The windowed scalogram difference: a novel wavelet tool for comparing time series

V. J. Bolós¹, R. Benítez², R. Ferrer³, R. Jammazi⁴

¹ Dpto. Matemáticas para la Economía y la Empresa, Facultad de Economía, Universidad de Valencia. Avda. Tarongers s/n, 46022 Valencia, Spain. e-mail: vicente.bolos@uv.es

² Dpto. Matemáticas, Centro Universitario de Plasencia, Universidad de Extremadura. Avda. Virgen del Puerto 2, 10600 Plasencia (Cáceres), Spain. e-mail: rbenitez@unex.es

³ Dpto. Economía Financiera y Actuarial, Facultad de Economía, Universidad de Valencia. Avda. Tarongers s/n, 46022 Valencia, Spain. e-mail: roman.ferrer@uv.es

⁴ Ipag Business School France & Ecole Nationale des Sciences de l'Informatique, Manouba, Tunisia. e-mail: jamrania2@yahoo.fr

> September 2015 (Revised: February 2019)

Abstract

We introduce a new wavelet-based tool called *windowed scalogram difference* (WSD), which has been designed to compare time series. This tool allows quantifying if two time series follow a similar pattern over time, comparing their scalograms and determining if they give the same weight to the different scales. The WSD can be seen as an alternative to another tool widely used in wavelet analysis called *wavelet squared coherence* (WSC) and, in some cases, it detects features that the WSC is not able to identify. As an application, the WSD is used to examine the dynamics of the integration of government bond markets in the euro area since the inception of the euro as a European single currency in January 1999.

1 Introduction

Quantifying relationships between time series has been historically one of the most frequently addressed issues by most scientific disciplines. A large number of mathematical and statistical methods have been developed and applied for measuring the strength and direction of relationships between time series. The great majority of these techniques have focused on the time domain. Correlation and regression analysis constitute the first and most popular tools to quantify the association between time series. Subsequently, a number of more sophisticated time series methods, including cointegration analysis [1], Granger causality tests [2], vector

autoregressive (VAR) models [3] or generalized autoregressive conditional heteroscedasticity (GARCH) models [4,5] have been also used for the same purpose. In addition, several newly introduced techniques, such as the combined cointegration approach [6], the quantileon-quantile method [7], the quantile correlation approach [8], the nonlinear autoregressive distributed lag (NARDL) model [9], or the quantile autoregressive distributed lag (QADL) method [10] are also very useful to assess the linkages among time series. An obvious limitation of these approaches is that they are restricted to one or at most two time scales, i.e. the short run and the long run. In some fields, such as economics and finance, traditional time domain models are insufficient to describe precisely the linkage between variables. For example, financial markets are complex systems consisting of thousands of heterogeneous agents making decisions over a different time frame (from minutes to years), so that the relationships between economic and financial variables may vary across time scales associated to different investment horizons of market participants (see [11]). To remedy this situation, a body of literature seeking to characterize the connection between time series at different frequencies has been also developed. The Fourier analysis represents the best exponent of this line of research focused on the frequency domain, although it has serious shortcomings. In particular, under the Fourier transform the time information is completely lost, so it is hard to distinguish transient relations or to identify structural changes. Therefore, this approach is not suitable for non-stationary processes (see [12]).

In this context, the wavelet theory is a very versatile methodology that allows to study a wide range of different signal properties. Due to this great flexibility, wavelet methods have been applied to many disciplines such as geophysics [13,14], meteorology [15,16], engineering [17,18], medicine [19,20], image analysis [21,22], economics [11,23], or, for instance, recently they have been used for measuring the degree of non-periodicity of a signal [24]. Hence, the wavelet analysis emerges as an appealing alternative to the Fourier transform that takes into account both time and frequency domains simultaneously, whose primary advantage is its ability to decompose any signal into time scale components. This property offers a unique opportunity to study relationships between time series in both, time and frequency domains, at the same time. In fact, wavelet techniques can reveal interactions which would be, otherwise, hard to detect by using any other statistical procedure.

The aim of this paper is to propose a novel wavelet-based tool, called *windowed scalogram* difference (WSD), which has been designed to compare time series. As its name suggests, this new measure is based on the concept of wavelet scalogram, restricted, however, to a finite window in time and scale. The main feature of the WSD is that it allows to assess whether two time series, measured preferably in the same units, follow a similar pattern over time and/or across scales (or frequencies) through the comparison of their respective scalograms for different windows in time and scale. The WSD can be regarded as an alternative tool to the widely applied *wavelet squared coherence* (WSC) [14,16], in the sense that both measures serve to evaluate the level of association between two time series, although from slightly different perspectives. As a matter of fact, in some cases (see Figure 1), the WSD detects certain features that the WSC is not able to identify.

The paper is organized as follows. Section 2 introduces the concept of WSD, including some practical aspects and simulation results on the validity of this tool. In Section 3, the WSD is applied to real data to test its validity, examining the dynamics of the integration of government bond markets in the euro area since the inception of the euro in January 1999. Finally, Section 4 concludes the paper.

