Universidade do Minho

Documentos de Trabalho

Working Paper Series

"Decomposing time-frequency economic relationships"

Luís Aguiar-Conraria

Maria Joana Soares

Nuno Azevedo

NIPE WP 16/ 2007

NÚCLEO DE INVESTIGAÇÃO EM POLÍTICAS ECONÓMICAS

UNIVERSIDADE DO MINHO

“Decomposing time-frequency economic relationships”

Luís Aguiar-Conraria
Maria Joana Soares
Nuno Azevedo

NIPE[*] WP 16/ 2007

URL:
http://www.eeg.uminho.pt/economia/nipe

[*] NIPE – *Núcleo de Investigação em Políticas Económicas* – is supported by the Portuguese Foundation for Science and Technology through the *Programa Operacional Ciência, Teconologia e Inovação* (POCI 2010) of the *Quadro Comunitário de Apoio III*, which is financed by FEDER and Portuguese funds.

Decomposing time-frequency economic relationships

Luís Aguiar-Conraria* Maria Joana Soares† Nuno Azevedo‡

July 27, 2007

Abstract

Many economic processes are a combination of components operating on different frequencies. Several questions about the data are connected to the understanding of the behavior at different frequencies. Fourier analysis allow us to study the cyclical nature of a time-series in the frequency domain. However, under the Fourier transform, the time information of a time series is completely lost. Some authors have proposed Wavelet analysis, which performs the estimation of the spectral characteristics of a time-series as a function of time, as an alternative to the Fourier transform.

In this paper, we suggest two tools that generalize wavelet methods: cross wavelets and wavelet coherency. With these tools, we are able to use wavelet analysis to directly study the interactions (such as covariance, correlation and causality) between two time-series at different frequencies and how they evolve over time.

Keywords: Business cycles, time-frequency analysis, non-stationary time series, wavelets, cross wavelets, wavelet coherency.

JEL Classification: C10, E32.

*Departamento de Economia, Universidade do Minho, E-mail address: lfaguiar@eeg.uminho.pt
†Departamento de Matemática, Universidade de Minho, E-mail address: jsoares@math.uminho.pt
‡Faculdade de Engenharia, Universidade do Porto, E-mail address: ncazevedo@gmail.com

1 Introduction

Many economic processes are a combination of components operating on different frequencies. Several questions about the data are connected to the understanding of the behavior at different frequencies. Sometimes the same economic agent may be simoultaneasly operating at different horizons. For example, the Central banks may have different objectives in the short and long run, and may be operating simultaneously at different timescales (e.g. see Ramsey and Lampart 1998a and 1998b).

Fourier analysis allow us to study the cyclical nature of a time-series in the frequency domain. However, under the Fourier transform, the time information of a time series is completely lost. This does not mean that they are not useful. For example, spectral techniques can be used to identify seasonal components, such as a Christmas effect (see Wen 2002). But these classical techniques can only be used for time-series in which the statistical properties do not vary with time, i.e. are stationary. Unfortunately, macroeconomic time-series are noisy, complex and, tipically strongly non-stationary.

To overcome the problems of analysing non-stationary Gabor (1946) have suggested the use of the Short Time Fourier Transform (STFT). The basic idea is to break a time series into smaller sub-samples and apply Fourier transform to each sub-sample. However, as Raihan et al. (2005) pointed out, once it is selected the length of the window is fixed. However a longer window implies the loss of information along the time dimension, and a shorter window implies the loss of information along the frequency dimension.

As an alternative to STFT, wavelet analysis has been proposed. Wavelet analysis performs the estimation of the spectral characteristics of a a time-series as a function of time.This approach reveals how the different the periodic components of the time-series change over time.

One major advantage afforded by the wavelet transform is the ability to perform natural local analysis of a time series in the sense that the length of wavelets varies endogenously.

It stretches into a long wavelet function to measure the low frequency movements; and it compresses into a short wavelet function to measure the high frequency movements. In order to capture abrupt changes, for example, one would like to have very short functions (narrow windows). At the same time, in order to isolate slow and persistent movements, one would like to have very long functions (wide windows). This is exactly what can be achieved with the wavelet transform.

While the Fourier transform breaks down a time series into constituent sinusoids of different frequencies and infinite duration in time, the wavelet transform expands the time series into shifted and scaled versions of a function – the so-called *mother-wavelet* – that has limited spectral band and limited duration in time. We know from the Heisenberg uncertainty principle that there is always a trade-off between localization in time and localization in frequency; in particular, we cannot ask for a function to be, simultaneously, band and time limited. However, a mother wavelet can be chosen with a fast decay in time and frequency which, for all pratical purposes, corresponds to an effective band and time limiting; see Daubechies (1992).

As a coherent mathematical body, wavelet theory was born in the mid-1980s (Grossmann and Morlet 1984, Goupillaudand et al 1984). After 1990, the literature rapidly expanded and wavelet analysis is now used extensively in physics, geo-physics, astronomy, epidimiology, signal processing, oceanography, etc. Interestingly, and in spite of all its potential advantadges, this technique is very rarely used in Economics. The pioneering work of Ramsey and Lampart (1998a and 1998b) and Ramsey (1999), as noted by Crowley (2007) is unknown by most of the economists, who seem fixated on traditional econometric methods because, overlooking the potential for using wavelets in economic data. Among the notable exceptions to this rule, one can point at Wen e os outros. For a more thorough (and excellent) review see ??????.

Probably, one of the reasons why wavelets are not more popular in the economics literature is related to the dificulty to simultaneosly analyse two (or more) time-series simultaneously. In Economics, these techiques have either been applied to a single time-series (e.g. Raihan et al. 2005) or used to individually analise two time-series (one each time), whose decompositions are then studied using traditional time domain methods, such us correlation analysis or Granger causality (see Ramsey and Lampart, 1998a and 1998b).

In this paper, we present two tools, Cross Wavelet Transform and Cross Wavelet Coherence, proposed by Hudgins et al (1993) and Torrence and Compo (1998) that may help to overcome this problem. Cross wavelets and wavelet coherency generalize wavelet methods, allowing the analysis of time-frequency dependencies between the two time-series. With these tools, we are able to use wavelet analysis to directly study the interactions (such as covariance, correlation and causality) between two time-series at different frequencies and how they evolve over time.

This paper proceeds as follows. Section 2 rigorously introduces wavelets. It starts by discussing the Continuous Wavelet Transform (CWT), its localization properties and discusses in some detail one of the most popular wavelets, the Morlet wavelet. Section 3 describes the Cross Wavelet Transform (XWT) and the Wavelet Coherence (WTC) and discusses how to assess their statistical significance. Section 4 applies CWT, XWT and WTC to macroeconomic data and discusses its insights. Section 5 concludes.

