Dynamical Decomposition of Political Time-Series: An Application of Wavelet Analysis to Electoral Cycles in the United States Presidential Elections*

Luís Aguiar-Conraria† Pedro C. Magalhães‡ Maria Joana Soares§

August 24, 2009

1 Introduction

Political agents operate simultaneously at different time horizons. Presidents and prime-ministers may seek reelection at the end of their terms, but they may also seek many other shorter- or longer-term goals, ranging from weathering current political setbacks or successfully appoint officials to transforming things like the foreign policy profile of their countries or the social bases of support for their parties. Besides, most - if not all - political processes of any interest are the result of a combination of the actions of several agents, who have themselves different short- and long-term objectives, causing political time-series to combine components that operate on different frequencies. And in spite of the relative slowness of social, cultural and institutional change, it is probably unwise to assume, especially over prolonged periods of time, that the underlying processes generating the time series data we observe are themselves time invariant.

The almost trivial statements above raise a number of rather non-trivial problems in what concerns the analysis of time series data and the kind of assumptions made about the underlying process that generates them. Several tools in econometrics that aim at uncovering relations at different frequencies and dynamic relationships in time series data have already made their way

*We thank Bernard Cazelles for providing us the code used in Cazelles *et al.* (2007). Luís Aguiar-Conraria acknowledges financial support from Fundação para a Ciência e a Tecnologia, project "Oil shocks and the Macroeconomy: Econometric estimation, economic modeling and policy implications", PTDC/ECO/64750/2006.

†NIPE and Economics Departament, University of Minho, E-mail address: lfaguiar@eeg.uminho.pt

‡Social Science Institute, University of Lisbon, E-mail address: pedro.magalhaes@ics.ul.pt

§Mathematics Departament, University of Minho, E-mail address: jsoares@math.uminho.pt

into Political Science. Spectral analysis has been used to determine which frequencies play relevant roles in explaining the overall variance of time series, by decomposing the observed pattern over time into a spectrum of cycles of different lengths, just as a prism decomposes white light into a spectrum of colors of different wavelengths or frequencies. In Political Science, spectral analysis as been applied to the study of terrorist attacks (Enders and Sandler 2000, 2002 and 2006), wars (Beck 1991), military expenditures (Williams and McGinnis 1992 and Gerace 2002), policy agendas (Howlett 1998), government popularity (Goodhart and Bhansali 1970; Miller and Mackie 1973, Richards 1992) and election returns (Lin and Guillén 1998; Merrill, Grofman, and Brunell 2008).

However, there is an important limitation with the spectral analysis approach. First, while spectral analysis techniques are only appropriate for strictly stationary time-series, many, if not all, economic and political time-series are, in fact, noisy, complex and strongly non-stationary. Second, under the Fourier transform used for spectral analysis, the time information of a time-series becomes very difficult to uncover. Because of this loss of information, it is hard to distinguish transient relations or to identify structural changes. As Goldstein (1988, chapter 8) argues, when studying social phenomena, one should not expect to encounter cycles of perfect fixed periodicity, unlike some cycles we observe in physical processes. Goldstein argues that attempts to decompose long-run political cycles into their components using spectral analysis (based on the Fourier transform) are misleading, because they implicitly assume that these political cycles are well approximated by sine and cosine functions of fixed periodicity.

Beck (1991) responds to Goldstein's criticism of spectral analysis. Beck's argument is, in its essence, a statistical one: "any stationary series has a spectral representation, and frequency analysis is appropriate whenever ARMA analysis is adequate". Beck's argument is correct. The spectral representation theorem indeed guarantees that the analysis of stationary processes in the "frequency domain" is equivalent to "time domain" analysis, based on the autocovariance function. But Beck's argument could be reversed. Instead of arguing that if an ARMA representation of a stationary time-series is adequate then spectral representation is also adequate, one could argue that if the spectral representation is inadequate to study political time-series then the spectral representation theorem guarantees that ARMA representations are also inadequate. When we estimate some ARMA process, we are estimating some coefficients that summarize the dynamics of the time series for the chosen time period. As Lebo and Box-Steffensmeier (2008) write, "[t]here is much more we may wish to know — how does it vary over time, how volatile, and what are

the impacts of different circumstances? For example, a regression coefficient can tell us how subjective economic evaluations affect leader or party approval over a long period of data, but it is much less useful in determining how those effects may vary in the months leading up to and following an election". The development of dynamic conditional correlation (DCC) models under the ARCH/GARCH framework (Engle 2002) provided a way to estimate time-varying correlations as a function of past correlations and volatilities. Lebo and Box-Steffensmeier (2008) used DCC to address the lack of constancy of evaluations of the economy as predictors of presidential approval. The implication is that, although very useful, ARMA (and VARMA) models cannot capture all the features of the data one may be interested in. DCC is indeed an important breakthrough in addressing the kind of time-varying relationships of which spectral analysis inevitably loses sight. Still, this method (and others, like Kalman filtering) is completely silent about relationships across frequencies.

Wavelet analysis helps overcoming these problems in the analysis of the cyclical components of a time series and of the frequencies that explain its variance. It 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. As a coherent mathematical body, wavelet theory was born in the mid-1980s (Grossmann and Morlet 1984, Goupillaud *et al.* 1984).[1] After 1990, the literature rapidly expanded and wavelet analysis is now extensively used in Physics, Geophysics, Astronomy, Epidemiology, Signal Processing, Oceanography, etc. Unfortunately, and in spite of all its potential, to our knowledge, this technique has never been used in Political Science and it is rarely used in Economics. The pioneering work of Ramsey and Lampart (1998), Ramsey (1999), and Gençay *et al.* (2001) is largely unknown among social scientists, who reveal a strong preference for traditional time-series methods, overlooking the potential for using wavelets to analyze time-series data.[2]

Probably, the most compelling reason for wavelets not being more popular among us is related with the difficulty of performing multivariate analysis with it. To overcome this problem, in the

[1] See Hubbard (1996) for a very nice non-technical historic account of wavelet analysis.

[2] For a detailed review of wavelet applications to economic data, the reader is referred to Crowley (2007).

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.* (1993) and Torrence and Compo (1998) have been applied in different scientific fields, ranging from Medicine (Zhan *et al.* 2006) and Epidemiology (Cazelles *et al.* 2007) to Astrophysics (Kelly *et al.* 2003, and Bloomfield *et al.* 2004) and Geophysics (Jevrejeva *et al.* 2003 and Grinsted *et al.* 2004). Gallegati (2008) and Aguiar-Conraria *et al.* (2008) showed that cross-wavelet analysis could be also be applied to study pairs of Economic time-series. While the (single) wavelet power spectrum describes the evolution of the volatility 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 two variables in the time-frequency space. One can also look at the wavelet coherency as a localized correlation coefficient in the time-frequency space. 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.

The paper proceeds as follows. In section 2, we present the continuous wavelet transform, discuss its localization properties and the optimal characteristics of the Morlet wavelet. We also describe the wavelet power spectrum, the cross-wavelet power spectrum, the wavelet coherency and the phase-difference and illustrate the usefulness of these tools using simulated data. In section 3, we propose a new metric that will allow us to estimate a dissimilarity matrix between the electoral cycles of different states. In section 4, we apply these tools to study presidential electoral cycles in and across the United States. Section 5 concludes.