2 The windowed scalogram difference (WSD)

This section starts presenting some basic notions of wavelet theory and recalling the concept of wavelet scalogram. Subsequently, the concept of WSD is formally introduced as a tool for measuring the degree of similarity between two time series. Finally, some important practical aspects for the application of the WSD are discussed.

2.1 Basic concepts of Wavelets

A wavelet is a function $\psi \in L^2(\mathbb{R})$ with zero average (i.e. $\int_{\mathbb{R}} \psi = 0$), normalized ($\|\psi\| = 1$) and "centered" in the neighborhood of t = 0 ([25]). Scaling ψ by s > 0 and translating it by $u \in \mathbb{R}$, we can create a family of *time-frequency atoms* (also called *daughter wavelets*), $\psi_{u,s}$, as follows

$$\psi_{u,s}(t) := \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right). \tag{1}$$

Given a time series $f \in L^2(\mathbb{R})$, the continuous wavelet transform (CWT) of f at time u and scale s with respect to the wavelet ψ is defined as

$$Wf(u,s) := \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t)\psi_{u,s}^*(t) \,\mathrm{d}t,$$
(2)

where * denotes the complex conjugate. The CWT allows us to obtain the frequency components (or *details*) of f corresponding to scale s and time location u, thus providing a time-frequency decomposition of f.

On the other hand, the dyadic version of (1) is given by

$$\psi_{j,k}(t) := \frac{1}{\sqrt{2^k}} \psi\left(\frac{t-2^k j}{2^k}\right),\tag{3}$$

where $j, k \in \mathbb{Z}$ (note that there is an abuse of notation between (1) and (3), nevertheless the context makes it clear if we refer to (1) or (3)). It is important to construct wavelets so that the family of dyadic wavelets $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{R})$. Thus, any function $f \in L^2(\mathbb{R})$ can be written as

$$f = \sum_{j,k \in \mathbb{Z}} d_{j,k} \psi_{j,k},\tag{4}$$

where $d_{j,k} := \langle f, \psi_{j,k} \rangle$ is the discrete wavelet transform (DWT) of f at time $2^k j$ and scale 2^k . In fact, the DWT is the particular dyadic version of the CWT given by (2).

The *scalogram* of a time series f at a given scale s > 0 can be defined as

$$\mathcal{S}(s) := \left(\int_{-\infty}^{+\infty} |Wf(u,s)|^2 \,\mathrm{d}u \right)^{1/2}.$$
(5)

The scalogram of f at s is the L^2 -norm of Wf(u, s) (with respect to the time variable u) and captures the "energy" of the CWT of the time series f at this particular scale. It allows for the identification of the most representative scales of a time series, that is, the scales that contribute most to its total energy. The rationale behind the use of this measure is that if two time series show a similar pattern, then their scalograms should be very similar. In this regard, it is important to point out certain requirements for two time series have the same scalogram. The next proposition can be easily proved by considering appropriate changes of variables.

Proposition 2.1. Let $f \in L^2(\mathbb{R})$ be a time series and $c \in \mathbb{R}$. Then, -f(t), f(t) + c and f(t+c) have the same scalogram as f(t). Moreover, if the wavelet ψ is symmetric or antisymmetric, i.e. $\psi(-t) = \pm \psi(t)$ (e.g. Haar, Mexican Hat, Morlet, etc.), then f(-t) has also the same scalogram.

It is worth highlighting that most wavelets are "almost" symmetric or antisymmetric (e.g. Daubechies). In this case,

$$\pm f(\pm t + c_1) + c_2 \tag{6}$$

has approximately the same scalogram as f(t), where $c_1, c_2 \in \mathbb{R}$. So, we will say that (6) follows the same pattern as f(t).

2.2 The scalogram difference

The scalogram of a time series can be redefined by making a change of variable. Taking into account the decomposition of a function by means of the DWT (see (4)), it is convenient to use base 2 power scales, and thus

$$\mathcal{S}(k) := \left(\int_{-\infty}^{+\infty} |Wf\left(u, 2^k\right)|^2 \,\mathrm{d}u \right)^{1/2},\tag{7}$$

where $k \in \mathbb{R}$ is the binary logarithm of the scale (again, there is an abuse of notation that will be clarified by the context, this time between (5) and (7)), which is called *log-scale*. Note that in (7) we use the CWT and $k \in \mathbb{R}$, while in the framework of the DWT $k \in \mathbb{Z}$ (e.g. in (3)).

Hence, the *scalogram difference* of two time series f, g at log-scale k and log-scale radius r can be defined as

$$\mathcal{SD}_{r}(k) := \left(\int_{k-r}^{k+r} \left(\frac{\mathcal{S}(\kappa) - \mathcal{S}'(\kappa)}{\mathcal{S}(\kappa)} \right)^{2} \, \mathrm{d}\kappa \right)^{1/2}, \tag{8}$$

where S, S' represent the scalogram of f, g, respectively. It is expected that for two time series with similar behavior, their scalogram difference takes very small values.