2 Wavelets: The dynamical decomposition of time

2.1 The wavelet

We start by introducing some mathematical notation. In what follows, $L^2(\mathcal{R})$ denotes the set of square integrable functions, i.e. the set of functions defined on the real line such that

$$\int_{-\infty}^{\infty} |x(t)|^2 \, dt < \infty. \tag{1}$$

Since the above quantity is usually referred to as the energy of the function x, this space is also known as the space of functions with finite energy. As it is well known, one can define in $L^2(\mathcal{R})$ an inner product

$$\langle x, y \rangle := \int_{-\infty}^{\infty} x(t) \, y^*(t) \, dt \tag{2}$$

and an associated norm $\|x\| := \langle x, x \rangle^{\frac{1}{2}}$. Here, and throughout the paper, the asterisk superscript will be used to denote complex conjugation and the symbol := means “by definition”.

Given a function $x(t) \in L^2(\mathcal{R})$, we will denote by $X(f)$ the Fourier transform of $x(t)$:

$$X(f) := \int_{-\infty}^{\infty} x(t) \, e^{-i2\pi f t} dt. \tag{3}$$

We recall the well-known Parseval relation, valid for all $x(t), y(t) \in L^2(\mathcal{R})$:

$$\langle x_t, y(t) \rangle = \langle X(f), Y(f) \rangle, \tag{4}$$

from which the Plancherel identity (which sates that the energy of a function is preserved by the Fourier transform) immediately follows:

$$\|x(t)\|^2 = \|X(f)\|^2; \tag{5}$$

see, for example, Körner (1988).

The minimum requirements imposed on a function $\psi(t)$ to qualify for being a *mother (admissible or analyzing) wavelet* are that $\psi \in L^2(\mathcal{R})$ and also fulfills a technical condition, known as the *admissibility condition*, which reads as follows:

$$0 < C_\psi := \int_{-\infty}^{\infty} \frac{|\Psi(f)|}{|f|} df < \infty. \tag{6}$$

The wavelet ψ is usually normalized so that it has unit energy, i.e. $\|\psi\|^2 = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt = 1$.

The square integrability of ψ is a very mild decay condition; the wavelets used in practice have much faster decay; typical behavior will be exponential decay ($|\psi(t)| \leq Me^{C|t|}$,for some constants C and M) or even compact support.

For functions with sufficient decay[1] it turns out that the admissibility condition (6) is equivalent to requiring

$$\Psi(0) = \int_{-\infty}^{\infty} \psi(t) \, dt = 0. \tag{7}$$

This means that the function ψ has to wiggle up and down the t-axis, i.e. it must behave like a wave; this, together with the decaying property, justifies the choice of the term wavelet (originally, in French, *ondelette*) to designate ψ.

2.2 The continuous wavelet transform

Starting with a mother wavelet ψ, a family $\psi_{s,\tau}$ of "wavelet daughters" can be obtained by simply scaling ψ by s and translating it by τ

$$\psi_{s,\tau}(t) := \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right) \quad ,s, \tau \in \mathcal{R}, s \neq 0. \tag{8}$$

The parameter s is a scaling or dilation factor that controls the length of the wavelet (the

[1] The exact minimum decay requirements are that $\psi \in L^2(\mathcal{R})$ also satisfies $\int_{-\infty}^{\infty} (1+|t|)^\alpha |\psi(t)| \, dt < \infty$ for some $\alpha > 0$; see e.g. Daubechies (1992). p. 25.

factor $1/\sqrt{|s|}$ being introduced to guarantee preservation of the unit energy, $\left\|\psi_{s,\tau}\right\| = 1$) and τ is a location parameter that indicates where the wavelet is centered. Scaling a wavelet simply means stretching it(if $|s| > 1$), or compressing it (if $|s| < 1$).[2]

Given a function $x(t) \in L^2(\mathcal{R})$ (a time series), its continuous wavelet transform (CWT), with respect to the wavelet ψ, is a function $W_x(s,\tau)$ obtained by projecting $x(t)$, in the L^2 sense, onto the over-complete family $\left\{\psi_{s,\tau}\right\}$:

$$W_x(s,\tau) = \left\langle x, \psi_{s,\tau} \right\rangle = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|s|}} \psi^* \left(\frac{t-\tau}{s} \right) dt. \tag{9}$$

The importance of the admissibility condition (6) comes from the fact that it guarantees that it is possible to recover $x(t)$ from its wavelet transform; see e.g. Daubechies (1992). p. 25:

$$x(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \left[W_x(s,\tau)\, \psi_{s,\tau}(t)\, d\tau \right] \frac{ds}{s^2}. \tag{10}$$

Since we can go from $x(t)$ to its wavelet transform, and from the wavelet transform back to $x(t)$, we can conclude that both are representations of the same mathematical entity. They just present information in a different manner, allowing us to gain insights that would, otherwise, remain hidden. It is also important to observe that the energy of $x(t)$ is preserved by the wavelet transform, in the sense that

$$\|x\|^2 = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \left| W_x(s,\tau) \right|^2 d\tau \right] \frac{ds}{s^2} \tag{11}$$

and that a Parseval type identity also holds

$$\langle x, y \rangle = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \left[W_x(s,\tau)\, W_y^*(s,\tau)\, d\tau \right] \frac{ds}{s^2} \tag{12}$$

for $x, y \in L^2(\mathcal{R})$.

Because the wavelet function $\psi(t)$ may, in general, be complex, the wavelet transform

[2] Note that for negative s, the function is also reflected.

W_x may also be complex. The transform can then be divided into its real part, $\mathcal{R}\{W_x\}$, and imaginary part, $\mathcal{I}\{W_x\}$, or in its amplitude, $|W_x|$, and phase, $\tan^{-1}\left(\frac{\mathcal{I}\{W_x\}}{\mathcal{R}\{W_x\}}\right)$. For real-valued wavelet functions the imaginary part is zero and the phase is undefined. Therefore, in order to separate the phase and amplitude information of a time series it is important to make use of complex wavelets. In particular, it is convenient to choose $\psi(t)$ to be *progressive* or *analytic*, i.e. to be such that $\Psi(f) = 0$ for $f < 0$; in this case, if $x(t)$ is real, a variant of the reconstruction formula, in which the parameter s can be restricted to positive values only, is possible:

$$x(t) = \frac{2}{C_\psi} \int_0^\infty \left[\int_{-\infty}^\infty \mathcal{R}\left(W_x(s,\tau)\, \psi_{s,\tau}(t) \right) d\tau \right] \frac{ds}{s^2} \tag{13}$$

and one also has

$$\|x\|^2 = \frac{2}{C_\psi} \int_0^\infty \left[\int_{-\infty}^\infty |W_x(s,\tau)|^2 \, d\tau \right] \frac{ds}{s^2} \tag{14}$$

and

$$\langle x, y \rangle = \frac{2}{C_\psi} \int_0^\infty \left[W_x(s,\tau)\, W_y^*(s,\tau)\, d\tau \right] \frac{ds}{s^2}; \tag{15}$$

see Daubechies (1992), pp. 27-28, Kaiser 1994, pp. 70-73 or Mallat (1998), pp.82-83 for more details about analytic wavelets. Throughout the rest of the paper, since, in the practical applications, we will use an analytic wavelet, we always assume that the scaling parameter s takes positive values only.