2 Wavelets: Frequency Analysis Across Time

To overcome the problems of analyzing non-stationary data, Gabor (1946) 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, this approach is inefficient because the frequency resolution is the same across all different frequencies. 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. It is true that the Heisenberg uncertainty principle proves that there will always be 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 practical purposes, corresponds to an effective band and time limiting; see Daubechies (1992).

2.1 The Wavelet

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 $\|x\| := \int_{-\infty}^{\infty} |x(t)|^2 dt < \infty$, with the usual inner product, $\langle x, y\rangle := \int_{-\infty}^{\infty} x(t) y^*(t) dt$. The asterisk superscript denotes complex conjugation. Given a function $x(t) \in L^2(\mathbb{R})$, $X(f) := \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} dt$ will denote the Fourier transform of $x(t)$. 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 immediately follows: $\|x(t)\|^2 = \|X(f)\|^2$. 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:

$$0 < C_\psi := \int_{-\infty}^{\infty} \frac{|\Psi(f)|}{|f|} df < \infty, \tag{1}$$

where $\Psi(f)$ is the Fourier transform of $\psi(t)$, see Daubechies (1992, p. 24).

The wavelet ψ is usually normalized to have unit energy: $\|\psi\|^2 = \int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$. 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 or even compact support. For functions with sufficient decay it turns out that the admissibility condition (1) is equivalent to requiring $\Psi(0) = \int_{-\infty}^{\infty} \psi(t) dt = 0$. 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

$$\psi_{s,\tau}(t) := \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right), \quad s,\tau \in \mathbb{R}, s \neq 0. \tag{2}$$

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, $\left\|\psi_{s,\tau}\right\| = 1$) 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 $|s| < 1$).[3]

Given a function $x(t) \in L^2(\mathbb{R})$ (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 $\left\{\psi_{s,\tau}\right\}$:

$$W_x(s,\tau) = \left\langle x, \psi_{s,\tau}\right\rangle = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{|s|}}\psi^*\left(\frac{t-\tau}{s}\right)dt. \tag{3}$$

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 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. It is also convenient to choose $\psi(t)$ to be *progressive* or *analytic*, i.e. to be such that $\Psi(f) = 0$ for $f < 0$.[4]

The importance of the admissibility condition (1) comes from the fact that it guarantees that it is possible to recover $x(t)$from its wavelet transform. When ψ is analytic, if $x(t)$ is real,[5] the reconstruction formula is given by

$$x(t) = \frac{2}{C_\psi}\int_0^{\infty}\left[\int_{-\infty}^{\infty}\mathcal{R}\left(W_x(s,\tau)\psi_{s,\tau}(t)\right)d\tau\right]\frac{ds}{s^2}. \tag{4}$$

[3] Note that for negative s, the function is also reflected.
[4] Note that an analytic function is necessarily complex.
[5] See Aguiar-Conraria *et al.* (2008) for the case of complex $x(t)$.

Therefore, we can easily go from $x(t)$to its wavelet transform, and from the wavelet transform back to $x(t)$. Note that one can limit the integration over a range of scales, performing a band-pass filtering of the original series. See Daubechies (1992, pp. 27-28) or Kaiser (1994, pp. 70-73) for more details about analytic wavelets.

2.3 Localization Properties

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

$$\mu_t = \int_{-\infty}^{\infty} t \left|\psi(t)\right|^2 dt. \tag{5}$$

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

$$\sigma_t = \left\{ \int_{-\infty}^{\infty} (t-\mu_t)^2 \left|\psi(t)\right|^2 dt \right\}^{\frac{1}{2}}. \tag{6}$$

In a total similar manner, one can also define the center μ_f and standard deviation 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 $\Psi(f)$. 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 Heisenberg 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 $\sigma_t \sigma_f$.

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 tells us 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. To be more precise, the Heisenberg uncertainty principle establishes that the uncertainty is bounded from below by the quantity $1/4\pi$:

$$\sigma_t \sigma_f \geq \frac{1}{4\pi}. \tag{7}$$

If the mother wavelet ψ is centered at μ_t, has standard deviation σ_t and its wavelet transform $\Psi(f)$ is centered at μ_f with a standard deviation σ_f, then one can easily show that the daughter wavelet $\psi_{\tau,s}$ will be centered at $\tau + s\mu_t$ with standard deviation $s\sigma_t$, whilst its Fourier transform $\Psi_{s,\tau}$ will have center $\frac{\mu_f}{s}$ and standard deviation $\frac{\sigma_f}{s}$.

From the Parseval relation, we know that $W_x(s,\tau) = \langle x(t), \psi_{s,\tau}(t) \rangle = \langle X(f), \Psi_{s,\tau}(f) \rangle$. Therefore, the continuous wavelet transform $W_x(s,\tau)$ gives us local information within a time-frequency window $[\tau + s\mu_t - s\sigma_t, \tau + s\mu_t + s\sigma_t] \times \left[\frac{\mu_f}{s} - \frac{\sigma_f}{s}, \frac{\mu_f}{s} + \frac{\sigma_f}{s}\right]$. In particular, if ψ is chosen so that $\mu_t = 0$ and $\mu_f = 1$, then the window associated with $\psi_{\tau,s}$ becomes

$$[\tau - s\sigma_t, \tau + s\sigma_t] \times \left[\frac{1}{s} - \frac{\sigma_f}{s}, \frac{1}{s} + \frac{\sigma_f}{s}\right] \tag{8}$$

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 $s\sigma_t$, and information about $X(f)$ for frequency values near the frequency $f = \frac{1}{s}$, with precision $\frac{\sigma_f}{s}$. Therefore, small/large values of s correspond to information about $x(t)$ in a fine/broad scale and, even with a constant area of the windows, $A = 4\sigma_t\sigma_f$, 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 $f = 1/s$). This is one major advantages afforded by the wavelet transform, when compared to the Short Time Fourier Transform: its 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.

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. 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 one has in mind. We choose a complex wavelet, as it yields a complex transform, with information on both the amplitude and phase, important to study cycles synchronism between different time-series.

We will use the Morlet wavelet, introduced in Goupillaud *et al.* (1984):

$$\psi_\eta(t) = \pi^{-\frac{1}{4}} \left(e^{i\eta t} - e^{-\frac{\eta^2}{2}} \right) e^{-\frac{t^2}{2}}. \tag{9}$$

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

$$\psi_\eta(t) = \pi^{-\frac{1}{4}} e^{i\eta t} e^{-\frac{t^2}{2}} \tag{10}$$

of (9) is normally used (and still referred to as a Morlet wavelet). Our results in the next section, were obtained with the particular choice $\eta = 6$.

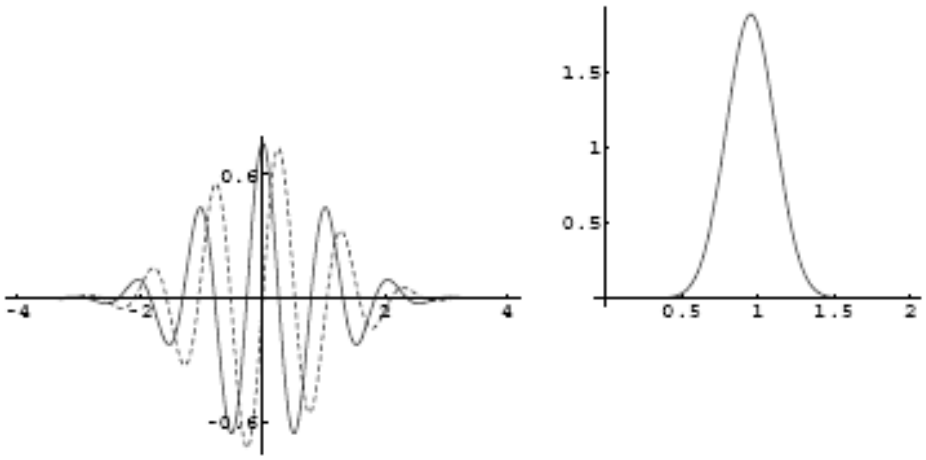


Figure 1: On the left: the Morlet wavelet $\psi_6(t)$ — real part (solid line) and imaginary part (dashed line). On the right: its Fourier transform.

