Using wavelets to decompose the time–frequency effects of monetary policy

Luís Aguiar-Conraria\textsuperscript{a,}\textsuperscript{*}, Nuno Azevedo\textsuperscript{b}, Maria Joana Soares\textsuperscript{c}

\textsuperscript{a}Departamento de Economia, Universidade do Minho, Portugal
\textsuperscript{b}Faculdade de Ciências, Universidade do Porto, Portugal
\textsuperscript{c}Departamento de Matemática, Universidade de Minho, Portugal

Received 11 September 2007; received in revised form 27 December 2007
Available online 16 January 2008

Abstract

Central banks have different objectives in the short and long run. Governments operate simultaneously at different timescales. Many economic processes are the result of the actions of several agents, who have different term objectives. Therefore, a macroeconomic time series is a combination of components operating on different frequencies. Several questions about economic time series are connected to the understanding of the behavior of key variables at different frequencies over time, but this type of information is difficult to uncover using pure time-domain or pure frequency-domain methods.

To our knowledge, for the first time in an economic setup, we use cross-wavelet tools to show that the relation between monetary policy variables and macroeconomic variables has changed and evolved with time. These changes are not homogeneous across the different frequencies.

\textcopyright 2008 Elsevier B.V. All rights reserved.

Keywords: Monetary policy; Time–frequency analysis; Non-stationary time series; Wavelets; Cross-wavelets; Wavelet coherency

1. Introduction

Central banks have different objectives in the short and long run and they operate simultaneously at different timescales — see Ref. [40]. Moreover, many economic processes are the result of the actions of several agents, who have different term objectives, with some agents focusing on daily movements and co-movements, while other agents are concerned about longer horizons. Therefore, a macroeconomic time series is a combination of components operating on different frequencies. On top of this, some interesting relations may exist between two macroeconomic time series at different frequencies. For example, it is possible that monetary policies have different impacts in the short or long run, therefore affecting the economy in different ways at different frequencies. Or, it is possible that monetary...
authorities react to inflation news in the short run, while, in the long run, the price level is essentially determined by the money supply. Finally, it is possible that the effects of a certain policy evolve with time, as institutions and policymakers change. While several questions about time series economic data are connected to the understanding of the behavior of key variables at different frequencies over time, this type of information is difficult to uncover using pure time-domain or pure frequency-domain methods.

We use wavelets to analyze the impact of interest rate price changes on some macroeconomic variables: industrial production, inflation and the monetary aggregates M1 and M2. Specifically, we utilize the continuous wavelet power spectrum and three cross-wavelet tools: the cross-wavelet power spectrum, the wavelet coherency and the wavelet phase difference. With these instruments we are able to unravel some economic time–frequency relations that have remained hidden so far.

With the wavelet power spectrum, we show that the “great moderation” is not a recent phenomenon and that the observed reduction in the volatility of production happened in the 1950s and not in the 1980s as has been erroneously assumed in the economic literature. The volatility was temporarily revived during the oil crisis of the 1970s and 1980s, mostly at business cycle frequencies. This is specifically one of the advantages of the wavelet analysis: the possibility of uncovering transient relations. The same conclusion is reached about inflation. The volatility of inflation decreased at the same time. Therefore, the great moderation is not just a real phenomenon but also a nominal phenomenon.

With the cross-wavelet tools we show that the relation between monetary policy variables (money aggregates and interest rates) and macroeconomic variables (industrial production and inflation) has changed and evolved with time and is not homogeneous across the different frequencies. For example, in the 1970s and 1980s we observe that interest rates reacted procyclically with inflation at the business cycle frequencies, and that at lower frequencies this helped to control inflation. We also find evidence that in the 1950s interest rates were reacting to industrial production, in the 2–4 year period, but that in the 1970s and 1980s, especially in the 4–12 year period, we observe an anti-phase relation with interest rates leading, meaning that increases in the interest rates had contractionary effects, supporting the conclusions of some authors [1,30] who argued that monetary policy reinforced the recessionary effects of the oil shocks. We also find evidence of a structural break in the relation between the interest rates and the monetary aggregates. In the late 1970s, early 1980s, at the business cycle frequency, the broader definition of money, M2, stopped lagging the interest rates. At lower frequencies, corresponding to 15–20 year period oscillations, M1 started leading the interest rates. This suggest that the Fed has been following a monetary targeting type of policy, even if they are not explicit about it (except between the mid-1970s and mid-1980s).

The paper proceeds as follows. In Section 2, we discuss the main advantages of wavelet analysis, its applications to economics and some of the typical difficulties in applying it to study economic relations. We present the continuous wavelet transform, and discuss its localization properties and the optimal characteristics of the Morlet wavelet. Section 3 describes the wavelet power spectrum, the cross-wavelet power spectrum, the wavelet coherency, and the phase difference. In Section 4, we apply these tools to study the effects and the effectiveness of monetary policy. Section 5 concludes.