In view of the energy preservation formula (14), and in analogy with the terminology used in the Fourier case, the function $|W_x(s,\tau)|^2$ is usually referred to as the wavelet power spectrum (sometimes also called the scalogram, see Flandrin 1988).

2.3 Localization properties

Let the wavelet ψ be normalized so that $\|\psi\| = 1$ and define its center μ_t by

$$\mu_t = \int_{-\infty}^\infty t\, |\psi(t)|^2 \, dt. \tag{16}$$

In other words, the center of the wavelet is simply the mean of the probability distribution obtained from $|\psi(t)|^2$. As a measure of concentration of ψ around its center one usually takes the variance σ_t:

$$\sigma_t = \left\{ \int_{-\infty}^{\infty} (t-\mu_t)^2 |\psi(t)|^2 \, dt \right\}^{\frac{1}{2}}. \tag{17}$$

In a total similar manner, one can also define the center μ_f and variance σ_f of the Fourier transform $\Psi(f)$ of ψ.

The interval $[\mu_t - \sigma_t, \mu_t + \sigma_t]$ is the set where ψ attains its "most significant" values whilst the interval $\left[\mu_f - \sigma_f, \mu_f + \sigma_f\right]$ plays the same role for $\Psi(f)$ of ψ. The rectangle $[\mu_t - \sigma_t, \mu_t + \sigma_t] \times \left[\mu_f - \sigma_f, \mu_f + \sigma_f\right]$ in the $(t, f)-$plane is called the Heisenber box or window in the time-frequency plane. We then say that ψ is localized around the point $\left(\mu_t, \mu_f\right)$ of the time-frequency plane with uncertainty given by $\sigma_t \sigma_f$.

In quantum mechanics, the uncertainty principle, first established by Werner Karl Heisenberg, gives a lower bound on the product of the standard deviations of position and momentum for a system, implying that it is impossible to have a particle that has an arbitrarily well-defined position and momentum simultaneously. The Heisenberg uncertainty principle establishes that the uncertainty is bounded from below by the quantity $1/4\pi$, i.e. one has

$$\sigma_t \sigma_f \geq \frac{1}{4\pi}. \tag{18}$$

It is also known that the equality in (18) is attained if and only if the function ψ is a (translated and modulated) gaussian: $\psi(t) = a e^{i\mu_f t} e^{-b(t-\mu_t)^2}$;see Messiah (1961).

It follows from the Parseval relation (4) that

$$\begin{aligned} W_x(s,\tau) &= \left\langle x(t), \psi_{s,\tau}(t) \right\rangle \\ &= \left\langle X(f), \Psi_{s,\tau}(f) \right\rangle \end{aligned} \tag{19}$$

where $X(f)$ and $\Psi_{s,\tau}(f)$ are the Fourier transforms of $x(t)$ and $\psi_{s,\tau}(t)$,respectively.

If the mother wavelet ψ is centered at μ_t and has variance σ_t and its wavelet transform $\Psi(f)$ is centered at μ_f with a variance σ_f, then one can easily show that the daughter wavelet $\psi_{\tau,s}$ s will be centered at $\tau + s\mu_t$ with variance $s\sigma_t$, whilst its Fourier transform $\Psi_{s,\tau}$ will have center $\frac{\mu_f}{s}$ and variance $\frac{\sigma_f}{s}$. Hence, (19) shows that the continuous wavelet transform $W_x(s,\tau)$ gives us local information within a time-frequency window

$$[\tau + s\mu_t - s\sigma_t, \tau + s\mu_t + s\sigma_t] \times \left[\frac{\mu_f}{s} - \frac{\sigma_f}{s}, \frac{\mu_f}{s} + \frac{\sigma_f}{s}\right] \tag{20}$$

In particular, if ψ is chosen so that $\mu_t = 0$ and $\mu_f = 1$,then the window associated with $\psi_{\tau,s}$ becomes

$$[\tau - s\sigma_t, \tau + s\sigma_t] \times \left[\frac{1}{s} - \frac{\sigma_f}{s}, \frac{1}{s} + \frac{\sigma_f}{s}\right] \tag{21}$$

In this case, the wavelet transform $\{\mathcal{W}_\psi f\}(s,\tau)$ will give us information on $x(t)$ for t near the instant $t = \tau$, with precision $s\sigma_t$, and information about $X(f)$ for frequency values near the frequency $f = \frac{1}{s}$, with precision $\frac{\sigma_f}{s}$. Therefore:

- small values of s correspond to information about $x(t)$ in a fine scale and about $X(f)$ in a broad scale,

- large values of s correspond to information in a broad scale about $x(t)$ and in a fine scale about $X(f)$,

- Although the area of the windows is constant at all scales, $A = 4\sigma_t\sigma_f$, their dimensions change according to the scale. The windows stretch for large values of s (broad scales s – low frequencies $f = 1/s$ and compress for small values of s (fine scale s – high frequencies $f = 1/s$).

2.4 The Morlet wavelet: optimal joint time-frequency concentration

There are several types of wavelet functions available with different characteristics, such as, Morlet, Mexican hat, Haar, Daubecies, etc; see, e.g. Daubechies (1992), Mallat (1998) or Meyer (1993). Since the wavelet coefficients $W_x(s, \tau)$ contain combined information on both the function $x(t)$ and the analyzing wavelet $\psi(t)$, the choice of the wavelet is an important aspect to be taken into account. This will depend maily on the particular application one has in mind. In this paper we choose a complex wavelet, as it yields a complex transform, with information on both the amplitude and phase, which is essential for the analysis we want to perform. One of the most popular wavelets used is the Morlet wavelet, first introduced in Goupillaudand (1984), which is defined as

$$\psi_\eta(t) = \mu^{-\frac{1}{4}} \left(e^{i\eta t} - e^{-\frac{\eta^2}{2}} \right) e^{-\frac{t^2}{2}}, \tag{22}$$

the term $e^{-\frac{\eta^2}{2}}$ being introduced to guarantee the fulfillment of the admissibility condition; however, for $\eta \geq 5$ this term becomes negligible and the simplified version

$$\psi_\eta(t) = \mu^{-\frac{1}{4}} e^{i\eta t} e^{-\frac{t^2}{2}} \tag{23}$$

of (22) is normally used (and still referred to as a Morlet wavelet). Our results in the next section, were obtained with the particular choice $\eta = 6$.

This wavelet has interesting characteristics. First of all, it is (almost) analytic. The Fourier transform of the "true" Morlet wavelet (22) is, in fact, supported in $(0, \infty)$, but that of (23) has some mass on $(-\infty, 0)$; for $\eta > 5$, this mass is, however, negligible, so, for all practical purposes, the wavelet can be considered as analytic; see Foufoula-Gergiou and Kumar (1993).