This wavelet has interesting characteristics. First of all, for $\eta > 5$, for all practical purposes, the wavelet can be considered as analytic; see Foufoula-Georgiou and Kumar (1994).

The wavelet (10) is centered at the point $\left(0, \frac{\eta}{2\pi}\right)$ of the time-frequency plane; hence, for the particular choice $\eta = 6$, one has that the frequency center is $\mu_f = \frac{6}{2\pi}$and the relationship between the scale and frequency is simply $f = \frac{\mu_f}{s} \approx \frac{1}{s}$. Therefore there is biunivocal relation between scale and frequency and we will use both terms interchangeably

It is simple to verify that the time standard deviation is $\sigma_t = 1/\sqrt{2}$ and the frequency standard deviation is $\sigma_f = 1/\left(2\pi\sqrt{2}\right)$. Therefore, the uncertainty of the corresponding Heisenberg box attains the minimum possible value $\sigma_t \sigma_f = \frac{1}{4\pi}$. In this sense, the Morlet wavelet has optimal joint

time-frequency concentration.

2.5 Wavelet Power Spectrum

In view of the energy preservation formula, and in analogy with the terminology used in the Fourier case, we simply define the (local) wavelet power spectrum as $S_x(s,\tau) = |W_x(s,\tau)|^2$, which gives us a measure of the local variance. Torrence and Compo (1998) showed how the statistical significance of wavelet power can be assessed against the null hypothesis that the data generating process is given by an $AR(0)$ stationary process with a certain background power spectrum (P_f):

$$D\left(\frac{|W_x(s,\tau)|^2}{\sigma_x^2} < p\right) = \frac{1}{2} P_f \chi_v^2(p), \tag{11}$$

at each time τ 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 $s \approx \frac{1}{f}$ — 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.

2.5.1 Example

As we have argued before, the main advantage of wavelet analysis over spectral analysis is the possibility of tracing transitional changes across time. To illustrate this, consider the following experiment with simulated data. We generate 50 years of monthly data according to the following data generating process:

$$\begin{aligned} y_t &= \cos\left(\frac{2\pi}{p_1}t\right) + \cos\left(\frac{2\pi}{p_2}t\right) + \varepsilon_t, \qquad t = \frac{1}{12}, \frac{2}{12}, \cdots, 50 \\ &\text{with} \quad \begin{cases} p_1 = 5 & \text{if } 20 \le t < 30 \\ p_1 = 3 & \text{otherwise} \end{cases} \quad \text{and} \quad p_2 = 10 \end{aligned} \tag{12}$$

Formula (12) tells us that the time series y_t is the sum of two periodic components and a white noise.[6] The second period component represents a 10 year cycle while the first periodic component shows some transient dynamics. In the beginning, it represents a 3 year cycle that, temporarily,

[6] This formulation is not as restrictive as it may seem. An autoregressive process of order 2, or higher, with an oscillatory behavior, will have a solution that involves sines and cosines. We, therefore, could have generated similar time series using a more common auto regressive process. We chose to explicitly have a cosine because the period of the oscillation is observed directly.

changes to a 5 year cycle between the second end the third decades.

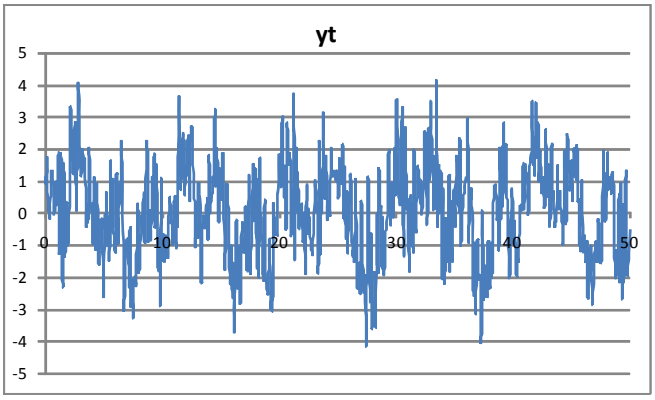


Figure 2: $y_t = \cos\left(\frac{2\pi}{p_1}t\right) + \cos\left(\frac{2\pi}{p_2}t\right) + \varepsilon_t$

This change in the dynamics is very difficult to spot in Figure 2. If we use traditional the traditional spectral analysis, this information on the transient dynamics is completely lost, as we can see in Figure 3.c. The power spectral density estimate is able to capture the 3 and the 10 year cycles but it completely fails to capture the 5 year cycle that occurred for a decade. Comparing with Figure 3.b, we observe that spectral analysis gives us essentially the same information as the global wavelet power spectrum, which is an average, across time, of the wavelet power spectrum.

On the other hand, Figure 3.a shows the wavelet power spectrum itself. There we can see the biggest advantage of wavelet analysis over spectral analysis. With wavelets we are able to estimate the power spectrum as a function of time and, therefore, we do not lose the time dimension. The wavelet power spectrum is able to capture not only the 3 and 10 year cycles, but also to capture the change that occurred between years 20 and 30.

2.6 Cross-Wavelets and Phase-Differences

2.6.1 Cross-Wavelet Power

The cross-wavelet transform of two time-series, $x(t)$ and $y(t)$, first introduced by Hudgins *et al.* (1993), is simply defined as

$$W_{xy}(s,\tau) = W_x(s,\tau)\, W_y^*(s,\tau), \tag{13}$$

where W_x and W_y are the wavelet transforms of x and y, respectively. The cross-wavelet power is given by $|W_{xy}|$. While we can interpret the wavelet power spectrum as depicting the local variance

