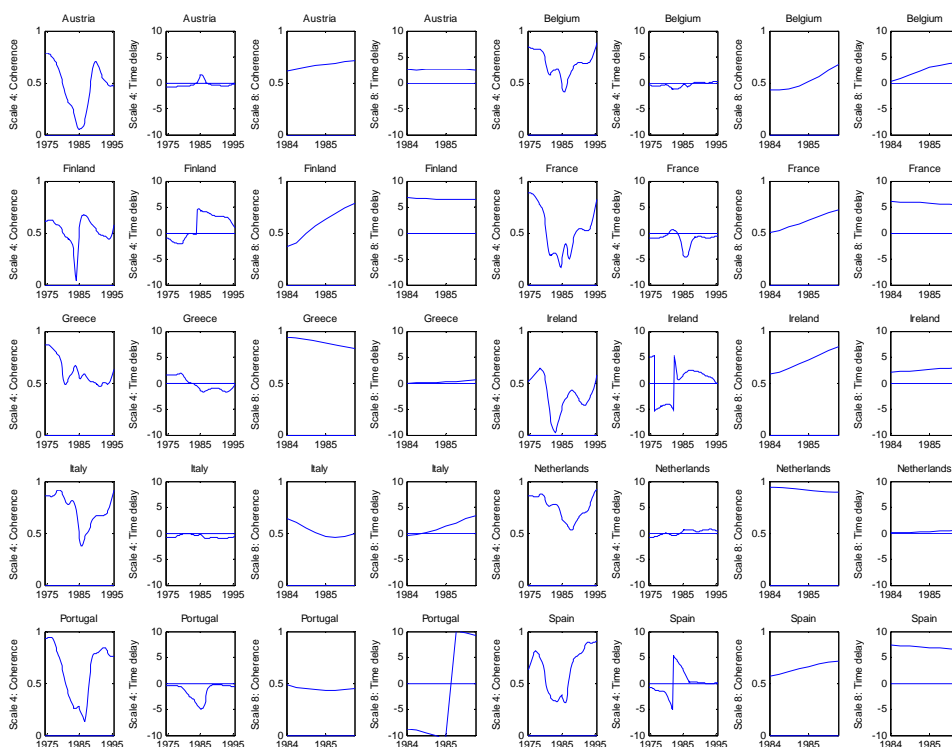


Figure 11. Running wavelet coherence and wavelet time delay for scales 4 and 8 corresponding to oscillations with period lengths 8–16 (2–4 years) and 16–32 (4–8 years); the first stage filters are la12 and its one-sample shifted version and the complex filter for the higher levels is k4l2; data windows consist of 30 non-boundary wavelet coefficients for scale 4, and 40 – for scale 8, circularly shifted to align them to the real data; step = 1.

the majority of countries in the sample at the third decomposition level (for shorter cycles). For Austria and the Netherlands, high correlations are also present at the fourth level. These two countries' cycles show the strongest relationship with the German business cycle. Positive delays mean that the German cycle is behind the other countries' cycles, as is the case for the long French cycle and the short Irish cycle. Instantaneous dependencies with the German business cycle take place for countries such as Belgium, the Netherlands, Greece and Italy as well as for the

lower decomposition levels for almost all countries in the sample. The local analysis in Figure 11, performed in windows of a scale-dependent size, reveals that the shorter cycles are becoming more synchronised from the middle of our sample, while in the case of the longer cycles the lead-lag relations are quite stable over time (partially because only a couple of coefficient windows were available), although the wavelet coherences are rising.

The overall conclusion from the real and complex wavelet analysis is that the synchronisation between euro zone business cycles started to rise after the first important steps toward European integration were taken. This is in line with the endogeneity hypothesis of the optimum currency area criteria as stated by Frankel and Rose (1998). Finally, Figure 12 presents a comparison of local wavelet time delay estimates obtained using the WCC and WPA methods. It turns out that these methods produce largely similar results.

4 Conclusions

The non-decimated real and complex discrete wavelet transforms provide a summary of the evolutionary spectral and cross-spectral properties of processes under scrutiny with high computational efficiency, good localisation properties and without an excessive redundancy of information that occurs when using the continuous wavelet methodology. These features together with a 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 discussed some of the questions arising in the discrete wavelet analysis of popular bivariate spectral quantities such as the amplitude, coherence and phase spectra and the frequency-dependent time delay. In particular, we showed how the wavelet bivariate spectra can serve to approximate the corresponding Fourier quantities and discussed certain implementation issues. Our simulation study of the properties of two wavelet estimators of the time delay parameter pointed out the 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.