The wavelet (23) is centered at the point $\left(0, \frac{\eta}{2\pi}\right)$ of the time-frequency plane; hence, the for the particular choice $\eta = 6$, one has that the frequency center is

$$\mu_f = \frac{6}{2\pi} \approx 1 \tag{24}$$

and the relationship between the scale and frequency is simply

$$f = \frac{\mu_f}{s} \approx \frac{1}{s} \tag{25}$$

It is also very simple to verify that the time variance is $\sigma_t = 1/\sqrt{2}$ and the frequency variance is $\sigma_f = 1/\left(2\pi\sqrt{2}\right)$. Therefore, the uncertainty of the corresponding Heisenberg box attains the minimum possible value $\sigma_t \sigma_f$ and one can thus say that the Morlet wavelet has optimal joint time-frequency concentration.[3]

2.5 Transform of finite discrete data

If one is dealing with a discrete time series $\{x_n, n = 0, ..., N-1\}$ of N observations with a uniform time step δt, the integral in (9) has to be discretized and is, therefore, replaced by a summation over the N time steps; the CWT of the time series $\{x_n\}$ is thus given by

$$W_m^x(s) = \frac{\delta t}{\sqrt{s}} \sum_{n=0}^{N-1} x_n \psi^* \left((n-m) \frac{\delta t}{s} \right). \tag{26}$$

Although it is possible to calculate the wavelet transform using the above formula for each value of s and m, one can also identify the computation for all the values of m simultaneously as a simple convolution of two sequences; in this case, one can follow the standard procedure and calculate this convolution as a simple product in the Fourier domain, using the fast Fourier transform (FFT) algorithm to go forth and back from time to spectral domain; this is precisely the technique prescribed by Torrence and Compo (1998).[4]

[3] This could be antecipated by noting that ψ_η is a simple modulated Gaussian.

[4] A program code based on the above procedure is also available at the the site http://paos.colorado.edu/research/wavelets/.

As with other types of transforms, the CWT applied to a finite length time series inevitably suffers from border distortions; this is due to the fact that the values of the transform at the beginning and the end of the time series are always incorrectly computed, in the sense that they involve "missing" values of the series which are then artificially prescribed; the most common choices are zero padding – extension of the time series by zeros – or periodization. Since the "effective support" of the wavelet at scale s is proportional to s, these edge-effects also increase with s. The region in which the transform suffers from these edge effects is called the cone of influence (COI). In this area of the time-frequency plane the results are unreliable and have to be interpreted carefully. In this paper, the cone of influence is defined, following Torrence and Compo (1998), as the e-folding time of the wavelet at the scale s, that is, so that the wavelet power of a Dirac δ at the edges decreases by a factor of e^{-2}. In the case of the Morlet wavelet this is given by $\sqrt{2}s$, and in all the pictures is marked as a shadow in the wavelet plot.

3 Cross Wavelet Transform and Cross Wavelet Coherence

Probably, one of the reasons why wavelets are not more popular in the Economics literature is because it has been a difficult task to use wavelets to analyze two, or more, time series together. Torrence and Compo (1998) and Grinsted et al. (2004) showed how the Cross Wavelet Transform (XWT) and wavelet coherence can be used to quantify the relationships between two time series in the time-frequency space.

The XWT of two time series, $x = \{x_n\}$ and $y = \{y_n\}$, first introduced by Hudgins et al (1993) is simply defined as

$$W_n^{xy} = W_n^x W_n^{y*}, \tag{27}$$

where W_n^x and W_n^{yx} are the wavelet transforms of x and y, respectively. The cross wavelet power is given by $|W_n^{xy}|$.

While a wavelet power spectrum depicts the variance of a time series, with times of large variance showing large power, the cross–wavelet power of two time series depicts the covariance between these time series at each scale or frequency. Therefore, cross–wavelet power gives us a quantified indication of the similarity of power between two time series.

As in the Fourier spectral approaches, the wavelet coherence (WTC) can be defined as ratio of the the cross-spectrum to the product of the spectrum of each series, and can be thought of as the local correlation between two CWTs. Here, again, we follow Grinsted et al (2004) and define the wavelet coeherence between two time series as $x = \{x_n\}$ and $y = \{y_n\}$ as follows:

$$R_n^2(s) = \frac{\left|S\left(s^{-1}W_n^{xy}(s)\right)\right|^2}{S\left(s^{-1}\left|W_n^x\right|^2\right)S\left(s^{-1}\left|W_n^y\right|^2\right)}, \tag{28}$$

where S denotes a smoothing operator in both time and scale. The smoothing can be achieved by a convolution in time and scale. For the Morlet wavelet, a suitable smoothing operator is suggested by Torrence and Webster (1998) and also used by Grinsted et al

(2004). In this case, the time convolution is done with a Gaussian $e^{-t^2/\left(2s^2\right)}$, which is the absolute value of the wavelet function in each scale, and the scale convolution is performed by a rectangular window Π with a length of $0.6s$, (the factor 0.6 is the empirical scale decorrelation length for Morlet wavelet):

$$S_t\left(W, s\right) = W_n\left(s\right) c_1 e^{-\frac{t^2}{2s^2}} \tag{29}$$

and

$$S_s\left(W, n\right) = W_n\left(s\right) * c_2 \Pi\left(0.6s\right), \tag{30}$$

where c_1 and c_2 are normalizing constants and is the rectangle function; see Grinsted et al (2004). In practice both convolutions are done discretely and therefore the normalization coefficients are determined numerically.

Smoothing is a necessary step, because, without that step, coherence is identically one at all scales and times. A similar procedure is used in Fourier analysis.

3.1 Statistical significance

1. We have exact probability distributions for the CWT
2. We have exact probability distributions for the XWT
3. Torrence and Compo (1998) suggest that this can be estimated by conducting repeated Monte Carlo simulations of white or colored random noise.

The computed wavelet results are only considered statistical significant if they are above a given confidence level defined by the random noise simulations; more details can be found in [TC98], [TW99] and [JMG03].

A MatLab software package for performing and displaying the XWT and WTC, which also computes the levels of significance as described above, was developed by A. Grinsted, J. C. Moore, and S. Jevrejeva and can be found at http://www.pol.ac.uk/home/ research/waveletcoherence/. All the numerical results presented in this paper were obtained using the above package.

4 Interest Rates and the Economic Activity: a simple application

In this section we apply the methods analyzed earlier, to decompose the evolution of time-frequency relations between Economic variables.

The data used are monthly. We have a measure for interest rates (Moody's Seasoned Aaa Corporate Bond Yield) running from 1919:01 to 2007:04, and a measure for inflation rate, based on the Consumer Price Index), running from 1921:02 to 2007:4. To measure Economic Activity we use the Industrial Production Index, available from 1921:1 to 2007:4. We also have data for money stock. We have data for M1 (since 1947:01) and M2 (since 1948:1). All data is available at the Federal Reserve Bank of St. Louis (data for M1 and M2 were complemented with the estimations provided by Rasche, 1987).