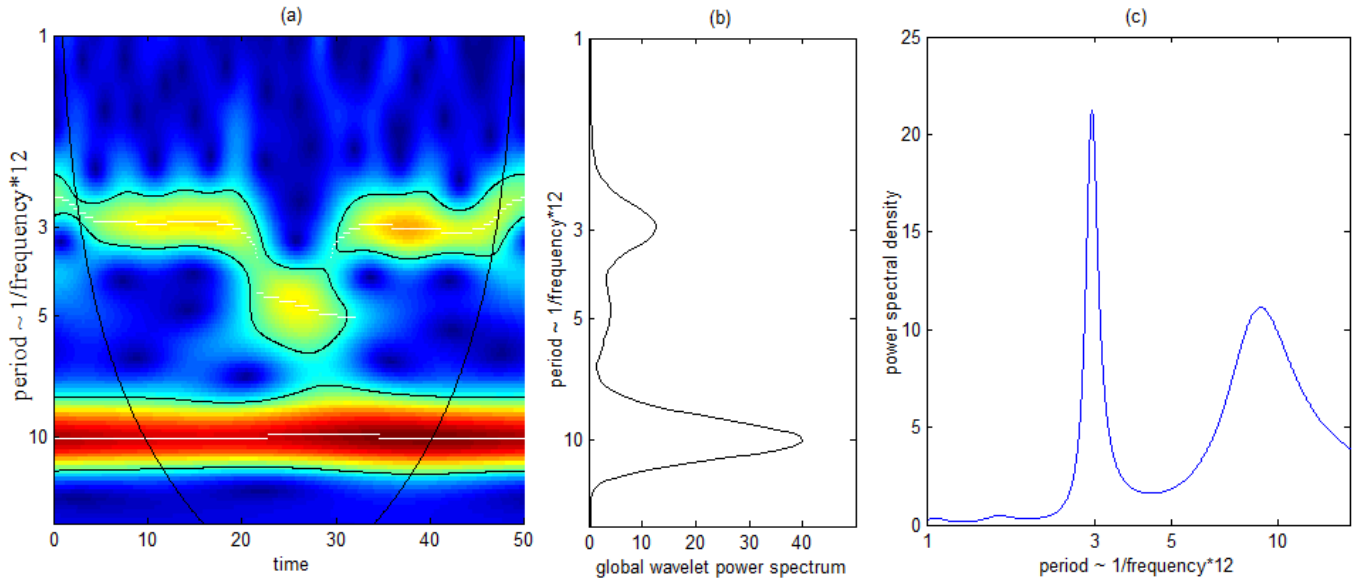


Figure 3: **(a) Wavelet Power Spectrum** — The black contour designates the 5% significance level based on a AR(1) null. The cone of influence, which indicates the region affected by edge effects, is shown with a thin black line. The color code for power ranges from blue (low power) to red (high power). The white lines show the maxima of the undulations of the wavelet power spectrum.
(b) Global Wavelet Power Spectrum — average wavelet power for each frequency.
(c) Power Spectral Density — The power spectral density is calculated using a parametric spectral estimation method, which implements the Yule-Walker algorithm.

of a time-series, the cross-wavelet power of two time-series depicts the local covariance between these time-series at each scale and frequency. Therefore, cross-wavelet power gives us a quantified indication of the similarity of power between two time-series. Torrence and Compo (1998) also derived the cross-wavelet distribution assuming that the two time-series have Fourier Spectra P_f^x and P_f^y.Under the null, the cross-wavelet distribution is given by

$$D\left(\frac{|W_x W_y^*|}{\sigma_x \sigma_y} < p\right) = \frac{Z_v(p)}{v}\sqrt{P_f^x\ P_f^y} \tag{14}$$

,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 χ^2 distributions. For more general data generating processes one has to rely on Monte Carlo simulations.

2.6.2 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. The wavelet coherency between two time-series, $x(t)$ and $y(t)$, is defined as follows:

$$R_{xy}\left(s,\tau\right)=\frac{\left|S\left(W_{xy}\left(s,\tau\right)\right)\right|}{\left|S\left(W_{xx}(s,\tau)\right)\right|^{\frac{1}{2}}\left|S\left(W_{yy}(s,\tau)\right)\right|^{\frac{1}{2}}}, \tag{15}$$

where S denotes a smoothing operator in both time and scale. Smoothing is necessary. Without that step, coherency is identically one 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 by a rectangular window; see Cazelles *et al.* (2007) 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.

2.6.3 Phase Difference and the Instantaneous Time Lag

The phase-difference gives us information about the delays of the oscillations between two time-series, $x(t)$ and $y(t)$, as a function of time and frequency. As we said before, the phase of a given time-series, ϕ_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

$$\phi_{x,y}(s,\tau)=\tan^{-1}\left(\frac{\mathcal{I}\left\{W_{xy}(s,\tau)\right\}}{\mathcal{R}\left\{W_{xy}(s,\tau)\right\}}\right),\qquad \text{with}\ \ \phi_{x,y}\in\left[-\pi,\pi\right]. \tag{16}$$

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

With the phase difference one can calculate the instantaneous time lag between the two time-series:

$$\Delta T\left(s,\tau\right)=\frac{\phi_{x,y}\left(s,\tau\right)}{2\pi F\left(\tau\right)}, \tag{17}$$

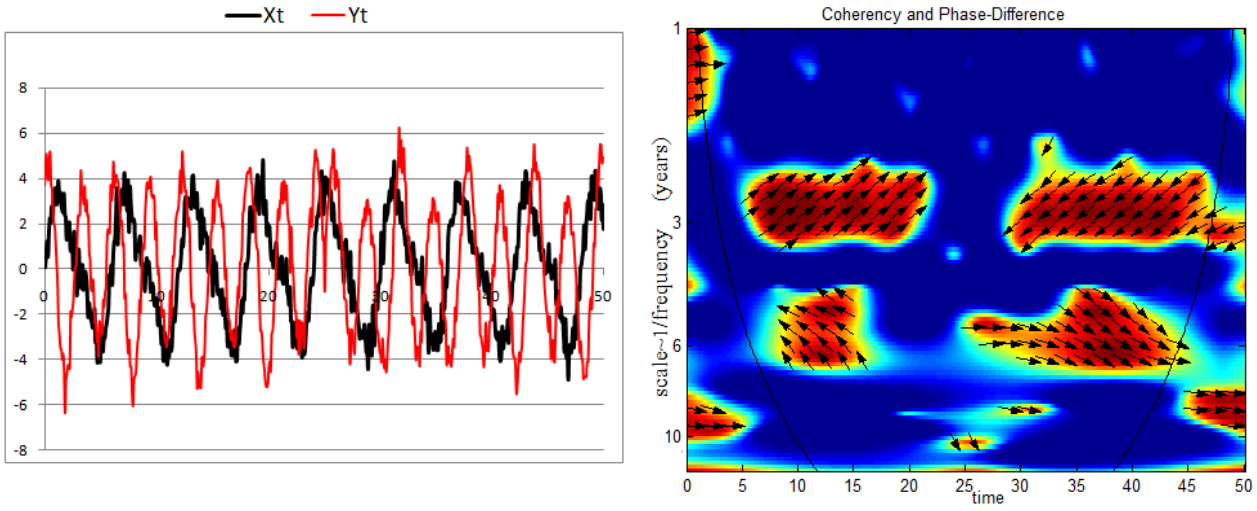


Figure 4: **Cross-Wavelet Coherency** — The color code for coherency ranges from blue (low coherency – close to zero) to red (high coherency – close to one). The cone of influence, which indicates the region affected by edge effects is shown with a thin black line.
Phase-Difference — the phase-difference 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 x_t lagging. To the right and down, with x_t leading. Arrows pointing to the left mean that the variables are out of phase. To the left and up, with x_t leading. To the left and down, with x_t lagging.

where $F(\tau)$ is the instantaneous frequency defined in a given frequency band as the first normalized moment in frequency of W_{xy}:

$$F(\tau) = \frac{\int_{f_1}^{f_2} f \left|W_{xy}(f,\tau)\right| df}{\int_{f_1}^{f_2} \left|W_{xy}(f,\tau)\right| df} \tag{18}$$

2.6.4 Example

Again, we illustrate these concepts with the aid of a numerical example. As before, we also show how wavelet analysis is suitable to identify changes in the dynamic relations of two time series. Consider two time series that share two common cycles, with some delays. Again, this process can be seen as the solution to some VAR.