Remark 2.1. Obviously, equation (8) has sense only when the two series considered are expressed in the same unit of measure. Otherwise, it will be necessary to somehow normalize the scalograms, but depending on the normalization method, some artificial results could be added. Due to this, it is recommended to consider only series with the same measurement unit. Nevertheless, there are some normalization alternatives for series with different units of measure. For example, if we are interested only in a finite interval of log-scales $[k_{min}, k_{max}]$ (e.g. if data are only available in this interval), we can normalize the scalograms so that their L^2 -norms are the same (e.g. 1) in that interval

$$\overline{\mathcal{S}}(k) = \left(\int_{k_{min}}^{k_{max}} |\mathcal{S}(\kappa)|^2 \,\mathrm{d}\kappa\right)^{-1/2} \mathcal{S}(k).$$
(9)

In this way, the total energy of the CWT of both series will be the same and so, we can compare the relative contributions of each scale^{*}. Note that for a proportionality constant $\alpha > 0$, the scalogram of αf is equal to $\alpha S(k)$ and, therefore, by taking $\alpha := \left(\int_{k_{min}}^{k_{max}} |S(\kappa)|^2 d\kappa\right)^{-1/2} > 0$ the scalogram of αf coincides with $\overline{S}(k)$. Hence, multiplying the original time series by this proportionality constant α implies a scalogram normalization[†].

Finally, we can apply (8) (using (9) instead of (7)) for obtaining a *normalized* scalogram difference. We can arrive at the same definition if we previously normalize the time series, i.e. by considering $\alpha f, \alpha' g$ instead of f, g, where α, α' are the appropriate proportionality constants. Thus, from now on we assume that the time series are normalized or they use the same unit of measure.

Remark 2.2. Note that (8) computes the difference relative to the scalogram of the first time series and, therefore, it is not commutative. Hence, instead of (8), it is preferable to consider the following *commutative scalogram difference*

$$\frac{1}{2} \left(\int_{k-r}^{k+r} \left(\frac{\mathcal{S}(\kappa) - \mathcal{S}'(\kappa)}{\mathcal{S}(\kappa)} + \frac{\mathcal{S}(\kappa) - \mathcal{S}'(\kappa)}{\mathcal{S}'(\kappa)} \right)^2 \, \mathrm{d}\kappa \right)^{1/2}.$$
(10)

^{*}It is worth mentioning that the interval considered $[k_{min}, k_{max}]$ must contain all the relevant log-scales because we are computing the total energy by means of the scalogram in this interval

[†]There are other possibilities that work well even if the interval of log-scales is not finite or the scalogram is not in L^2 , such as to normalize the scale of maximum amplitude (called *peak scale*) if it exists. In this case, the proportionality constant α must be the quotient between these amplitudes.

Remark 2.3. Some problems can arise in (8) or (10) when a scalogram is zero or close to zero for a given log-scale because the scalogram difference can take extremely high values or produce numerical errors. An easy solution is to multiply both scalograms by the same constant in order to make that their means (if they exist) are large in comparison with 1 (e.g. 100), and then to add 1 to both scalograms. So, in (8) and (10), it is recommended to consider

$$\beta \mathcal{S}(k) + 1, \qquad \beta \mathcal{S}'(k) + 1 \tag{11}$$

instead of the original scalograms S(k), S'(k), respectively, where β is a parameter given by 100/min {mean(S(k)), mean(S'(k))}. These new scalograms given by (11) will be practically proportional to the original ones and so, the results will be practically the same except, obviously, for the log-scales in which a scalogram takes values close to zero.

Remark 2.4 (Discrete and finite time series). In practice, time series are sampled with a finite frequency producing a finite sequence of values, which can be considered a sampling of a realization of the infinite ensemble of the real time series. When dealing with multiresolution analysis of a finite length time series, *border* or *edge effects* inevitably appear (see [25, Section 3.3]).

Let us consider a discrete set of times t_0, \ldots, t_N with stepsize Δt , i.e. $t_i = t_0 + i\Delta t$ for $i = 0, \ldots, N$, and a finite time series of N data f_0, \ldots, f_{N-1} defined over t_0, \ldots, t_{N-1} . We can still use the CWT and apply the definition (7) by considering, for example, the corresponding step function defined from the original time series. We could also consider a piecewise linearization instead of the step function, but the results are very similar for not too short series. Moreover, there are some alternatives to make f vanish outside the interval $[t_0, t_N]$, such as using *periodic wavelets*, folded wavelets or boundary wavelets (see [25]). However, these methods either produce large amplitude coefficients at the boundary or seriously complicate the calculations.