2. Wavelets: The dynamical decomposition of time

In the economic literature it is common to utilize Fourier analysis to uncover relations at different frequencies. For example, spectral techniques can be used to identify some stylized business cycle facts [21,28], and seasonal components [37,50], or to highlight different relations among economic variables at distinct frequencies, such as the paradoxical relations between production and inventories [51]. In spite of its utility, under the Fourier transform, the time information of a time series is completely lost. Because of this loss of information, it is hard to distinguish transient relations or to identify structural changes. Moreover, these techniques are only appropriate for time series with stable statistical properties, i.e. stationary time series. Unfortunately, typical economic time series are noisy, complex and strongly non-stationary.

To overcome the problems of analyzing non-stationary data, Gabor [13] introduced the short time Fourier transform. The basic idea is to break a time series into smaller sub-samples and apply the Fourier transform to each sub-sample. However, as Raihan et al. [42] pointed out, this approach is inefficient because the frequency resolution is the same across all different frequencies. As an alternative, wavelet analysis has been proposed.

1 The total stock of money in the economy consists primarily of currency in circulation, checking accounts and savings accounts. M1, also known as narrow money, includes currency in circulation and checking accounts. M2, also known as broad money, includes M1 and the savings accounts.
Wavelet analysis performs the estimation of the spectral characteristics of a time series as a function of time, revealing how the different periodic components of the time series change over time. 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.\(^2\)

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, 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.

As a coherent mathematical body, wavelet theory was born in the mid-1980s \(^{23,20}\). After 1990, the literature rapidly expanded and wavelet analysis is now extensively used in physics, geophysics, astronomy, epidemiology, signal processing, oceanography, etc. Interestingly, and in spite of all its potential, this technique is infrequently used in economics. The pioneering work of Goffe \(^{19}\), Ramsey and Lampart \(^{40,41}\), Ramsey \(^{38}\), and Gençay et al. \(^{15,16}\) is largely unknown to the majority of the economists, who reveal a strong preference for traditional econometric methods, overlooking the potential for using wavelets to analyze economic data. Among the exceptions to this rule, one can point to recent works on finance, for example, Refs. \(^{6,18,44,9,14}\).

One peculiarity of the applications of wavelets to economics is the almost exclusive use of the discrete wavelet transform. Common to all the cited economic applications of wavelets is the use of the discrete wavelet transform, or one of its variants.\(^4\) For example, Gallegati and Gallegati use the maximal overlap discrete wavelet transform to decompose the industrial production of the G-7 countries since 1961 on several scales. Then, to analyze the evolution of the volatility of the of real economic activity, they estimate the wavelet variance at each scale, for each decade separately. This obsession with the discrete wavelet transform is difficult to understand. Sometimes, the same type of analysis could be done more easily and straightforwardly using the continuous wavelet transform. For example, looking at Fig. 2, one can immediately infer the evolution of the variance of the industrial production index at the several timescales along the decades and extract the conclusions of Ref. \(^{14}\) with one picture.

Probably, there are two reasons for wavelets not being more popular among economists. First, in most economic applications the (discrete) wavelet transform has mainly been used as a low and high pass filter, it being hard to convince an economist that the same could not be learned from the data using the more traditional, in economics, band pass-filtering methods — see Ref. \(^{2}\). The second reason is related to the difficulty of analyzing simultaneously two (or more) time series. In economics, these techniques have either been applied to analyze individual time series (e.g. \(^{14}\)) or used to individually analyze several time series (one each time), whose decompositions are then studied using traditional time-domain methods, such as correlation analysis or Granger causality (e.g. Refs. \(^{40,41}\)).

\(^2\) We know from the Heisenberg uncertainty principle that there is always a trade-off between localization in time and localization in frequency. However, a mother wavelet can be chosen with a fast decay in time and frequency which, for all practical purposes, corresponds to an effective band and time limiting; see Ref. \(^{8}\).

\(^3\) For detailed reviews of wavelet applications to economic and financial data, the reader is referred to Refs. \(^{17,39}\) or, more recently, Ref. \(^{7}\).