$$x_t = \sin\left(\frac{2\pi}{3}t\right) + 3\sin\left(\frac{2\pi}{6}t\right) + \varepsilon_{x,t}, \qquad t = 0, \frac{1}{12}, \frac{2}{12}, \cdots, 50 \tag{19}$$

$$y_t = \begin{cases} 4\sin\left(\frac{2\pi}{3}\left(t+\frac{4.5}{12}\right)\right) - \sin\left(\frac{2\pi}{6}\left(t-\frac{9}{12}\right)\right) + \varepsilon_{y,t}, & \text{for } t = 0, \frac{1}{12}, \frac{2}{12}, \cdots, 25 \\ -4\sin\left(\frac{2\pi}{3}\left(t+\frac{4.5}{12}\right)\right) + \sin\left(\frac{2\pi}{6}\left(t-\frac{9}{12}\right)\right) + \varepsilon_{y,t}, & \text{for } t = 25+\frac{1}{12}, 25+\frac{2}{12}, \cdots, 50 \end{cases} \tag{20}$$

Looking at the formulas, it is clear that x_t and y_t share a 3 and a 6 year cycles. However,

these changes evolve with time and they are different from cycle to cycle. Consider the first cycle, the 3 year cycle. The y_t cycle precedes the x_t cycle by 4 and a half months. For the first half of the sample, the relation between them is positive, while, in the second half, it becomes negative. These features are captured in Figure 4. That both series have a strong 3 year cycle is revealed by the regions of strong coherency at that frequency. That in the first half of the sample the series are in phase is shown by the arrows pointing to the right. That in the second half of the sample the series are out of phase is shown by the arrows point to the left. The direction of the arrows are given by the angle provided by equation 16 (the phase difference). Arrows pointing upwards and to the right (angle between 0 and $\frac{\pi}{2}$) show that the series are in phase, with y_t leading; while arrows pointing downwards and to the left (angle between $-\pi$ and $-\frac{\pi}{2}$)show that the series are out of phase, with y_t leading.

Looking at the 6 year cycle, we observe that in the first half the series are out of phase, and that they are in phase in the second half. x_t leads y_t throughout the sample.

2.7 Transform of finite discrete data

If one is dealing with a discrete time-series $x = \{x_n,\ n = 0, \ldots, T-1\}$ of T observations with a uniform time step δt, the integral in (3) has to be discretized and is, therefore, replaced by a summation over the T time steps; also, it is convenient, for computational efficiency, to compute the transform for T values of the parameter τ, $\tau = m\delta t;\ m = 0, \ldots, T-1$. In practice, naturally, the wavelet transform is computed only for a selected set of scale values $s \in \{s_k, k = 0, \ldots, F-1\}$ (corresponding to a certain choice of frequencies f_k). Hence, our computed wavelet spectrum of the discrete-time series x will simply be a $F \times T$ matrix W_x whose (k, m) element is given by

$$W_x(k,m) = \frac{\delta t}{\sqrt{s_k}} \sum_{n=0}^{T-1} x_n \psi^* \left((n-m) \frac{\delta t}{s_k} \right) \quad k = 0, \ldots, F-1, \quad m = 0, \ldots, T-1. \tag{21}$$

Although it is possible to calculate the wavelet transform using the above formula for each value of k 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 (1998).[7]

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 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 $\sqrt{2}s$.

3 Wavelet Spectra Distance Matrix

To estimate a dissimilarity matrix between two variables, e.g. the electoral cycles of different states, based on their wavelet spectra, one has to find a metric to measure the distances between two wavelet spectra. Comparing time-series based on their wavelet spectra is, in a sense, like comparing two images. Direct comparison is not suitable because there is no guarantee that regions of low power will not overshadow the comparison. It would be like comparing two pencil-drawing sketches based mainly on the color of the paper, disregarding the sketches themselves. We build on the work of Rouyer *et al.* (2008) and use the singular value decomposition (SVD) of a matrix to focus on the common high power time-frequency regions. This method is similar to Principal Component Analysis, but while with the latter finds linear combinations that maximize the variance, subject to some orthogonality conditions, the method we use extracts the components that maximize covariances instead, subject to similar orthogonality conditions. Therefore, the first extracted components correspond to the most important common patterns between the two wavelet spectra. With that information, and after defining a metric to measure the pairwise distance between the several extracted components, we can build a dissimilarity matrix between the several analyzed countries. At that point, it is easy to implement clustering and multidimensional mapping

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

algorithms to visualize the estimated distances.

Consider the covariance matrix $C_{xy} := W_x W_y^H$, where W_y^H is the conjugate transpose, also known as the Hermitian transpose, of W_y. Its SVD decomposition yields

$$C_{xy} = U \Sigma V^H,$$

where the matrices U and V are unitary matrices (i.e. $U^H U = V^H V = I$), whose columns, $\mathbf{u}_k$ and $\mathbf{v}_k$ are, respectively, the singular vectors for W_x and W_y, and Σ is a diagonal matrix with the singular values ordered from highest to lowest, $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_F \geq 0$. The number of nonzero singular values is equal to the rank of the matrix C_{xy}. The SVD of C_{xy} guarantees that the singular vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ solve the problem of maximizing

$$\mathbf{p}_k^H C_{xy} \mathbf{q_k} = \mathbf{p}_k^H W_x W_y^H \mathbf{q_k} = \mathbf{p}_k^H W_x \left(\mathbf{q}_k^H W_y\right)^H$$

for all vectors $\mathbf{p}_k$ and $\mathbf{p}_k$ satisfying the orthogonality constraints $\mathbf{p}_k^H \mathbf{p}_j = \delta_{kj}$, $\mathbf{q}_k^H \mathbf{q}_j = \delta_{kj}$, $j = 1, \ldots, k$, where δ_{kj} is the Kronecker delta. In other words, the so-called *leading patterns*, obtained by projecting each spectrum W_x and W_y onto the respective singular vectors,

$$L_x^k := \mathbf{u}_k^H W_x \quad \text{and} \quad L_y^k := \mathbf{v}_k^H W_y, \tag{22}$$

are the linear combinations of the columns of W_x and W_y, respectively, that maximize their mutual covariance (subject to the referred orthogonality constraints). Moreover, since $U^H W_x W_y V = U^H C_{xy} V = \Sigma$, it follows immediately that the (squared) covariance of the k^{th} leading patterns is given by

$$|L_x^k \, (L_y^k)^H|^2 = \sigma_k^2.$$

On the other hand, the (squared) covariance of W_x and W_y is given by $\|C_{xy}\|_F^2$, where $\|.\|_F$ is the Frobenius matrix norm, defined by $\|A\|_F := \sqrt{\sum_{ij} |a_{ij}|^2}$. But, since this norm is invariant under a unitary transformation, we have

$$\|C_{xy}\|_F^2 = \|U^H C_{xy} V\|_F^2 = \|\Sigma\|_F^2 = \sum_{i=1}^{F} \sigma_i^2.$$

The (squared) singular values, σ_k^2, are the weights to be attributed to each leading pattern and are

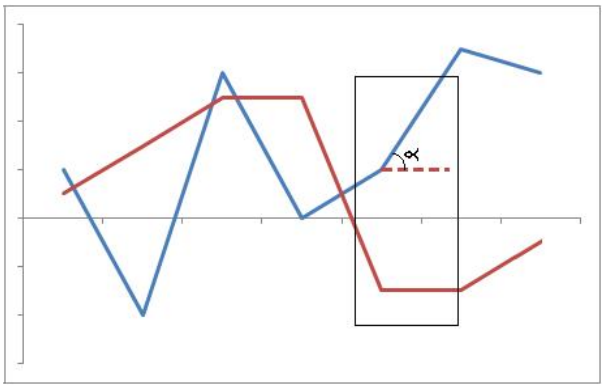


Figure 5: Angles between real vectors.

equal to the (squared) covariance explained by each pair of singular vectors (axis).