On the other hand, we also have a finite interval of scales to be studied: usually the minimum scale is assumed to be $s_{min} = 2\Delta t$ and the maximum scale is given by $s_{max} = N\Delta t/l$, where l is the size of the original wavelet function that we use. For instance, the size of the Daubechies n wavelet is 2n - 1, and the size of the Morlet wavelet is considered to be 8. In this case, the limits of the log-scales are $k_{min} = \log_2(s_{min}) = 1 + \log_2(\Delta t)$ and $k_{max} = \log_2(s_{max})$. Thus, we can adapt the expression (8) to this situation, so that the scalogram difference can be written as

$$\mathcal{SD}_{r}(k) := \left(\frac{2r}{k_{right} - k_{left}} \int_{k_{left}}^{k_{right}} \left(\frac{\mathcal{S}(\kappa) - \mathcal{S}'(\kappa)}{\mathcal{S}(\kappa)}\right)^{2} \mathrm{d}\kappa\right)^{1/2},\tag{12}$$

where $k_{left} := \max(k - r, k_{min})$ and $k_{right} := \min(k + r, k_{max})$. The factor $\frac{2r}{k_{right} - k_{left}}$ is optional to counteract the border effects in the log-scale interval. Its commutative version derived from (10) can be also considered.

2.3 The windowed scalogram difference (WSD)

The windowed scalogram of a time series f centered at time t with time radius τ can be defined as

$$\mathcal{WS}_{\tau}(t,k) := \left(\int_{t-\tau}^{t+\tau} |Wf\left(u,2^k\right)|^2 \,\mathrm{d}u \right)^{1/2}.$$
(13)

The windowed scalogram is simply the scalogram presented in (7) restricted to a given finite time interval $[t - \tau, t + \tau]$. Its principal feature is that it allows determining the relative importance of the different scales around a given time point.

Based on the above concept, the windowed scalogram difference (WSD) of two time series f, g centered at (t, k) with time radius τ and log-scale radius r is given by

$$\mathcal{WSD}_{\tau,r}(t,k) := \left(\int_{k-r}^{k+r} \left(\frac{\mathcal{WS}_{\tau}(t,\kappa) - \mathcal{WS}'_{\tau}(t,\kappa)}{\mathcal{WS}_{\tau}(t,\kappa)} \right)^2 \,\mathrm{d}\kappa \right)^{1/2},\tag{14}$$

. 10

where WS_{τ}, WS'_{τ} denote the windowed scalogram of f, g, respectively. The commutative version of the WSD (adapted from (10)) is recommended. Moreover, the reasoning given in Remark 2.3 can be easily adapted in order to avoid problems when the windowed scalograms take values close to zero.

As can be seen in (14), the WSD measures the difference between the windowed scalograms of two time series. It enables us to quantify the level of similarity between two time series for different finite time and scale intervals.

Remark 2.5 (Discrete and finite time series). By considering finite time series defined over a discrete set of times (as in Remark 2.4) and the corresponding step function, there also arise the aforementioned *border effects* in the windowed scalogram when $t - \tau < t_0$ or $t + \tau > t_N$. In this case, the expression (13) can be also adapted, writing the windowed scalogram as

$$\mathcal{WS}_{\tau}(t,k) := \left(\frac{2\tau}{t_{right} - t_{left}} \int_{t_{left}}^{t_{right}} |Wf\left(u, 2^k\right)|^2 \,\mathrm{d}u\right)^{1/2},\tag{15}$$

where $t_{left} := \max(t - \tau, t_0)$ and $t_{right} := \min(t + \tau, t_N)$. The factor $\frac{2\tau}{t_{right} - t_{left}}$ is optional to counteract border effects in the time interval.

Analogously, the WSD in (14) can also be modified to reduce border effects. So, it can be rewritten as

$$\mathcal{WSD}_{\tau,r}(t,k) := \left(\frac{2r}{k_{right} - k_{left}} \int_{k_{left}}^{k_{right}} \left(\frac{\mathcal{WS}_{\tau}(t,\kappa) - \mathcal{WS}_{\tau}'(t,\kappa)}{\mathcal{WS}_{\tau}(t,\kappa)}\right)^2 \,\mathrm{d}\kappa\right)^{1/2},\qquad(16)$$

where $k_{left} := \max(k - r, k_{min}), k_{right} := \min(k + r, k_{max})$, and k_{min}, k_{max} are those considered in Remark 2.4.

2.4 Wavelet squared coherence

The WSD can serve as an alternative or complement to the *wavelet squared coherence* (WSC) (see [14, 16]), which represents a widely employed measure in the wavelet framework. Both tools are very helpful to assess the degree of association between two time series, but they concentrate on slightly different aspects of the relationship. According to [16], the WSC between two time series f(t) and g(t) is defined by

$$\mathcal{WSC}(u,s) = \frac{|S\left(s^{-1}W_{fg}(u,s)\right)|^2}{S\left(s^{-1}|Wf(u,s)|^2\right)S\left(s^{-1}|Wg(u,s)|^2\right)},\tag{17}$$

where $W_{fg}(u,s) = Wf(u,s)Wg^*(u,s)$ is the cross-wavelet spectrum and S is a smoothing operator in both time and frequency. This smoothing operator is the only parameter that can be changed and, in this paper, we will always use a gaussian filter (following [16]). The WSC (17) ranges from 0 (no correlation) to 1 (perfect correlation) and is analogous to the squared correlation coefficient in linear regression. This concept is particularly useful for determining the regions in the time-frequency domain where two time series have a significant co-movement or interdependence, reflecting the local linear correlation between the series.