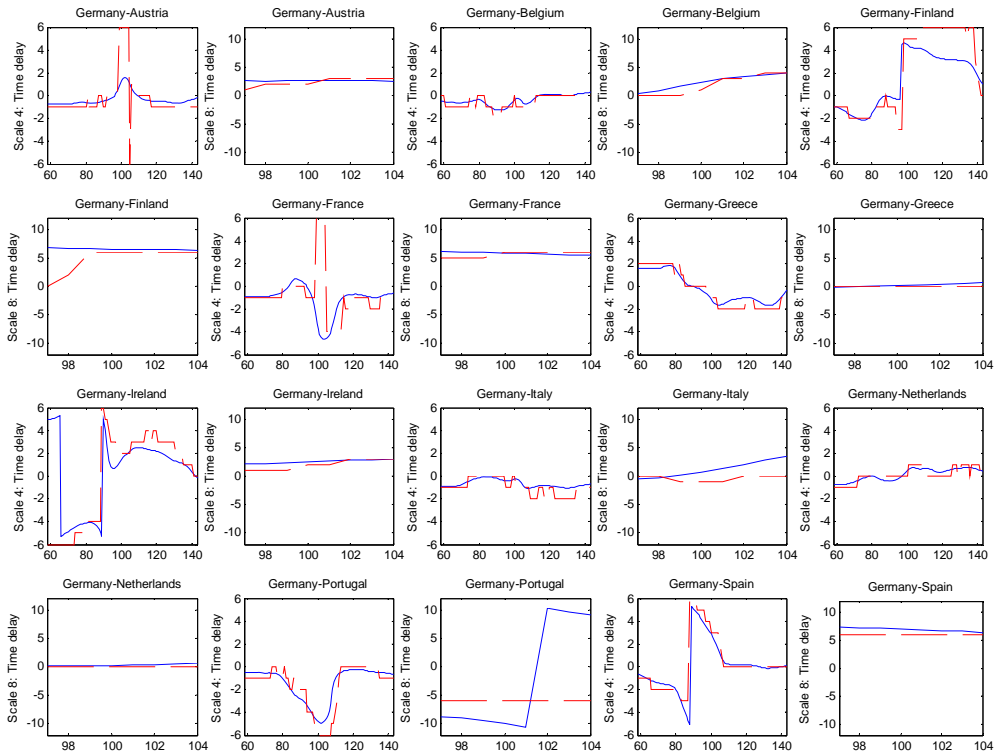


Figure 12. Running wavelet estimates of time delay for scales 4 and 8 corresponding to oscillations with period lengths 8–16 (2–4 years) and 16–32 (4–8 years); the solid blue line is the result obtained with the WPA-based estimator and the dashed red line – with the WCC; in the WPA-based method the first stage filters are la12 and its one-sample shifted version and the higher level filter is k412; the WCC is based on la12; the search ranges for the WCC are $[-6, 6]$ and $[-12, 12]$ for scales 4 and 8, respectively; data windows consist of 30 wavelet coefficients unaffected by circularity for scale 4, and 40 – for scale 8, the coefficients were circularly shifted to align them to the real data; step = 1; the numbers on the horizontal axis are the mid-points of the subsamples.

The complex discrete wavelet methodology was illustrated with an examination of business cycle synchronisation in the euro zone. The study was also supplemented with a wavelet analysis of the 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 the early start of the Great Moderation in Europe.

Acknowledgements

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 synchronisation and forecasting”.

References

- Aguiar-Conraria, L., and Soares, M.J. (2009). Business Cycle Synchronisation 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
- Antoniadis, A., and Gijbels, I. (2002). Detecting Abrupt Changes by Wavelet Methods. *Journal of Nonparametric Statistics*, 14 (1–2): 7–29.
<http://www.tandfonline.com/doi/abs/10.1080/10485250211396>
- Blanchard, O., and Simon, J. (2001). The Long and Large Decline in U.S. Output Volatility. *Brookings Papers on Economic Activity*, 32 (1): 135–174.
<http://ideas.repec.org/a/bin/bpeajo/v32y2001i2001-1p135-174.html>
- Carter, G.C. (1987). Coherence and Time Delay Estimation. *Proceedings of the IEEE*, 75 (2): 236–255.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1457990&tag=1

- Cohen, E.A.K., and Walden, A.T. (2010). A Statistical Study of Temporally Smoothed Wavelet Coherence. *IEEE Transactions of Signal Processing*, 58 (6): 2964–2973.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?tp=&arnumber=5411786
- Crowley, P.M., and 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., and 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., and 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.
<http://ideas.repec.org/a/bla/jecsur/v22y2008i2p234-273.html>
- 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.
<http://onlinelibrary.wiley.com/doi/10.1002/fut.20294/abstract>
- Frankel, J., and Rose A. (1998). The Endogeneity of the Optimum Currency Area Criteria. *The Economic Journal*, 108 (449): 1009–1025.
<http://ideas.repec.org/a/ecj/econjl/v108y1998i449p1009-25.html>
- Gallegati, M., and Gallegati M. (2007), Wavelet Variance Analysis of Output in G-7 Countries. *Studies in Nonlinear Dynamics and Econometrics*, 11(3), art. 6.
<http://ideas.repec.org/a/bpj/sndecm/v11y2007i3n6.html>
- Gençay, R.F., Selçuk, F., and Whitcher, B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, San Diego.