Data for industrial production and the money stocks were transformed in logarithms. The trend was removed using a wavelet based filter (see Cronley 2007), which has properties similar to a band pass filter. In Figures 1-5 we can see the continuous wavelet power spectrum of the several variables.

In Figure 2,[5] we see the time-scale decomposition of interest rates. It is clear that most of the action, specially at high scales (low frequencies) appears after 1960s, suggesting a structural change in that decade. Figure 3 tells us that until 1950s inflation rate variance was quite high both at low and high scales. Again in 1970s and 1980s, probably as a consequence of very active oil shocks, the variance of the inflation rate became higher, but in this case, the effect is clearer at medium and high scales, suggesting that we were facing permanent shocks to inflation.

Figure 4, suggests that the variance, at all scales of the industrial production was quite high until 1950s. After that it has been steadily decreasing, with an exception between mid

[5] The thick black contour designates the 5% significance level. The cone of influence where edge effects might distort the picture is shown as a lighter shade.

1970s and mid 1980s, when the variance at the business cycle frequency (2 to 8 years) was quite high. It has become common in the literature to argue that in the last decade we observed of decrease in the volatility of GDP in the United States. Some authors call it the "Great Moderation". In reality, we can observe that this is a secular, and not decadal, trend. Before the 2nd World War, the volatility was quite high at all scales (at least scales above 6 months). In the 1960s, the volatility decreased at all scales, and the increase at the business cycle frequency in the 1970s, probably due to the oil shocks, was temporary.

Figures 5 and 6 are interesting, because we can compare the different evolution of the behavior of two different monetary aggregates. The volatility of M1 is very high at very low scale (high frequencies), which is something we do not observe for M2. It is also very clear the different behavior in the 1970s, with M2 with a very high power in the $3 \sim 6$ year scale, while M1 only became more active after 1980, suggesting a change in the monetary policy.

In Figures 7 to 10,[6] we can observe the estimated cross wavelet between the interest rates and several other variables. We focus on the interest rates because according to Sims (1980, 1992) the role of money in output determination is very minor, when interest rates are included in the system. According to his results interest rates play a leading role both in determining output and inflation.

In Figure 7, we can see that in the 1970s and 1980s the covariance between inflation and interest rates was quite high in the $3 \sim 20$ year scale. Here can see the big advantage of using wavelets. Note that the causality is not the same at the different scales. Arrows pointing down and to the right (in the 3-8 year scales) suggest that these variables are procyclical, with the inflation rates leading. In the 12-20 year scales arrows point down and to the left, suggesting that the variables behave anti-cyclically and with the interest rates leading. This suggests that, at the business cycle frequency, interest rate increases follow inflation

[6] The 5% significance level is shown with a thick contour. The relative phase relationship is shown as arrows. In-phase pointing right, anti-phase pointing left.

In-phase relations with interest rates leading (lagging) pointing up (down) and to the right. Anti-phase realations with interest rates leading (lagging) pointing down (up) and to the left.

increases, but in the longer run the increase in the interest make inflation rates fall. This type of conclusion would not be easy to get if one is restricted to time series or frequency analysis methods.

Figure 8 tells us that in 1920s and 1930s increases in the interest rates preceded decreases in the industrial production suggesting that Friedman was right, when blaming the contractionary monetary policy for the big recession. In late 1950s and in the decade of 1960, long run changes (at the 10-14 year scales) in the interest rates caused anti-cyclical movements in the industrial production. In the 1970s and 1980s this effect was extended to the business cycle frequency (4-10 years). Interestingly, and starting in 1980, coinciding with Paul Volker as a governor of the Federal Reserve, one can see that interest rates, in the 2-4 year band, reacted procyclically with industrial production, having contractionary effects in the longer run.

Figures 9 and 10, describe to us the time-scale relation between interest rates and M1 and M2 after the second World War. A structural change in this relation has clearly happen in 1970s, probably coinciding with the end of the monetary targeting. Generally, arrows point to the left, suggesting an obvious anticyclical relation, higher interest rates correspond to contractionary monetary policies. Arrows pointing down, typically at lower scales, higher frequencies, suggest that interest rates lead the money stock, while in the long run (higher scales) the opposite happens.

While the cross wavelet transform gives us something that can be interpreted as the covariance between two variables at different time-scales, the wavelet coherency can be interpreted as the correlation; therefore complementing the previous analysis and highlighting relations that could, otherwise, remain hidden.

In Figure 10, we can observe that the relation between interest rates and inflation has changed a lot. In the 1930s the relation is not very strong, except for an island in the 6-9 year band, where arrows pointing upwards and to the left suggest that inflation was the leading variable, with anticyclical effetcs on the interest rates. After 1960s, strong medium and long run relations are uncovered confirming the conclusions we reached with the cross wavelet analysis. More iluminating is Figure 11. A negative relation between interest rates and industrial production is uncovered in the 1930s at the business cycle frequency. In 1950s, interest rates react procyclically to changes in the industrial production (1-5 year band). In the 1970s and 1980s, in the 3-10 year band, arrows suggest that increases in the interest rates had contractionary effects, confirming the analysis of the authors (examples???) who argued that monetary policy reinforced the effects of the oil shocks. Still in the 1980s, but at lower scales (0.5-1.5 and 2-5 year band), interest rates seem to follow industrial production, with the negative long run effects already noted. After 2000, this pattern seems to have moved to higher scales (but since this island is under the cone of influence it is still to early for decent inference).

Figures 12 and 13 show that the relation between money stocks and interest rates has changed quite often since 1950. For example, if we focus on Figure 12, we can see that in the 1950s the interest rate was the leading variable, with M2 reacting anticyclically. In the decade of 1960, in the 1-3 year band, M2 became the leading variable, with the interest rate responding in a procyclical way. In the 1970s, at the business cycle frequencies, 2-8 year band, interest rates and M2 varied anticyclically, and it is not clear which mariable was the leading one. But in the 1980s, at higher scales, 6-8 year M2 became the leading variable, while at lower scales, 1-3 year, interest rates were leading. A similar pattern continued to be observed in the 1990s.

References

[Dau92] Daubechies, I. (1992) **Ten Lectures on Wavelets**, *CBMSNSF Regional Conference Series in Applied Mathematics,*.vol. 61 SIAM, Philadelphia,

[Fla88] Flandrin, P. (1988), **Time-frequency and time-scale**, in *IEEE Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, pp 77–80, Minneapolis, Minnesota.

[FK93] Foufoula-Gergiou, E. and Kumar, P. (1993) , **Wavelts in Geophysics**, volume 4 of *Wavelet Analysis and Its Applications*. Academic Press, Boston, 1993.

[GGM84] Goupillaudand, P., A. Grossman, and J. Morlet (1984), **Cycle-octave and related transforms in seismic signal analysis**, *Geoexploration*, 23, 85–102.

[Kai94] Kaiser, G. (1994), **A Friendly Guide to Wavelets**. Birkhäuser, Basel, 1994.