In contrast, the WSD compares the behavior of two time series through their respective scalograms for different windows in time and scale, thus allowing to ascertain the particular scales and time intervals in which both time series exhibit a similar pattern, comparing their scalograms and determining if they give the same weight to the different scales. Thus, the WSD is able to detect features that go unnoticed by the WSC (see Figure 1). It is worth highlighting that the great flexibility of the WSD arises from the possibility of shifting the length of time and scale windows. Nevertheless, as it is stated in Remark 2.1, it is recommended that the two series have the same unit of measurement to avoid spurious results due to normalization.

2.5 Practical aspects of the WSD

Despite the modifications made in the definition of the WSD in order to mitigate border effects in time and log-scale intervals, these effects do not disappear completely and, therefore, large scales and times close to the boundaries are still affected. An effective method for further minimizing border effects is to perform a Monte Carlo simulation, computing the WSDs of a large number of pairs of random time series with the same length as the original signals f, g. Next, at each (t, k) the original WSD is divided by the mean at (t, k) of these WSDs, thus obtaining a modified WSD in which values greater than 1 denote significant differences between the patterns of f and g.

To facilitate comparison with the WSC, for which high values indicate a high degree of similarity, it is worthy to plot the $\log_2(\mathcal{WSD}^{-1})$ rather than \mathcal{WSD} . In this way, on the one hand, we have a direct relationship between the value of $\log_2(\mathcal{WSD}^{-1})$ and the level of similarity between the patterns of the two time series. On the other hand, the logarithmic scale enhances the plot clarity. Moreover, if, as stated above, the original WSD is divided by the mean of a Monte Carlo simulation, then negative values of $\log_2(\mathcal{WSD}^{-1})$ stand for low similarity and positive values stand for high similarity.

Additionally, it is worth mentioning that the WSD is defined in relative terms, i.e. we can not compare the scalogram differences of two distinct pairs of time series that use different units of measure or are not normalized as discussed in Remark 2.1. So, it is important to conduct a statistical significance analysis, e.g. by using Monte Carlo techniques (taking advantage of the previous simulation).

For example, Figure 1 shows a comparison between the graphs corresponding to the WSC[‡] and the logarithm of the inverse of the commutative WSD for the same pair of time series by employing the Morlet wavelet[§]. These time series have 1 500 values and are generated according to the following processes

$$f(t) := N(0,1) + \sin(t/10)$$

$$g(t) := N(0,1) + \sin(t/10) + \chi_{[500,1000]}(t)\sin(t/2),$$
(18)

for t = 1, ..., 1500, where $\chi_{[500,1000]}$ is the characteristic function of the interval [500, 1000]. The time series f is the sum of two components. The first component, N(0, 1), represents a random number generated from a normal distribution with a mean of 0 and a variance of 1, while the second component, $\sin(t/10)$, is a sine with a period of $20\pi \approx 62.83$. In turn, the time series g has an additional component, $\sin(t/2)$, a sine with a period of $4\pi \approx 12.57$, which only applies if $t \in [500, 1000]$. For computation of the WSD we have employed a window with time radius 50 and log-scale radius 5/12. The WSC and the logarithm of the inverse of the WSD are displayed by using contour plots as they involve three dimensions: scale, time and level of association between the two time series considered. Instead of the scale, we have

[‡]The computations of the WSC in this study have been performed by using a MatLab program written by C. Torrence and G. P. Compo available at http://paos.colorado.edu/research/wavelets/.

[§]The Morlet wavelet has become one of the most popular wavelet families because of its optimal joint frequency concentration (see [14]). Moreover, the Morlet wavelet simplifies the interpretation of the wavelet analysis as it implies a very simple inverse relationship between scale and frequency (see [12, 19]).



Figure 1: Contour plots of the wavelet squared coherence (WSC, left) and the logarithm of the inverse of the commutative windowed scalogram difference (WSD, right) for the pair of time series given by (18). The color scale on the right of the graphs shows the level of comovement (WSC) or similarity (WSD). The lighter the color the higher the comovement or similarity of the two time series considered. The black contour line designates areas in which the WSC or WSD are significant at the 5% level, which have been estimated using Monte Carlo simulation. The thin black line denotes the cone of influence. The rectangular areas in the margins of the WSD where there are also border effects are shown by a thin gray line. The WSD has been calculated by using a window with time radius 50 and log-scale radius 5/12. Only the WSD detects the lack of connection between both time series at scale 4π in the interval $t \in [500, 1000]$ (white dashed line).

represented the corresponding Fourier period which, in the case of the Morlet wavelet used here, is calculated by multiplying the scale by the Fourier factor 1.033044 aprox. (see [15]). Fourier period and time are represented on the vertical and horizontal axes, respectively. The association between the two time series is depicted in both wavelet tools by a gray scale. The extent of linkage is interpreted in terms of the intensity of the gray color, so that the shades of gray range from white (high level of similarity) to black (low level of similarity). The thick black continuous line isolates regions where the association is statistically significant at the 5% level. The statistical significance level of the WSC and WSD is estimated through Monte Carlo simulation methods. Specifically, 1000 pairs of time series of the same length as the original time series and normally distributed with the same variance as the series under examination have been generated. Then, the WSD and WSC for each pair of random time series have been estimated and the corresponding results have been used for determining the significance level. Moreover, in both tools the thin black line indicates the cone of influence, below which edge effects become important. Hence, values outside the cone of influence should be interpreted carefully (see [15]). In addition, the rectangular areas in the margins of the WSD graph (thin gray lines) where there exist edge effects resulting from the use of windows with a certain size are also displayed.