If we denote by L_x and L_y the matrices whose rows are the leading patterns L_x^k and L_y^k, equation(22) shows that $L_x = U^H W_x$ and $L_y = V^H W_y$, from where we immediately obtain

$$W_x = U L_x = \sum_{k=1}^{F} \mathbf{u}_k L_x^k, \qquad W_y = V L_y = \sum_{k=1}^{F} \mathbf{v}_k L_y^k.$$

In practice, we select a certain number $K < F$ of leading patterns, guaranteeing, for example, that the fraction of covariance $\left(\sum_{k=1}^{K} \sigma_k^2\right) / \left(\sum_{k=1}^{F} \sigma_k^2\right)$ is above a certain threshold, and use

$$W_x \approx \sum_{k=1}^{K} \mathbf{u}_k L_x^k, \quad W_y \approx \sum_{k=1}^{K} \mathbf{v}_k L_y^k.$$

What we have done so far is to reduce the information contained in the two wavelet spectra to a few components. Now we need to find a metric to measure the distance between the most relevant components associated to the different wavelet spectra. We need to measure the distance between the leading patterns, L_x^k and L_y^k, and between the singular vectors, $\mathbf{u}_k$ and $\mathbf{v}_k$. To do so, we compare two vectors by measuring the angle between each pair of corresponding segments. This would be straightforward if all values were real (see Figure 5).

In our case, because we use a complex wavelet, we need to define an angle in a complex vector space. Unfortunately, very little guidance is available in the mathematical literature on angles in complex vector spaces and there are several possibilities, see Scharnhorst (2001).

Recall that, given two vectors $\mathbf{a}$ and $\mathbf{b}$ in the Euclidian vector space $\mathbb{R}^n$, with the usual inner product $\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{R}} = \mathbf{a}^T \mathbf{b}$ and norm $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle_{\mathbb{R}}}$, the angle between the two vectors, $\Theta = \Theta(\mathbf{a}, \mathbf{b})$,

can be found using the formula:

$$\cos(\Theta) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{R}}}{\|\mathbf{a}\|\|\mathbf{b}\|}, \quad \Theta \in [0, \pi]. \tag{23}$$

Now, assume that $\mathbf{a}$ and $\mathbf{b}$ are vectors in the vector space $\mathbb{C}^n$. There are two reasonable approaches to define a (real)-valued angle between $\mathbf{a}$ and $\mathbf{b}$. The first one is to consider the isomorphism

$$\begin{aligned} \phi : \mathbb{C}^n &\longrightarrow \mathbb{R}^{2n} \\ \mathbf{a} = (a_1, \ldots, a_n) &\mapsto \mathcal{R}(a_1), \mathcal{I}(a_1), \cdots \mathcal{R}(a_n), \mathcal{I}(a_n) \end{aligned}$$

and simply define the *Euclidean* angle between the complex vectors $\mathbf{a}$ and $\mathbf{b}$ as the angle (defined by using formula (23)) between the real vectors $\phi(\mathbf{a})$ and $\phi(\mathbf{b})$.

The other approach is based on the use of the Hermitian inner product $\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{C}} = \mathbf{a}^H \mathbf{b}$ and corresponding norm $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle_{\mathbb{C}}}$. We can then define the so-called *Hermitian angle* between the complex vectors $\mathbf{a}$ and $\mathbf{b}$, $\Theta_H(\mathbf{a}, \mathbf{b})$, by the formula

$$\cos(\Theta_H) = \frac{|\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{C}}|}{\|\mathbf{a}\|\|\mathbf{b}\|}, \quad \Theta_H \in \left[0, \frac{\pi}{2}\right]. \tag{24}$$

The measures are not equal, but they are related; see Scharnhorst (2001) for details. In what follows, we make use of the Hermitian angle.

The distance between the k^{th} leading patterns L_x^k and L_y^k is computed as:

$$d\left(L_x^k \ , \ L_y^k\right) = \frac{1}{T-1} \sum_{n=1}^{T-1} \Theta_H \left(\mathbf{l}_x^k(n) \ , \ \mathbf{l}_y^k(n)\right), \tag{25}$$

where $\mathbf{l}_x^k(n)$ is the two-component vector defined by the two "points"in $\mathbb{R} \times \mathbb{C}$, $P_n = (n, L_x^k(n))$ and $P_{n+1} = ((n+1), L_x^k(n+1))$, i.e. $\mathbf{l}_x^k(n) = (1, L_x^k(n+1) - L_x^k(n+1))$, where $L_x^k(n)$ denotes the n^{th} component of L_x^k. The distance between the singular vectors, $d(\mathbf{u}_k, \mathbf{v}_k)$ is defined in an analogous way.

To compare the wavelet spectra of country x and country y, we compute the following distance:

$$\text{dist}(W_x, W_y) = \frac{\sum_{k=1}^{K} \sigma_k^2 \left[d\left(L_x^k, L_y^k\right) + d(\mathbf{u}_k, \mathbf{v}_k)\right]}{\sum_{k=1}^{K} \sigma_k^2}, \tag{26}$$

where σ_k^2 are the weights equal to the squared covariance explained by each axis. This distance is computed for each pair of countries. With this information, we can fill a matrix of distances.

4 Two applications

4.1 Are there cycles in the American presidential election?

Are there partisan cycles in American politics? If so, for how long do they typically last? And has this cyclicality remained unchanged throughout modern history?

The notion that there is some sort of pendularity in the partisan control of the Presidency and Congress was first taken up by historians, who posed an alternation between periods of political and ideological dominance of liberal views and actors and periods of conservative dominance (Adams 1918; Schlesinger 1948). This basic notion was later seized upon by the electoral realignment literature, which claimed that certain elections (described as "critical" or "realigning") bring about sweeping changes in voters preferences and that such elections occur with a fixed periodicity (Burnham 1967 and 1970).

However, as the very concept of "realignment" and "realigning elections" came under severe criticism (Carmines and Stimson 1989; Shafer 1991; Mayhew 2000 and 2002), so did the notion that these cycles might even exist. Mayhew, in his discussion of the realignment literature, argues that a simple dichotomization of types of elections between "realigning" and "nonrealigning" is simply not borne out by the data and, as a result, "their periodicity is [also] in question" (Mayhew 2002: 60). Furthermore, there is evidence that, from the point of "party success" (i.e., victory in the presidential election, either in popular or electoral votes), the patterns seem "largely random from election to election" (Gans 1985: 236).

However, as Merrill, Grofman, and Brunell (2008) put it, deriving from the finding that elections cannot be clearly dichotomized as "realigning" and "non realigning" that cycles in aggregate voting behavior and election returns are necessarily absent is a stretch and there is the risk of "throwing out the baby with the bathwater" (2008:15). In fact, and quite interestingly, one of the studies that Mayhew uses to support the notion that elections cannot be sorted out in two types (Bartels 1998) is also clear in establishing the existence of the "long-term equilibrating movements" first detected by Stokes and Iversen (1962), which produce "fairly regular alternations between Republican and Democratic possession of the White House" (Bartels 1998: 293).