[Kör88] Körner, T. W. (1988), **Fourier Analyis**, Cambridge University Press, Cambridge, 1988.

[Mal98] Mallat, S. (1998), **A Wavelet Tour of Signal Processing**, Academic Press, New York, 1998.

[Mes61] Messiah, A. (1961), **Quantum Mechanics**, North-Holland, Amsterdam.

[Mey93] Meyer, Y. (1993), *Wavelets: Algorithms and Applications*, SIAM, Philadelphia.

[Sim80] Sims, C. A. (1980) Comparison of interwar and postwar business cycles: Monetarism reconsidered, *American Economic Review*, 70, 250–7.

[Sim92] Sims, C. A. (1992) Interpreting the macroeconomic time series facts, *European Economic Review*, 36, 975–1011.

[TC98] C. Torrence, C. and Compo, G. P. (1998), **A practical guide to wavelet analysis**, *Bulletin of the American Meteorological Society*, 79, 605–618.

[TW99] Torrence, C. and Webster, P. (1999), **Interdecadal changes in the esnomonsoon system**, *Journal of Climate*, 12,2679–2690.

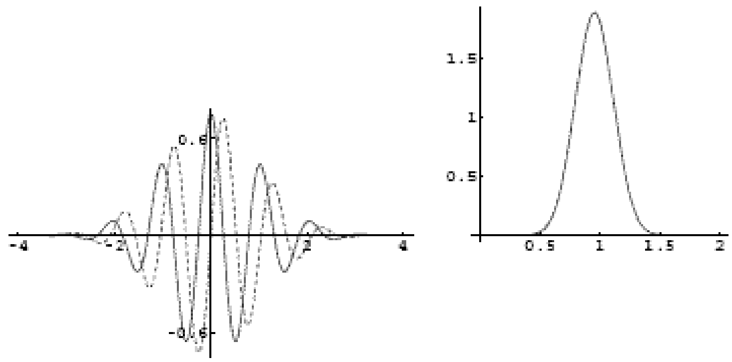

Figure 1: The Morlet wavelet $\psi_6\,(t)$: Real part – solid line and imaginary part – dashed line (on the left) and its Fourier transform (on the right).

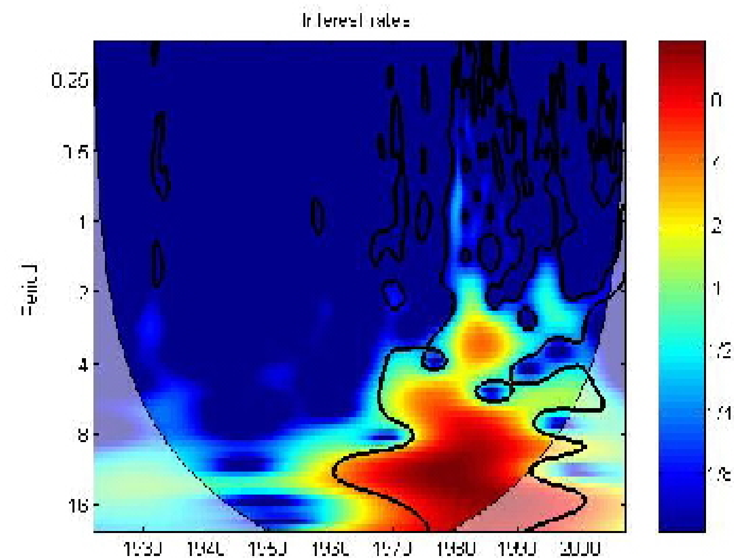

Figure 2: The continuous wavelet power spectrum of Interest Rates.

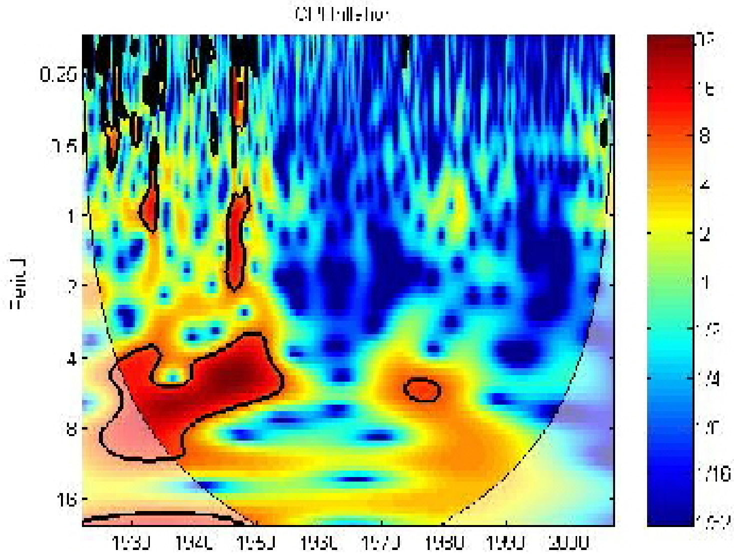

Figure 3: The continuous wavelet power spectrum of Inflation.

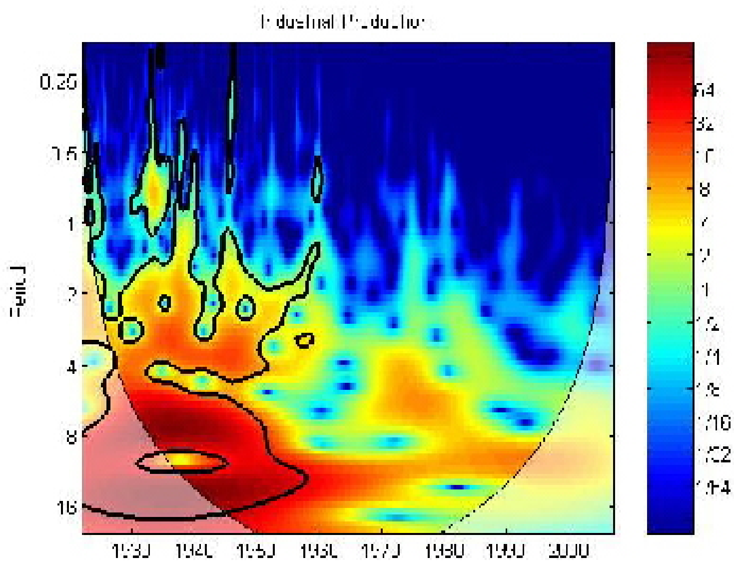

Figure 4: The continuous wavelet power spectrum of the Industrial Production Index.