The comparison of estimated WSC and WSD in Figure 1 reveals that both tools identify a similar pattern in the two time series under analysis at scale 20π (black dotted line) induced by the common term $\sin(t/10)$ during the full sample. However, only the WSD detects the lack of connection between both time series at scale 4π in the interval $t \in [500, 1000]$ (white dashed line) due to the presence of the term $\sin(t/2)$ in the second time series but not in the first one. Therefore, this figure clearly shows that the WSD enables one to detect dissimilarities between two series that remain unnoticed by the WSC.

Figures 2, 3 and 4 illustrate how different choices of the window size affect to the estimated WSD between the two time series considered. Visual inspection of these figures shows that



Figure 2: Logarithm of the inverse of the commutative WSD of the two time series reported in (18) for different choices of the window size (time radius \times log-scale radius). In particular, a time radius of 10, 25, 50 and 100 data points is used in each of the graphs, while a constant log-scale radius of 2/12 is utilized in the different graphs.



Figure 3: Logarithm of the inverse of the commutative WSD of the two time series reported in (18) for different choices of the window size (time radius \times log-scale radius). In particular, a time radius of 10, 25, 50 and 100 data points is used in each of the graphs, while a constant log-scale radius of 5/12 is utilized in the different graphs.



Figure 4: Logarithm of the inverse of the commutative WSD of the two time series reported in (18) for different choices of the window size (time radius \times log-scale radius). In particular, a time radius of 10, 25, 50 and 100 data points is used in each of the graphs, while a constant log-scale radius of 10/12 is utilized in the different graphs.

there is a similar general pattern in all of then, with a white strip at scale 20π covering the whole time range, and a black band at scale 4π through the time interval [500, 1000]. Obviously, for smaller windows more details but also more noise can be observed.

The optimum size of the window depends on the scales in which we are most interested but, generally, for a time series with N data points, a time radius between N/50 and N/10, and a log-scale radius between $0.2 \log_2 (N/6)$ and $\log_2 (N/6)$ (using the Morlet wavelet) seems to be a reasonable choice. However, a suitable parameters setting depends strongly on the characteristics that we want to study in the series and the level of detail of that study.

3 An application: integration of European government bond markets

The data for this application consist of yields on 10-year government bonds of five euro area peripheral countries, namely Greece, Ireland, Italy, Portugal and Spain, also called GIIPS countries, and Germany. The sample ranges from January 1999 to April 2013, thus covering the turbulent period which includes the recent global financial and Eurozone debt crises. Following the usual practice in the literature, Germany is taken as the benchmark country because German 10-year government bonds are typically seen as a safe haven. In line with, among others, [26] and [27], weekly data (sampled on Wednesdays) are used. Weekly changes in 10-year government bond yields are calculated as the first difference of 10-year bond yields between two consecutive observations. Bond data have been collected from Thomson Financial Datastream.

Figure 5 displays the graphs of the logarithm of the inverse of the commutative WSD between changes in 10-year government bond yields of each of the five EMU peripheral



Figure 5: Logarithm of the inverse of the commutative WSD between changes in yields on 10-year government bonds of each of GIIPS countries and Germany. The WSD has been calculated by using a window of time radius 25 and log-scale radius 4/12. The color scale on the right of the graphs shows the level of similarity. The lighter the color the higher the similarity of changes in 10-year bond yields of the respective GIIPS country and Germany. The thick black line designates areas in which the WSD is significant at the 5% level and the thin black line represents the cone of influence.

countries and those of Germany by using a window of time radius 25 (approximately half a year) and log-scale radius 4/12. The WSDs are visualized by contour plots, with Fourier period and time being represented in the vertical and horizontal axes, respectively. In order to facilitate the interpretation, the Fourier period is converted into time units (years) and it ranges from the highest scale of 0.0625 years, that is, one week (top of the plot) to the lowest scale of 4 years (bottom of the plot). The level of similarity between bond markets of each GIIPS country and Germany is indicated by color coding, which ranges from black (low similarity) to white (high similarity). The regions encircled by a thick black line represent areas where the WSD is significant at the 5% level. Monte Carlo methods are used to assess the statistical significance of the WSD. Specifically, the significance level is determined with 1000 pairs of random time series of the same length and with the same variance as the original series. The cone of influence, below which edge effects might distort the results of the WSD, is designated by a thin black line.