\(^4\) To our knowledge, Ref. \(^{42}\) is the only exception to this rule.
To overcome these problems, in the 1990s, wavelet tools were generalized to accommodate the analysis of time–frequency dependencies between two time series. The cross-wavelet power, the cross-wavelet coherency, and the phase difference, proposed by Hudgins et al. [24] and Torrence and Compo [49], have been applied in different scientific fields, ranging from medicine [52] and epidemiology [5] to astrophysics [27,4] and geophysics [25,22]. With cross-wavelet tools, we are able to use wavelet analysis to directly study the interactions between two time series at different frequencies and how they evolve over time. While the (single) wavelet power spectrum describes the evolution of the variance of a time series at the different frequencies, with periods of large variance associated with periods of large power at the different scales, the cross-wavelet power of two time series describes the local covariance...
between the time series — Ref. [4]. On the other hand, one can look at the wavelet coherency as a localized correlation coefficient in the time–frequency space — Ref. [22]. The phase can be viewed as the position in the pseudo-cycle of the series as a function of frequency; therefore the phase difference gives us information on the delay, or synchronization, between oscillations of the two time series — Ref. [4].

2.1. The wavelet

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

\[
\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty.
\]

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 is well known, one can define in \( L^2(\mathbb{R}) \) an inner product

\[
\langle x, y \rangle := \int_{-\infty}^{\infty} x(t) y^*(t) dt
\]

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

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

\[
X(f) := \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} dt.
\]

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

\[
\langle x(t), y(t) \rangle = \langle X(f), Y(f) \rangle,
\]

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

\[
\|x(t)\|^2 = \|X(f)\|^2;
\]

see, for example, Ref. [29].

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

\[
0 < C_\psi := \int_{-\infty}^{\infty} \frac{|\Psi(f)|}{|f|} df < \infty,
\]

where \( \Psi(f) \) is the Fourier transform of \( \psi(t) \); see Ref. [8] p. 24.

The wavelet \( \psi \) is usually normalized to have unit energy: \( \|\psi\|^2 = \int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1 \). The square integrability of \( \psi \) is a very mild decay condition; the wavelets used in practice have much faster decay; typical behavior will be exponential decay or even compact support.

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

\[
\psi(0) = \int_{-\infty}^{\infty} \psi(t) dt = 0
\]


This means that the function \( \psi \) 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 \( \psi \).
2.2. The continuous wavelet transform

Starting with a mother wavelet \( \psi \), a family \( \psi_{s, \tau} \) of “wavelet daughters” can be obtained by simply scaling \( \psi \) by \( s \) and translating it by \( \tau \):

\[
\psi_{s, \tau}(t) := \frac{1}{\sqrt{|s|}} \psi \left( \frac{t - \tau}{s} \right), \quad s, \tau \in \mathbb{R}, s \neq 0.
\]  

(8)

The parameter \( s \) is a scaling or dilation factor that controls the length of the wavelet (the factor \( 1/\sqrt{|s|} \) being introduced to guarantee preservation of the unit energy, \( \| \psi_{s, \tau} \| = 1 \)) and \( \tau \) 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 \( |s| < 1 \)).

Given a function \( x(t) \in L^2(\mathbb{R}) \) (a time series), its continuous wavelet transform (CWT) with respect to the wavelet \( \psi \) is a function \( W_x(s, \tau) \) obtained by projecting \( x(t) \), in the \( L^2 \) sense, onto the overcomplete family \( \{ \psi_{s, \tau} \} \):

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

(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. Ref. [8]:

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

(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 could, 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

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

(11)

and that a Parseval type identity also holds:

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

(12)

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

Because the wavelet function \( \psi(t) \) may, in general, be complex, the wavelet transform \( W_x \) may also be complex. The transform can then be divided into its real part, \( \mathcal{R}\{W_x\} \), and its imaginary part, \( \mathcal{I}\{W_x\} \), or in its amplitude, \( |W_x| \), and phase, \( \phi_x(s, \tau) = \tan^{-1} \left( \frac{\mathcal{I}\{W_x\}}{\mathcal{R}\{W_x\}} \right) \). The phase of a given time series \( x(t) \) can be viewed as the position in the pseudo-cycle of the series. 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:

\[
x(t) = \frac{2}{C_\psi} \int_{0}^{\infty} \left[ \int_{-\infty}^{\infty} \mathcal{R}(W_x(s, \tau) \psi_{s, \tau}(t)) \, d \tau \right] \, ds;
\]  

(13)

one also has

\[
\|x\|^2 = \frac{2}{C_\psi} \int_{0}^{\infty} \left[ \int_{-\infty}^{\infty} |W_x(s, \tau)|^2 \, d \tau \right] ds
\]  

(14)

Note that for negative \( s \), the function is also reflected.
and

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

(15)

see Refs. [8, pp. 27–28], [26, pp. 70–73] or [32, pp. 82–83] for more details about analytic wavelets. Throughout the rest of the paper, since, in the practical economic applications we deal with real data, we will use an analytic wavelet and always assume that the scaling parameter \( s \) takes positive values only.