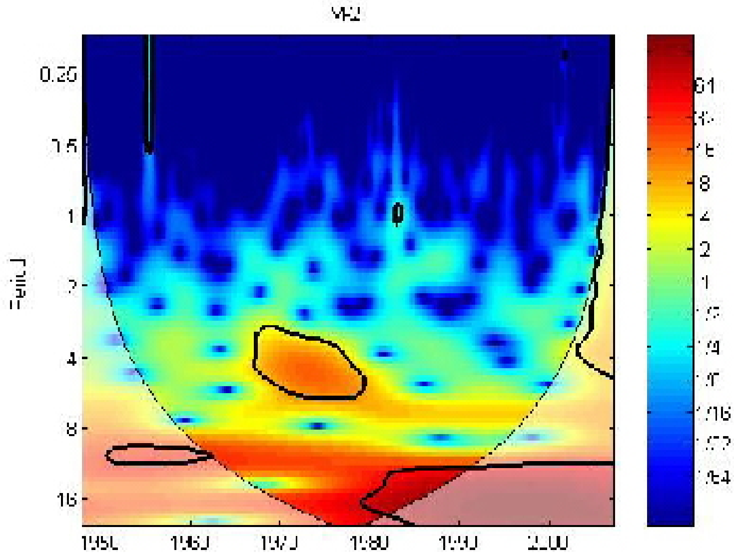

Figure 5: The continuous wavelet power spectrum of M2.

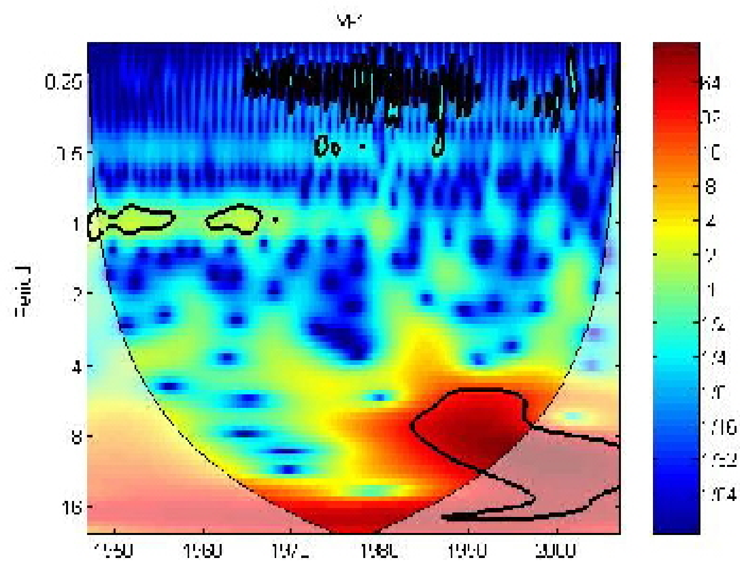

Figure 6: The continuous wavelet power spectrum of M1.

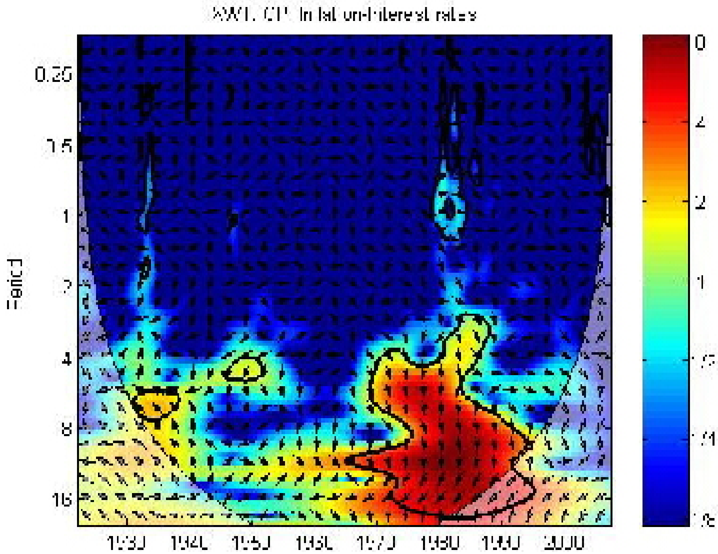

Figure 7: Cross Wavelet Transform between Inflation and Interest Rates

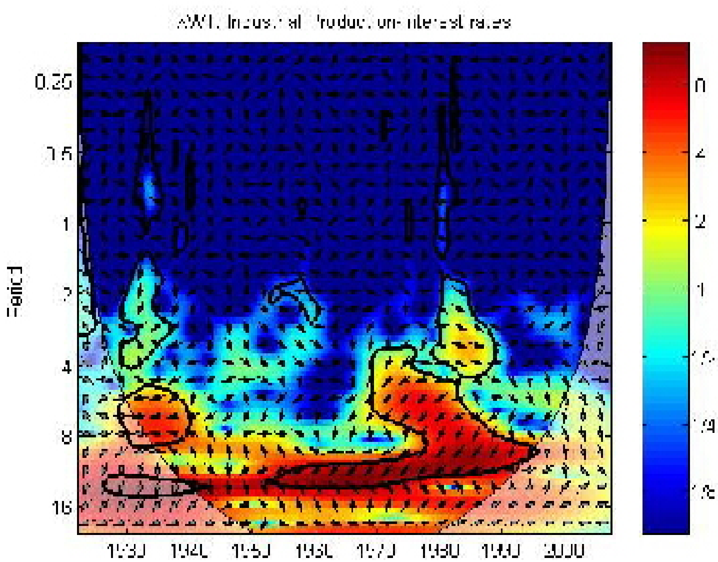

Figure 8: Cross Wavelet Transform between Industrial Production and Interest Rates