The WSDs show that the degree of bond market integration varies considerably over time and across scales. In particular, a virtually identical pattern is observed for changes in 10year bond yields of GIIPS countries and Germany, mainly for scales of less than two years, since the introduction of the euro until the intensification of the global financial crisis in the autumn of 2008. This finding implies an almost perfect bond market integration in the years



Figure 6: WSC between changes in yields on 10-year government bonds of each of GIIPS countries and Germany. The color scale on the right of the graphs shows the level of comovement. The lighter the color the higher the comovement of changes in 10-year bond yields of the respective GIIPS country and Germany. The thick black line designates areas in which the WSC is significant at the 5% level and the thin black line represents the cone of influence.

following the launch of the euro as a result of the removal of exchange rate risk, the nominal convergence of economic fundamentals and harmonization of fiscal and regulatory frameworks within the European Monetary Union (EMU). In contrast, the time interval after the collapse of the U.S. bank Lehman Brothers in September 2008 is characterized by the decoupling of 10-year government bond yields of GIIPS countries relative to those of Germany for virtually all scales. This divergence may be explained by the pessimistic economic outlook in euro area peripheral countries, the increased risk aversion of investors and the safe haven status of German debt after September 2008. These factors have led to an unprecedented rise in sovereign bond yields of European peripheral countries relative to German bond yields. The breakdown in bond market integration is first detected at lower scales and then it is gradually extended to all other scales, although there are small differences among countries. As can be seen, the highest level of bond market integration of GIIPS countries and Germany is consistently found at the scale of around one year, while for horizons of more than two years the extent of linkage is substantially weaker.

The estimated WSC are reported in Figure 6 for comparative purposes and their results are broadly consistent with those of the WSD. Indeed, a statistically significant coherence between changes in yields on 10-year government bonds of GIIPS countries and those of Germany is detected since the euro's introduction in 1999 until the worsening of the global financial crisis in autumn of 2008. Just as the WSD, the highest level of coherence is observed



Figure 7: Up: Logarithm of the inverse of the commutative WSD between changes in yields on 10-year government bonds of two European core countries (Finland and the Netherlands) and Germany. The WSD has been calculated by using a window of time radius 25 and log-scale radius 4/12. Down: WSC between changes in yields on 10-year government bonds of the same two European core countries and Germany. The thick black line designates areas in which the WSD or WSC is significant at the 5% level and the thin black line represents the cone of influence.

at the scale of around one year. Furthermore, the WSC also shows an abrupt rupture of the nearly-perfect integration between sovereign bond yields from late 2008.

In order to check whether the above findings on integration of bond markets of GIIPS countries and Germany hold for other European countries, the WSD is also calculated for some euro area core countries. Thus, Figure 7 reports the logarithm of the inverse of the commutative WSD between changes in 10-year sovereign bond yields of Finland and the Netherlands relative Germany. It is worth noting that the sharp decline in bond market integration since the aggravation of the financial crisis in late 2008 above documented for peripheral countries is not found for these two European core countries. Specifically, a certain reduction in the level of bond market integration is observed for these economies during the hardest stage of the European debt crisis. However, the values of the $\log_2 (WSD^{-1})$ are not only positive but greater than 1, suggesting a high level of comovement along all the sample. Therefore, this evidence seems to indicate that the fragmentation of government bond markets during the recent financial crisis period has primarily affected European peripheral countries.

Finally, we compare the results of the WSD and WSC analysis regarding the integration of sovereign debt markets of several European countries with those of a reference approach in the financial economics literature such as the DCC (Dynamic Conditional Correlation)-GARCH model developed by [28]. The DCC-GARCH model has become the most popular GARCH-type model due to its parsimony and flexibility in modelling the conditional correlation dynamics between asset returns. Figures 8 and 9 display the dynamic conditional correlation estimates between changes in 10-year government bond yields of each of the EMU countries



Figure 8: Dynamic Conditional Correlation between changes in yields on 10-year government bonds of each of GIIPS countries and Germany estimated with the DCC-GARCH model.

under consideration and those of Germany. The findings of the DCC-GARCH approach are largely consistent with those of the WSD and WSC. Thus, a significant decoupling of 10-year government bond yields of GIIPS countries relative to those of Germany is observed from the bankruptcy of Lehman Brothers in September 2008 and a reduction in the degree of integration of sovereign debt markets of Finland and the Netherlands and that of Germany is also found during the most acute phase of the European debt crisis. However, the main disadvantage of the DCC-GARCH analysis compared with the WSD and WSC is that the DCC-GARCH approach is not capable of distinguishing between different scales.

4 Concluding remarks

The main contribution of this paper is the introduction of the windowed scalogram difference (WSD), a new wavelet-based tool which is especially useful to assess if two time series follow a similar pattern over time and across scales. The WSD shares with the widely accepted wavelet squared coherence (WSC) the common objective of providing a measure of the level of association between two time series in the time-frequency domain. However, the WSD has two major advantages over the WSC. First, the WSD shows a greater flexibility as it allows one to change the size of the window considered depending on which time intervals and scales are more interesting. Second, the WSD can reveal certain aspects of the relationship between the time series that the WSC is not able to capture (see Figure 1). The only limitation of the WSD is the recommendation that the two series under analysis should use the same unit of measure or are normalized such as discussed in Remark 2.1.