2.3. Localization properties

Let the wavelet \( \psi \) be normalized so that \( \| \psi \| = 1 \) and define its center \( \mu_t \) by

\[ \mu_t = \int_{-\infty}^{\infty} t|\psi(t)|^2 \, dt. \]  

(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 \( \psi \) around its center one usually takes the standard deviation \( \sigma_t \):

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

(17)

In a totally similar manner, one can also define the center \( \mu_f \) and variance \( \sigma_f \) of the Fourier transform \( \Psi(f) \) of \( \psi \).

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

The uncertainty principle, first established by Werner Karl Heisenberg in the context of quantum mechanics, 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. In our context, the Heisenberg uncertainty principle establishes that the uncertainty is bounded from below by the quantity \( 1/4\pi \):

\[ \sigma_t \sigma_f \geq \frac{1}{4\pi}. \]  

(18)

It is also known that the equality in (18) is attained if and only if the function \( \psi \) is a (translated and modulated) Gaussian: \( \psi(t) = ae^{i\mu_1}e^{-b(t-\mu)^2} \); see Ref. [35].

It follows from the Parseval relation (4) that

\[ W_x(s, \tau) = \langle x(t), \psi_{s,t}(t) \rangle = \langle X(f), \Psi_{s,t}(f) \rangle \]  

(19)

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

If the mother wavelet \( \psi \) is centered at \( \mu_t \), has variance \( \sigma_t \) and its wavelet transform \( \Psi(f) \) is centered at \( \mu_f \) with a variance \( \sigma_f \), then one can easily show that the daughter wavelet \( \psi_{t,s} \) will be centered at \( \tau + s\mu_t \) with variance \( \sigma_t \), whilst its Fourier transform \( \Psi_{t,s} \) 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

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

(20)

In particular, if \( \psi \) is chosen so that \( \mu_t = 0 \) and \( \mu_f = 1 \), then the window associated with \( \psi_{t,s} \) becomes

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

(21)
In this case, the wavelet transform \( W_x(s, \tau) \) will give us information on \( x(t) \) for \( t \) near the instant \( t = \tau \), with precision \( \sigma_t \), 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 \( x(t) \) on a fine scale and about \( X(f) \) on a broad scale,
- large values of \( s \) correspond to information on a broad scale about \( x(t) \) and on 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 \( s \rightarrow \text{low frequencies } f = \frac{1}{s} \)) and compress for small values of \( s \) (fine scale \( s \rightarrow \text{high frequencies } f = \frac{1}{s} \)).

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, Daubechies, etc.; see, e.g. Refs. [8,32] or [36]. 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, which will depend on the particular application that one has in mind. We choose a complex wavelet, as it yields a complex transform, with information on both the amplitude and phase, essential for studying the business cycle synchronism between different time series. One of the most popular wavelets used is the Morlet wavelet, first introduced in Ref. [20], which is defined as

\[
\psi_\eta(t) = \pi^{-\frac{1}{4}} (\cos \eta t - e^{-\frac{\eta^2}{2}})e^{-\frac{t^2}{2}},
\]

(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. The simplified version

\[
\psi_\eta(t) = \pi^{-\frac{1}{4}} e^{i\eta t} e^{-\frac{t^2}{2}}
\]

(23)

of (22) is normally used (and still referred to as a Morlet wavelet). Fig. 1 and the 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 \((-\infty, 0)\):

\[
\Psi_\eta(f) = \pi^{\frac{1}{4}} \sqrt{2} e^{-\frac{1}{4} (2\pi f - \eta)^2}.
\]

(24)

For \( \eta > 5 \), this mass is, however, negligible, so, for all practical purposes, the wavelet can be considered as analytic; see Ref. [10].

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

\[
\mu_f = \frac{6}{2\pi} \approx 1
\]

(25)

and the relationship between the scale and frequency is simply

\[
f = \frac{\mu_f}{s} \approx \frac{1}{s}.
\]

(26)

It is simple to verify that the time variance is \( \sigma_t = 1/\sqrt{2} \) and the frequency variance is \( \sigma_f = 1/(2\pi \sqrt{2}) \). Therefore, the uncertainty of the corresponding Heisenberg box attains the minimum possible value \( \sigma_t\sigma_f = \frac{1}{4\pi} \) and the Morlet wavelet has optimal joint time–frequency concentration.\(^6\)

\(^6\) This could be anticipated by noting that \( \psi_\eta \) is a simple modulated Gaussian.
2.5. Transform of finite discrete data

If one is dealing with a discrete time series \( \{x_n, n = 0, \ldots, N-1\} \) of \( N \) observations with a uniform time step \( \delta t \), 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

\[
W_x^m(s) = \frac{\delta t}{\sqrt{s}} \sum_{n=0}^{N-1} x_n \psi^* \left( \frac{(n - m) \delta t}{s} \right), \quad m = 0, 1, \ldots, N - 1.
\] (27)

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 algorithm to go forth and back from time to spectral domain; this is the technique prescribed by Torrence and Compo [49].