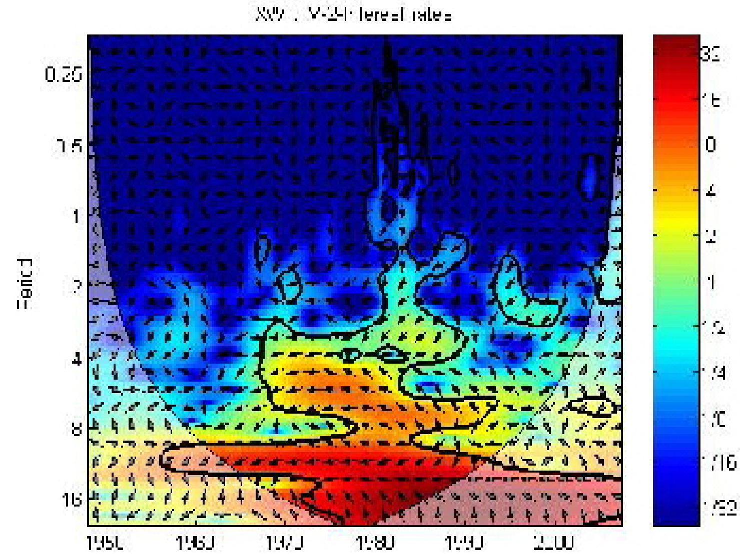

Figure 9: Cross Wavelet Transform between M2 and Interest Rates

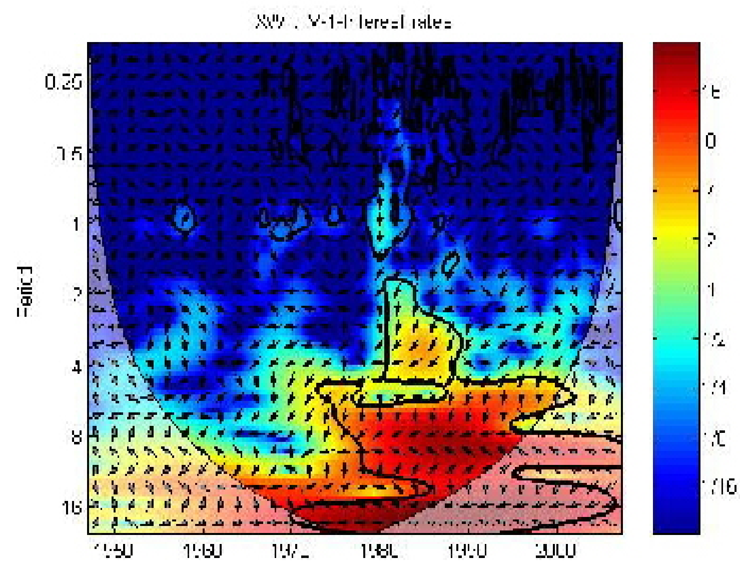

Figure 10: Cross Wavelet Transform between M1 and Interest Rates

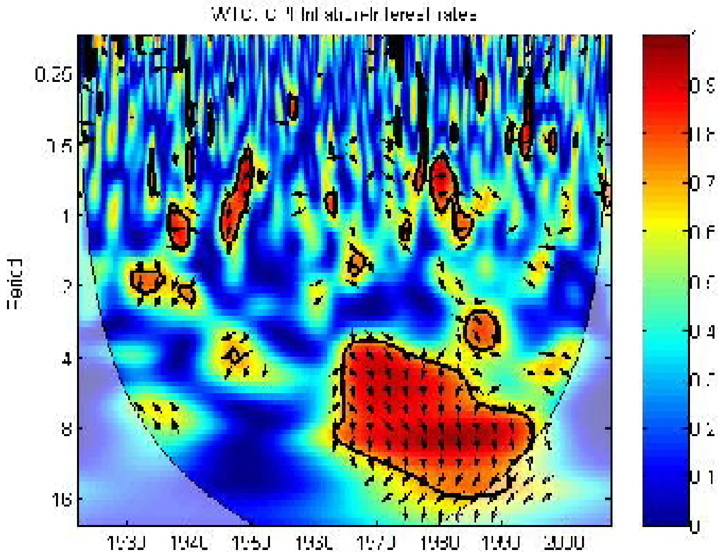

Figure 11: Squared wavelet coherence between Interest Rates and Inflation.

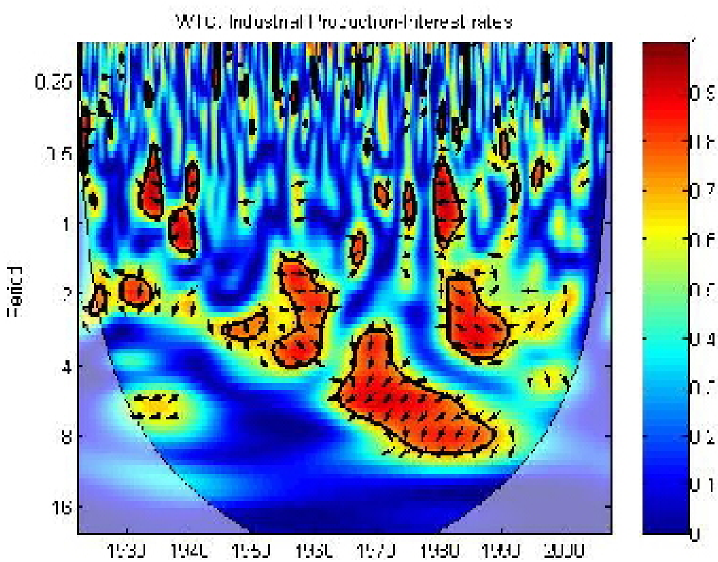

Figure 12: Squared wavelet coherence between Interest Rates and Industrial Production

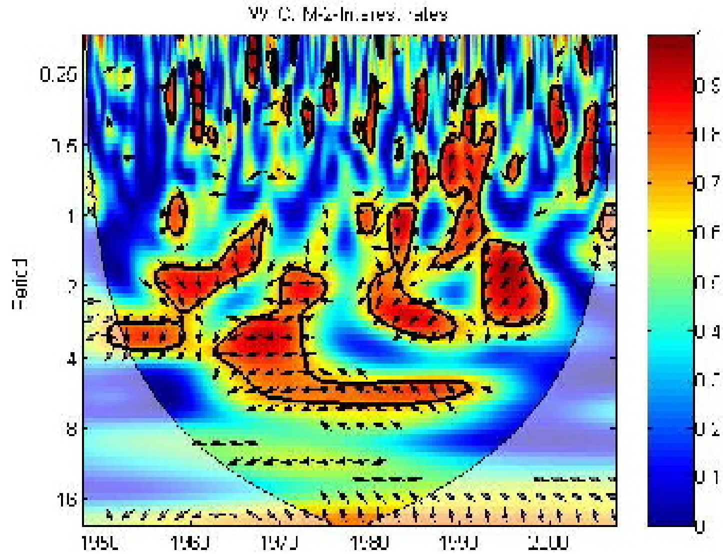

Figure 13: Squared wavelet coherence between Interest Rates and M2.

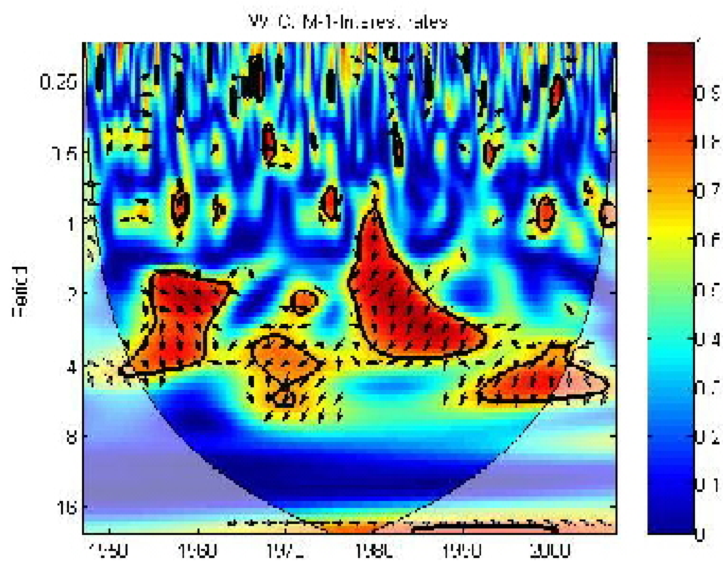

Figure 14: Squared wavelet coherence between Interest Rates and M1.