Figure 9: Dynamic Conditional Correlation between changes in yields on 10-year government bonds of two European core countries (Finland and the Netherlands) and Germany estimated with the DCC-GARCH model.

Even though the primary purpose of this study is to introduce the concept of WSD, the practical applicability of this tool is demonstrated in the context of the integration of sovereign debt markets in a number of countries in the euro area. For both tools, WSC and WSD, the results show that the extent of bond market integration of eurozone peripheral countries and Germany has undergone a dramatic reversal during the latest financial and debt crises.

To sum up, it can be concluded that the WSD appears as an interesting alternative tool to the WSC, and hence it can be used to compare different time series in future research.

References

- R. F. Engle, C. W. J. Granger. Cointegration and error correction: representation, estimation and testing. *Econometrica* 35 (1987), 251–276.
- [2] C. W. J. Granger. Investigating causal relations by econometric models and crossspectral methods. *Econometrica* 37 (1969), 424–438.
- [3] C. A. Sims. Macroeconomics and reality. *Econometrica* 48 (1980), 1–48.
- [4] R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50 (1982), 987–1007.
- [5] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31 (1986), 307–327.
- [6] C. Bayer, C. Hanck. Combining non-cointegration tests. J. Time Series Anal. 34 (2013), 83–95.
- [7] N. Sim, H. Zhou. Oil prices, US stock return, and the dependence between their quantiles. J. Bank. Financ. 55 (2015), 1–8.
- [8] G. Li, Y. Li, C. L. Tsai. Quantile correlations and quantile autoregressive modeling. J. Amer. Statist. Assoc. 110 (2015), 246–261.
- [9] Y. Shin, B. Yu, M. Greenwood-Nimmo. Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework. Chapter 9 of the book "Festschrift in honor of Peter Schmidt", 281–314. Springer New York (2014).

- [10] J. S. Cho, T. H. Kim, Y. Shin. Quantile cointegration in the autoregressive distributedlag modeling framework. J. Econometrics 188 (2015), 281–300.
- [11] F. In, S. Kim. An introduction to wavelet theory in finance: a wavelet multiscale approach. World Scientific (2013).
- [12] L. Aguiar-Conraria, M. J. Soares. The continuous wavelet transform: moving beyond uni- and bivariate analysis. J. Econ. Surveys. 28 (2014), 344–375.
- [13] P. Goupillaud, A. Grossman, J. Morlet. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration* 23 (1984), 85–102.
- [14] A. Grinsted, J. C. Moore, S. Jevrejeva. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Proc. Geoph.* 11 (2004), 561–566.
- [15] C. Torrence, G. P. Compo. A practical guide to wavelet analysis. B. Am. Meteorol. Soc. 79 (1998), 61–78.
- [16] C. Torrence, P. J. Webster. Interdecadal changes in the ENSO-monsoon system. J. Climate 12 (1999), 2679–2690.
- [17] G. Hariharan. An efficient wavelet based approximation method to water quality assessment model in a uniform channel. Ain Shams Eng. J. 5 (2014), 525–532.
- [18] P. Rajeev, K. K. Wijesundara. Energy-based damage index for concentrically braced steel structure using continuous wavelet transform. J. Constr. Steel Res. 103 (2014), 241–250.
- [19] B. Cazelles, M. Chavez, G. C. de Magny, J. F. Guégan, S. Hales. Time-dependent spectral analysis of epidemiological time-series with wavelets. J. R. Soc. Interface 4 (2007), 625–636.
- [20] C. H. Lin, Y. C. Du, T. Chen. Adaptive wavelet network for multiple cardiac arrhythmias recognition. *Expert Syst. Appl.* 34 (2008), 2601–2611.
- [21] R. Kronland-Martinet, J. Morlet, A. Grossmann. Analysis of sound patterns through wavelet transforms. Int. J. Pattern Recognit. Artif. Intell. 1 (1987), 273–302.
- [22] A. A. Petrosian, F. G. Meyer. Wavelets in signal and image analysis: From theory to practice. Springer Science+Business Media (2013).
- [23] P. M. Crowley. A guide to wavelets for economists. Journal of Economic Surveys 21 (2007), 208–267.
- [24] R. Benítez, V. J. Bolós, M. E. Ramírez. A wavelet-based tool for studying nonperiodicity. Comput. Math. Appl. 60 (2010), 634–641.
- [25] S. Mallat. A Wavelet Tour of Signal Processing: The Sparse Way. Academic Press (2008).
- [26] P. Abad, H. Chuliá, M. Gómez-Puig. EMU and European bond market integration. J. Bank. Financ. 34 (2010), 2851–2860.
- [27] L. Pozzi, G. Wolswijk. The time-varying integration of euro area government bond markets. Eur. Econ. Rev. 56 (2012), 36–53.
- [28] R. F. Engle. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econ. Stat. 20 (2002), 339– 350.