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. 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 Ref. [49], as the e-folding time of the wavelet at the scale \( s \), that is, so that the wavelet power of a Dirac \( \delta \) at the edges decreases by a factor of \( e^{-2} \). In the case of the Morlet wavelet this is given by \( \sqrt{2} s \).

3. Wavelet tools

3.1. Wavelet power spectrum

In view of the energy preservation formula (14), and in analogy with the terminology used in the Fourier case, we simply define the wavelet power spectrum as \( |W_x^m|^2 \), which gives us a measure of the local variance. Following [49], the statistical significance of wavelet power can be assessed against the null hypothesis that the data generating process is given by a stationary process with a certain background power spectrum (\( P_f \)). Torrence and Compo computed the white noise and red noise wavelet power spectra, from which they derived, under the null, the corresponding distribution for the local wavelet power spectrum,

\[
D \left( \frac{|W_x^m(s)|^2}{\sigma_s^2} < p \right) = \frac{1}{2} P_f \chi^2_v,
\] (28)

at each time \( n \) and scale \( s \). The value of \( P_f \) is the mean spectrum at the Fourier frequency \( f \) that corresponds to the wavelet scale \( s \) – in our case \( s \approx \frac{1}{f} \); see Eq. (26) – and \( v \) is equal to 1 or 2, for real or complex wavelets respectively. For more general processes, one has to rely on Monte Carlo simulations.

3.2. Cross-wavelets and phase differences

3.2.1. Cross-wavelet power

The cross-wavelet transform of two time series, \( x = \{x_n\} \) and \( y = \{y_n\} \), first introduced by Hudgins et al. [24], is simply defined as

\[
W_{xy}^m = W_x^m W_y^m*.
\] (29)

where \( W_x^m \) and \( W_y^m \) are the wavelet transforms of \( x \) and \( y \), respectively. The cross-wavelet power is given by \( |W_{xy}^m|^2 \). While we can interpret the wavelet power spectrum as depicting the local variance of a time series, the cross-wavelet

---

7 A program code based on the above procedure is available at the site http://paos.colorado.edu/research/wavelets/.
power of two time series depicts the local 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. For white and red noise processes, Torrence and Compo [49] showed that if two time series have Fourier Spectra $P_f^x$ and $P_f^y$, then the cross-wavelet distribution is given by

$$D \left( \frac{|W_x W_y^*|}{\sigma_x \sigma_y} < p \right) = \frac{Z_v(p)}{v} \sqrt{P_f^x P_f^y},$$

where $Z_v(p)$ is the confidence level associated with the probability $p$ for a pdf defined by the square root of the product of two $\chi^2$ distributions (see Ref. [49] for details). For more general data generating processes one has to rely on Monte Carlo simulations.

3.2.2. Wavelet coherency

As in the Fourier spectral approaches, wavelet coherency can be defined as the ratio of the cross-spectrum to the product of the spectrum of each series, and can be thought of as the local correlation, both in time and frequency, between two time series. We follow Refs. [25,4,5] and define the wavelet coherency between two time series $x = \{x_n\}$ and $y = \{y_n\}$ as follows:

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

where $S$ denotes a smoothing operator in both time and scale. Smoothing is necessary. Without that step, coherency is identically 1 at all scales and times. Smoothing is achieved by a convolution in time and scale. The time convolution is done with a Gaussian and the scale convolution is performed with a rectangular window; see Refs. [5,22] for details.

Theoretical distributions for wavelet coherency have not been derived yet. Therefore, to assess the statistical significance of the estimated wavelet coherency, one has to rely on Monte Carlo simulation methods.

3.2.3. Phase difference

As Bloomfield et al. [4] explain, this tool gives us information about the delays of the oscillations between two time series, $x = \{x_n\}$ and $y = \{y_n\}$, as a function of frequency. As we said before, the phase of a given time series, $\phi_x$, can be viewed as the position in the pseudo-cycle of the series. The phase difference, $\phi_{x,y}$, characterizes phase relationships between the two time series, i.e. their relative position in the pseudo-cycle. The phase difference is defined as

$$\phi_{x,y} = \tan^{-1}\left( \frac{I(W_n^{xy})}{R(W_n^{xy})} \right), \quad \text{with } \phi_{x,y} \in [-\pi, \pi].$$

A phase difference of zero indicates that the time series move together at the specified frequency. If $\phi_{x,y} \in (0, \frac{\pi}{2})$ then the series move in phase, but the time series $y$ leads $x$. If $\phi_{x,y} \in (-\frac{\pi}{2}, 0)$ then it is $x$ that is leading. A phase difference of $\pi$ (or $-\pi$) indicates an anti-phase relation. If $\phi_{x,y} \in (\frac{\pi}{2}, \pi)$ then $x$ is leading. Time series $y$ is leading if $\phi_{x,y} \in (-\pi, -\frac{\pi}{2})$.