What might be, therefore, the underlying cyclicality of those alternations? Here, in spite of different methodological approaches and different time ranges used, the results have not been dramatically different from each other. Using data from 1860 to 1980 and a time-domain analysis, Midlarsky (1984) detects 28-year intervals in the restoration points for Republican control of the presidency. In the context of election forecasting and using an autoregressive model, Norpoth (1995 and 2002) estimates the average length of the presidential vote staying above 50% as about 2.5 terms in office, i.e., a 20-year cycle. Finally, Lin and Guillén (1998), first, and Merrill, et al. (2008), later, using spectral analysis, detect a 26-years cycle for presidential elections. These results are interestingly similar to the cycle-lengths detected early on by historians (Schlesinger 1948), but rather different from the longer cycles assumed by realignment scholars (Burhnam 1967), where ascendancy of a particular party was argued to last for at least twice as long.

One major shortcoming of all these approaches, however, is their inability to determine whether these periodic components of the presidential vote can be said to explain well the variance of the time series throughout the entire time-span under examination or, instead, if certain frequencies are peculiar to particular time-periods. This is where, precisely, the main advantage of wavelet analysis over spectral analysis comes into play: the possibility of tracing transitional changes across time. We can illustrate this advantage with a replication of the analysis in Merrill et al. (2008). One of the series that they use is the share of democratic vote in the presidential elections between 1856 and 2004 – see Figure 6 (we added the 2008 election).

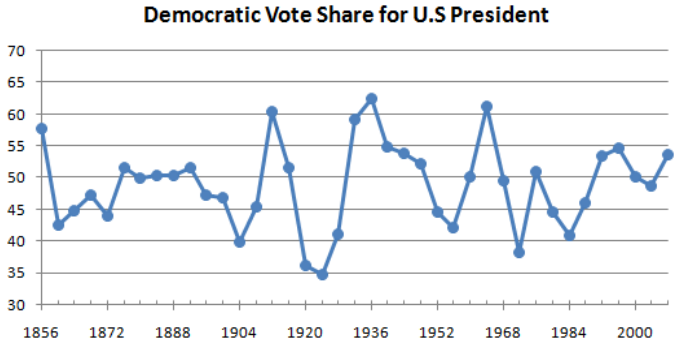


Figure 6: Democratic Vote Share for U.S President (1856 - 2008)

Merrill et al. (2008) estimate the power spectral density and identify a 26 year cycle given by the peak of the estimated power spectrum (see Figure 7.a). On the other hand, Figure 7.b gives us the wavelet power spectrum.

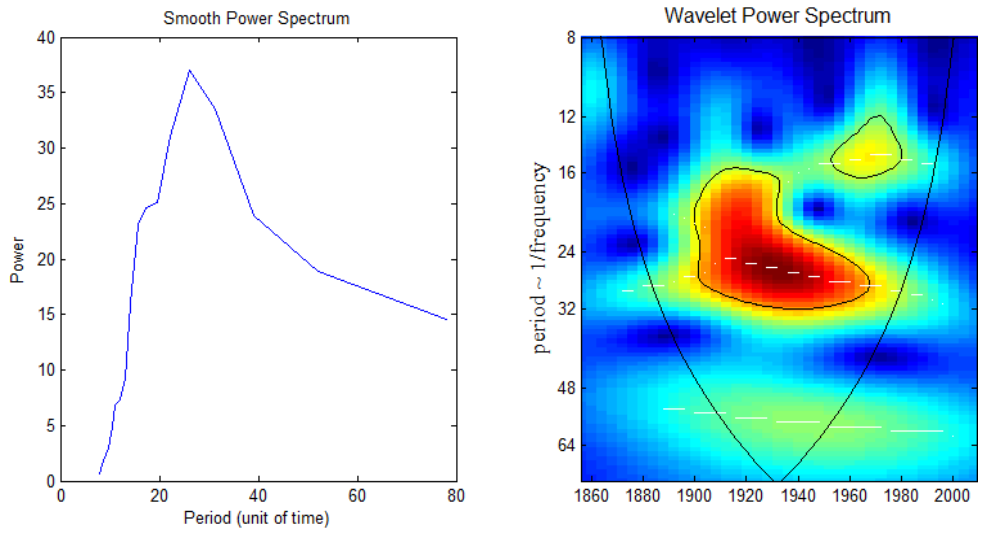


Figure 7: **(a) Power Spectral Density** of the Democratic Vote Share for Presidential elections .
(b) Wavelet Power Spectrum — The black contour designates the 5% significance level based on a AR(1) null. The cone of influence, which indicates the region affected by edge effects, is shown with a thin black line. The color code for power ranges from blue (low power) to red (high power). The white lines show the maxima of the undulations of the wavelet power spectrum

Two new different pieces of information result from this analysis. First, the wavelet spectrum identifies not only the 26-year cycle, but also a transitional 15-year cycle between the late 1950s and 1980. And although not statistically significant, it is also possible to identify a very long cycle of 60 years, basically equivalent to the one argued for in the realignment literature. Although that latter cycle is not statistically significant, the very fact that with 150 years of data one is able to see some (light) evidence of a 60-year cycle is quite striking, given that not even three full cycles can fit in the data.

Second, the statistically significant evidence for the strongest cycle – the 26-year one – is, in our wavelet spectrum, temporally localized. It starts in the turn of the 20th century but dissipates by the end of the 1960s. By then, it seems to have been replaced by a shorter 15-year cycle, where the electoral results of a particular party seem to improve (or worsen) for slightly less than two consecutive terms. These time references should not be strange to students of electoral realignment. While the 1894-96 elections mark one of the canonical moments for a change in the American party system, the end of the 1960s also marks the beginning of an age of more balanced and nationally competitive party system in presidential elections (Campbell 2006), a phenomenon captured in that 15-years cycle of fundamental alternation between the parties of 2nd term presidents and the

opposite parties. In other words, with wavelet analysis, we obtain a much more nuanced view of the cyclicality of American electoral politics, and one that is cogently connected to real world events that seem to represent structural breaks in the data.

4.2 Electoral Cycles Synchronism: how united are the United States of America?

Several studies of electoral realignments have already shown that the occurrence of critical elections and their enduring consequences cannot be described as uniform national events. "There is simply too much geopolitical diversity in the United States to justify such an expectation" (Nardulli 1995: 11). However, as Brown and Bruce argue (2008: 585), "most efforts tend to focus overwhelmingly on party competition at one level, either national or state, to the exclusion of the other".

As explained before, our entire discussion about the existence of cycles in presidential election results is mostly agnostic about the existence of "realigning elections" or "critical realignments" in the sense that the literature has employed these terms. However, the same sort of question about the national and subnational nature of realignments can be posed in what concerns the periodicity of presidential election results: can the cycles we have detected in the previous section be said to be common to all American states, or not?

We analyze the cycles in electoral behavior across the United States using State specific data on American Presidential elections since 1896 until 2008. We have data for 45 states (see appendix). For each state, we consider the votes on the Republican and the Democratic candidate and compute the democratic share.[8] All data was provided by the The American Presidency Project at UC Santa Barbara (http://www.presidency.ucsb.edu).

Based on formula (26) (multiplied by 100) we compute a pairwise dissimilarity index. In Table 1, we have the dissimilarity between each state's electoral cycle and the national cycle, which can be interpreted as the core cycle. In Table 2, we have the pairwise dissimilarity between states. Note that the higher the index, the lesser the synchronism between cycles. In that sense, the higher the index between two entities, the further away they are from each other.