- Gonçalves, E.S., Rodrigues, M., and Soares T. (2009). Correlation of Business Cycles in the Euro Zone. *Economics Letters*, 102 (1): 56–58.
<http://ideas.repec.org/a/eee/ecolet/v102y2009i1p56-58.html>
- Granger, C.W.J., and Hatanaka, M. (1964). *Spectral Analysis of Economic Time Series*. Princeton University Press, Princeton.
<http://www.cabdirect.org/abstracts/19651802116.html?freeview=true>
- Huang, N.E., and Shen, S.S.P. (eds.) (2005). *Hilbert-Huang Transform and Its Applications*. World Scientific, Singapore.
http://www.ebook3000.com/The-Hilbert-Huang-Transform-and-Its-Applications_39677.html
- Jagrič, T., and Ovin, R. (2004). Method of Analyzing Business Cycles in a Transition Economy: The Case of Slovenia. *The Developing Economies*, 42 (1): 42–62.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1746-1049.2004.tb01015.x/abstract>
- 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.
<http://www.ingentaconnect.com/content/ap/ha/2001/00000010/00000003/art00343>
- Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Academic Press, San Diego.
- Percival, D.B. (1995). On Estimation of the Wavelet Variance. *Biometrika*, 82 (3): 619–631.
<http://www.jstor.org/stable/2337538>
- Percival, D.B., and Walden, A.T. (2000), *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.
http://www.ebook3000.com/Wavelet-Methods-for-Time-Series-Analysis_24274.html

- Piersol, A.G. (1981). Time Delay Estimation Using Phase Data. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29 (3): 471–477.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1163555&tag=1
- Priestley, M.B. (1981). *Spectral Analysis and Time Series*. Academic Press, London.
- Raihan, S.M., Wen, Y., and 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., and Nunes, L.C. (2009). International Comovement of Stock Market Returns: A Wavelet Analysis. *Journal of Empirical Finance*, 16 (4): 632–639.
<http://ideas.repec.org/a/eee/empfin/v16y2009i4p632-639.html>
- Scarbrough, K., Ahmed, N., and 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.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1163615
- Selesnick, I.W. (2001). Hilbert Transform Pairs of Wavelet Bases. *IEEE Signal Processing Letters*, 8 (6): 170–173.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=923042
- Selesnick, I.W. (2002). The Design of Approximate Hilbert Transform Pairs of Wavelet Bases. *IEEE Transactions on Signal Processing*, 50 (5): 1144–1152.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=995070
- Selesnick, I.W., Baraniuk, R.G., and 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.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1550194
- Serroukh, A., Walden, A.T., and 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.
<http://www.jstor.org/stable/2669537>
- Tay, D.B.H., Kingsbury, N.G., and Palaniswami, M. (2006). Orthonormal Hilbert-pair of Wavelets with (Almost) Maximum Vanishing Moments. *IEEE Signal Processing Letters*, 13 (9), 533–536.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1673413

- Torrence, C., and Compo, G.P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79 (1): 61–78.
http://atoc.colorado.edu/research/wavelets/bams_abstract.html
- 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., and Craigmile, P.F. (2004). Multivariate Spectral Analysis Using Hilbert Wavelet Pairs. *International Journal of Wavelets, Multiresolution and Information Processing*, 2 (4): 567–587.
<http://www.citeulike.org/user/tingying/article/928669>
- Whitcher, B.J., Craigmile, P.F., and Brown, P. (2005). Time-varying Spectral Analysis in Neurophysiological Time Series Using Hilbert Wavelet Pairs. *Signal Processing*, 85 (11): 2065–2081.
<http://dl.acm.org/citation.cfm?id=1139865>
- Whitcher, B.J., Guttorp, P., and 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., and Lui, X. (2003). Modelling and Forecasting by Wavelets, and the Application to Exchange Rates. *Journal of Applied Statistics*, 30 (5): 537–553.
<http://ideas.repec.org/a/taf/japsta/v30y2003i5p537-553.html>
- Yogo, M. (2008). Measuring Business Cycles: A Wavelet Analysis of Economic Time Series. *Economics Letters*, 100 (2): 208–212.
<http://ideas.repec.org/a/eee/ecolet/v100y2008i2p208-212.html>

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

<http://dx.doi.org/10.5018/economics-ejournal.ja.2011-16>

The Editor