4. Interest rates and macroeconomic activity

The correlation between monetary aggregates and real economic activity is a very well documented fact for the developed economies (e.g. see Refs. [11], or [12]). Although this correlation is strong, economists do not agree about the causes [28,31]. Some view the monetary aggregates as, essentially, endogenous variables, which respond to changes in real economic activity. Others claim the opposite, that monetary shocks are the main source of economic fluctuations. Basically, although nobody disputes the correlation between economic activity and money, there is no consensus on the explanation.

---

8 We thank Bernard Cazelles for providing us the code used in Ref. [5]. The code can be found at http://ecologie.snv.jussieu.fr/cazelles/wavelets/. The code of Grinsted, Moore, and Jevrejeva can also be found at http://www.pol.ac.uk/home/research/waveletcoherence/.
According to Ramsey and Lampart [40] the ambiguity of these results may be due to the fact that there are several timescales involved in the relationships and that pure frequency analysis or pure time-domain analysis may mask shifting dynamical relationships between the variables. For example, the timescale at which changes in interest rates affect M1 demand is most likely different from that of its effects on M2. We build on and extend the work of Ramsey and Lampart [40] who used the discrete wavelet transform to decompose the relation between money and income. They separated the time series into a sequence of scales, decomposed orthogonally, and then, using Granger tests, they analyzed the relationships between the variables. They concluded that, at very high frequencies, income Granger-causes money, but, at business cycle frequencies, money Granger-causes income.

We use monthly data. We have a measure for interest rates (Moody’s Seasoned Aaa Corporate Bond Yield) running from 1921:01 to 2007:04, and a measure for the 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 stocks. We have data for M1 (since 1947:01) and M2 (since 1948:1). All data are available at the Federal Reserve Bank of St. Louis.9 Data for industrial production and the money stocks were transformed into logarithms. The trend was removed using a wavelet based filter.

In Fig. 2,10 we can see the continuous wavelet power spectra of the several variables. To assess the statistical significance it is difficult to rely on the formula given by Eq. (28). Its applicability to economic data is quite questionable. The assumption that economic time series follow white noise or an auto-regressive process of order 1 (red noise) is unrealistic. It is common to model economic time series as ARMA processes, which (when the invertibility condition is met) can be inverted into an infinite order auto-regressive process. We will take an agnostic approach and will consider two extreme cases: (1) we will assume that the series are well represented by a white noise and attest the wavelet power significance under this assumption; (2) we will estimate the statistical significance of the wavelet power spectrum from Monte Carlo simulations. We follow Ref. [48] to construct the surrogate series of (2) and randomize the phase of the original data: we apply the Fourier transform to the original time series, randomize the phase values and perform an inverse Fourier transformation with the original amplitude and the randomized phase values. With this transformation, the surrogate series approximately keep the mean, standard deviation, auto-correlation and partial auto-correlation functions and the power spectrum of the original data. Case (1) is equivalent to assuming an ARMA(0, 0) process and case (2) resembles an ARMA(p, q) process, where no pre-conditions are imposed on p and q.

Looking at the timescale decomposition of the several variables some interesting facts are revealed. Most of the action in the interest rates occurred after 1960, especially at high scales (low frequencies), indicative of a structural change in that decade.

The power of the industrial production, across scales, was quite high until 1950s. After that, the volatility at all frequencies steadily decreased, with an exception between the mid-1970s and mid-1980s, when the variance at the business cycle frequency (2–8 years) was quite high as a result of the severe oil crises that hit the world economy in 1973 and 1979 and lasted until the early 1980s. It has become common in the literature to argue that we have been observing, in the last decade, a decrease in the volatility of GDP in the United States (e.g. see Refs. [3,14]). It is known in the literature as the “great moderation”. In reality, we can observe that this is a secular, and not decadal, trend. Before the Second World War, the volatility was quite high at all scales (at least scales above 1 year). In the 1960s, the volatility decreased at all scales. It then increased, due to the oil shocks, at the business cycle frequency in the 1970s; however this increase was temporary.

Until 1950, the inflation rate variance was quite high both at low and high scales. Again in the 1970s and 1980s, as a consequence of very active oil price 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 long term shocks to inflation. Therefore, we also observe a great moderation on the volatility of inflation, which started in the 1950s and was temporarily revived during the oil crises in the 1973 and 1979. These results show that the “great moderation” is not just a real phenomenon but also a monetary phenomenon.

9 Data for M1 and M2 were complemented with the estimations provided by Rasche [43], available at http://www.msu.edu/∼rasche/research/money.htm.
10 Due to the extension of the significant regions under the white noise null, in Fig. 2, it is not easy to distinguish the significant regions from the non-significant regions. Significant regions are the ones associated with warmer colors (red, orange and yellow).
We observe a distinct evolution of the behavior of two different monetary aggregates. In the 1970s, M2 is high powered in the 3–6 year timescale, while M1 only became more active after 1980, indicating a structural change in the monetary policy. This coincides with the presidential election of Ronald Reagan and the beginning of the appointment of Paul Volcker as a chairman of the Federal Reserve. Volcker implemented a very restrictive monetary policy as a reaction to the inflationary pressures of the second oil shock.

To perform the cross-wavelet analysis we could focus on the wavelet coherency, instead of the wavelet cross spectrum for three reasons: (1) there is some redundancy between the two measures, (2) the wavelet coherency has the advantage of being normalized by the power spectrum of the two time series, and (3) Maraun and Kurths [33] show that the wavelet cross spectrum can show strong peaks even for the realization of independent processes suggesting the possibility of spurious significance tests. Still, we will use both tools, not only because they are largely unknown in the economics literature but also because the redundancy is not absolute and they do complement each other. The values for the significance were obtained from Monte Carlo simulations. Again, we took an agnostic approach considering the same two extreme cases that we used to test the wavelet power spectrum.

In Fig. 3 we have the wavelet cross power\textsuperscript{11} and in Fig. 4 we can observe the estimated wavelet coherency and the phase difference between the interest rates and several other variables. We focus on the interest rates because,

\textsuperscript{11} Again, in Fig. 3, when the null hypothesis is associated with a white noise, it may be difficult to distinguish the significant regions from the non-significant regions delimited by white lines. Significant regions are associated with warmer colors (red and yellow) and are the ones that include the phase difference arrows.
Fig. 4. Cross-wavelet coherency. The white contour designates the 5% significance level against a white noise null. The black contour designates the 5% significance level estimated from Monte Carlo simulations using phase randomized surrogate series. The cone of influence, which indicates the region affected by edge effects, is also shown with a dotted line. The color code for coherency ranges from blue (low coherency — close to zero) to red (high coherency — close to one). The phase difference between the two series is indicated by arrows. Arrows pointing to the right mean that the variables are in phase. To the right and up, with interest rates lagging. To the right and down, with interest rates leading. Arrows pointing to the left mean that the variables are out of phase. To the left and up, with interest rates leading. To the left and down, with interest rates lagging.

(according to Sims [45,46] and Loizides and Vamvoukas [31], interest rates play a leading role both in determining output and inflation and the role of money in output determination is very minor, once interest rates are included in the system.)

Maraun and Kurths were correct in their warning about the pitfalls of significance testing the cross-wavelet power. The regions of high significance do depend heavily on the null hypothesis. So the information extracted from Fig. 3 has to be treated with caution. More attention will be given to Fig. 4.

In the 1930s, the relation between inflation and interest rates is strong. The wavelet cross-spectrum shows high power (although this is not confirmed by the significance tests expect for the white noise null) and we observe two islands of high coherency (which are also statistically significant). The phase information about this period shows us an anti-phase relation with the inflation rate leading. In the 1970s and 1980s we observe a quite different behavior. Both the cross-wavelet spectrum and the coherency were quite high in the 3–20 year timescale. Moreover, information on the phases shows us that the relationship was not homogeneous across scales. In the 3–12 year scales these variables are in phase, with the inflation rates leading, indicating that, at the business cycle frequency, interest rate increases follow inflation increases, consistent with a central bank that follows a Taylor rule. In the 12–20 year band, the phase difference is consistently between π/2 and π, arrows pointing upwards and to the left, showing that the variables are

12 According to the Taylor rule the Federal Reserve should change the interest rates in response to real divergences of real GDP from potential GDP and divergences of actual rates of inflation from a target rate of inflation. See Ref. [47].
anti-phase, with interest rates leading. Therefore, in the long run, an increase in the interest rates precedes a decrease in inflation, a result consistent with the notion that in the long run a restrictive monetary policy (high interest rates) does help to control inflation. Note that if we were restricted to classical time series we would have no information about the frequency differences; hence this type of conclusion would not be easy to reach if we were to be restricted to classical methods.

It is difficult to find a strong relation between interest rates and the industrial production. The cross-wavelet power spectrum is high at low frequencies, but this relation is statistically non-significant (unless we consider the unrealistic white noise assumption). Looking at the wavelet coherency, we see some statistically significant islands. During the 1920s and 1930s, increases in the interest rates preceded decreases in the industrial production giving some support to Milton Friedman, who blamed the contractionary monetary policy for aggravating the effects of the big recession. But this effect is not strong and it is under the effect of the cone of influence. In the 1950s, we observe a region of high coherency in the 2–4 year band. Variables are in phase, with industrial production leading. Suggesting again that the Fed is following a Taylor rule at the business cycle frequency, but at higher scales (4–12), and specially in the 1970s and 1980s, we observe an anti-phase relation with interest rates leading, meaning that in the long run increases in the interest rates have contractionary effects. This supports the conclusions of some authors [1,30] who argued that monetary policy reinforced the recessionary effects of the oil shocks. Still in the 1980s, but at lower scales (2–4 year band), interest rates seem to follow industrial production, with the negative long run effects already noted. This analysis shows that a Taylor rule type of policy does have some effects in stabilizing the economy in the long run.

Looking at Figs. 3 and 4, one might think that the relation between interest rates and the different definitions of money (M1 and M2) is the same, especially if attention was not being paid to the arrows. The cross-wavelet spectrum is large at high scales and the significance tests are similar in the two situations. If we take the white noise as the null then all this area is significant and, if we consider the alternative, only in the late 1980s and early 1990s do we have a significant cross spectrum. The wavelet coherency is high at very large scales (above the 15 year scale) and at the business cycles frequencies, below eight years, we also see several statistically significant islands.

It is when we analyze at the behavior of the phase differences that we realize some important distinctions. Looking at the long run (15–20 year band), one can see that both measures of money are out of phase with the interest rates as expected (low interest rates are synonymous with loose monetary policies), but while the interest rates lag M2 for the entire sample, in the first half of the sample interest rates lead M1, and in the other half we observe the opposite. In the 4–8 year band, we observe a change in the behavior of M2 in the early 1970s, when the interest rates ceased being the leading variables. The relation between M1 and the interest rate remained almost unchanged, except for a period in the 1980s, coinciding with the time of Paul Volcker as the chairman of the Fed and with the period of M1 targeting implemented by the Fed in the mid-1970s. For most of the period, interest rates were the leading variable. These results are clearer in Fig. 5, which summarizes this information. So while the relation between the interest rates and M1 saw a structural change in the very long run (15–20 year band), we reach a different conclusion for M2, whose relationship with interest rates changed at the business cycle frequencies.
5. Conclusions

In this paper, we claimed that wavelet analysis can be very useful for analyzing economic relations and that it is better suited for dealing with economic data than the Fourier transform. We illustrated how wavelet analysis can naturally be applied to the study of business cycles (given its periodic nature), or to any field of economics, or finance, especially when there is a distinction between short and long run relations. The main advantage of the wavelet approach is the ability to analyze transient dynamics, for single time series or for the association between two time series.

We have also showed that some of the shortcomings that economists have found when applying wavelet techniques to study two or more time series disappear once the concept of the cross-wavelet is introduced. We used three tools that, to our knowledge, have not yet been used by economists: the cross-wavelet transform, the cross-wavelet coherency and the phase difference. While the wavelet power spectra quantifies the main periodic component of a given time series and its time evolution, the cross-wavelet transform and the cross-wavelet coherency are used to quantify the degree of linear relation between two non-stationary time series in the time–frequency domain. Phase analysis is a nonlinear technique that makes it possible to study synchronization and delays between two time series across different frequencies or timescales.

This paper’s main contribution to the literature is to clearly demonstrate the utility of wavelets and cross-wavelets for the analysis of economic time series and to illustrate how relationships between macroeconomic variables change over time and across different frequencies. In fact, wavelets allowed us to detect transient effects which would be very difficult to detect using classical econometric techniques. For example, we were able to see that the reduction in the US output and inflation volatility decreased in the 1960s at all frequencies (and not in the 1980s as is usually claimed), but that it was temporarily revived in the 1970s (especially at the business cycle frequency) probably because of the oil price shocks. We observed the same behavior for inflation rates and concluded that the “great moderation” is also a nominal phenomenon.

We were able to disentangle different short, medium and long run relations and to detail transient relations. For example, we showed that, in the 1970s and 1980s, at the business cycle frequency, inflation and interest rates were in phase, with the inflation rates leading, consistent with a central bank that follows a Taylor rule. In the long run, 12–20 year timescales, the phase difference showed that the variables are anti-phase, with interest rates leading, a result consistent with the notion that, in the long run, a restrictive monetary policy does help to control inflation.

We saw that after 1980, coinciding with Paul Volcker being a chairman of the Federal Reserve, interest rates, in the 2–4 year time band, were in phase with and reacting to industrial production, having contractionary effects in the longer run. Our results also gave some support to authors like Barsky and Kilian[1] and Leduc and Sill [30] who argued that, during the 1970s and early 1980s, monetary policy reinforced the recessionary effects of the oil shocks. About the same time, there was a structural break in the relation between interest rates and the monetary aggregates, M1 and M2. This supports and illuminates the results of Ref. [34] which found evidence for a structural break, in the early 1980s, in the relation between money and real output.

References