[8] The only exception is for the 1912 presidential run. In that election, Theodore Roosevelt failed to receive the Republican nomination. Roosevelt, created the Progressive party and ran for president, diveding the Republican electorate. For this specific election, we compare the votes of the Democratic candidate (Woodrow Wilson) with the total of the votes of the other two major contendants (William Taft and Roosevelt).

	United States			United States			United States
Ohio	5.80		Rhode Island	13.17		Nevada	18.35
Maine	8.14		Massachusetts	13.59		Vermont	20.03
New Hampshir	10.21		Maryland	13.79		Utah	20.79
New Jersey	10.38		Pennsylvania	13.91		Texas	25.30
California	10.48		Missouri	13.98		Arkansas	25.44
Wyoming	10.50		South Dakota	14.38		Tennessee	27.54
Iowa	10.54		Wisconsin	14.41		Kentucky	28.01
Connecticut	10.56		Delaware	14.48		Virginia	28.40
Indiana	10.99		Minnesota	15.18		Louisiana	32.36
New York	11.49		West Virginia	15.37		Georgia	33.63
North Dakota	11.52		Colorado	16.72		North Carolina	33.65
Illinois	11.64		Idaho	17.17		Florida	34.51
Kansas	11.86		Montana	17.35		Alabama	47.47
Michigan	11.93		Oregon	17.94		Mississippi	53.46
Nebraska	12.90		Washington	18.35		South Carolina	58.65

Table 1: Dissimilarity index between the National and the State electoral cycle

The 10 states with the electoral cycle more aligned with the national cycle are Ohio, Maine, New Hampshire, New Jersey, California, Wyoming, Iowa, Connecticut, Indiana and New York. Note that the fact that these states have an electoral cycle synchronized with the national cycle, does not mean that the candidate that wins in these states is the candidate that wins the country. It just means that the swings around the mean are synchronized.

Using Table 2, we can see that the most similar pair of states is New Jersey/New York, followed by New Hampshire/Maine and New Jersey/Connecticut. The most dissimilar are Pennsylvania/Mississippi, South Carolina/Indiana and South Carolina/Ohio.

Because Table 2 has too much information to be readable, more than 900 entries, we try to visualize this matrix, performing some clustering analysis (e.g. see Camacho *et al.* 2006). First we produce a hierarchical tree clustering. The idea is to group the states according to their similarities. We follow a bottom up approach. We start with the 45 countries and group, in cluster, the two most similar countries, say C_1 and C_2 (New Jersey and New York, to be more precise). In the second round, countries C_1 and C_2 are replaced by a a combination of the two, say C_{46}. Now one has to build a new matrix, not only with the distance between the 44 remaining states, but also with the distance between each state and C_{13} (which we consider to be the average of the individual distances). The procedure continues until there is only one cluster with all the countries. In Figure

8, we can see the result of this hierarchical clustering.

	AL	AR	CA	CO	CT	DE	FL	GA	ID	IL	IN	IA	KS	KY	LA	ME	MD	MA	MI	MN	MS	MO	MT	NE	NV	NH	NJ	NY	NC	ND	OH	OR	PA	RI	SC	SD	TN	TX	UT	VT	VA	WA	WV	WI	WY
AL	0																																													
AR	32	0																																												
CA	48	34	0																																											
CO	47	28	18	0																																									
CT	49	32	12	25	0																																								
DE	40	35	15	27	9	0																																							
FL	35	29	37	35	35	32	0																																						
GA	29	26	34	39	38	31	31	0																																					
ID	54	22	21	17	23	27	32	30	0																																				
IL	55	27	12	18	12	15	34	38	22	0																																			
IN	49	21	18	20	18	20	32	37	19	16	0																																		
IA	54	34	15	13	17	21	34	50	20	16	15	0																																		
KS	49	23	17	13	19	24	29	35	17	19	13	14	0																																	
KY	40	22	30	32	25	25	30	31	34	34	21	29	23	0																															
LA	38	29	34	41	30	28	20	31	32	33	30	34	36	31	0																														
ME	47	31	10	21	6	10	33	37	20	13	16	15	16	25	28	0																													
MD	48	31	14	24	10	13	36	32	26	12	19	21	22	24	32	11	0																												
MA	51	33	13	26	6	10	34	39	28	12	21	19	22	26	29	10	12	0																											
MI	49	29	9	14	14	17	37	37	21	9	17	15	17	32	37	12	14	15	0																										
MN	52	31	17	16	21	25	35	44	23	17	22	15	21	35	33	19	22	22	14	0																									
MS	21	25	56	50	55	54	45	17	46	54	57	57	51	38	27	57	51	50	46	46	0																								
MO	43	19	21	17	16	20	30	31	23	15	11	18	11	17	32	16	20	17	18	25	54	0																							
MT	50	31	18	15	22	25	31	40	11	23	17	17	17	31	30	19	25	28	18	21	45	21	0																						
NE	49	24	15	13	19	22	31	35	15	18	17	10	13	34	33	16	23	21	16	17	54	18	14	0																					
NV	53	25	14	15	22	22	31	29	16	21	20	18	15	34	39	18	21	22	14	20	40	20	12	15	0																				
NH	47	32	12	23	6	11	33	39	21	15	17	17	17	25	27	5	11	9	15	21	57	15	20	18	20	0																			
NJ	47	34	12	27	6	9	34	37	25	12	18	18	20	27	30	8	9	8	16	23	52	16	25	19	21	8	0																		
NY	49	35	13	26	6	10	34	35	27	11	19	18	21	26	30	9	9	7	15	22	50	17	27	20	19	9	3	0																	
NC	26	23	42	34	39	24	29	24	39	42	29	43	28	21	32	34	27	47	41	47	18	31	40	32	43	31	38	44	0																
ND	50	26	14	18	17	21	32	38	17	19	15	12	15	31	36	14	21	20	16	17	41	18	12	8	14	16	18	19	42	0															
OH	44	23	10	16	12	16	34	35	16	16	12	12	11	26	31	9	15	15	13	14	56	16	15	13	14	11	13	14	28	11	0														
OR	50	29	14	10	24	25	37	31	16	23	15	14	12	29	42	20	24	25	17	20	36	20	13	11	14	21	25	25	37	14	13	0													
PA	48	35	12	19	15	18	36	39	23	15	19	14	19	32	31	13	16	18	11	16	66	24	23	20	18	16	17	17	36	19	13	17	0												
RI	52	35	12	25	7	12	34	38	23	13	20	19	21	30	27	8	11	7	13	21	46	19	25	20	21	9	11	9	38	19	14	24	16	0											
SC	28	33	54	54	53	54	47	18	46	54	58	56	57	42	33	53	57	51	47	48	13	56	48	54	40	53	55	53	25	42	58	47	56	49	0										
SD	47	26	17	19	17	20	33	30	20	21	18	17	18	34	36	15	21	19	19	23	41	20	16	15	14	16	17	19	38	9	15	17	22	19	44	0									
TN	29	17	33	33	26	22	27	24	35	31	24	36	30	20	25	25	25	33	33	39	23	28	33	29	35	24	27	30	16	32	26	39	33	28	31	30	0								
TX	51	36	29	26	26	23	27	34	22	28	22	24	22	24	21	23	28	26	31	27	31	27	25	21	36	22	26	27	32	26	25	25	24	24	38	27	34	0							
UT	50	32	23	16	25	31	34	31	13	25	24	20	20	36	31	21	29	29	21	23	42	28	13	18	18	22	29	31	37	18	19	14	21	25	43	24	38	23	0						
VT	53	39	19	27	13	17	37	36	30	17	24	25	26	31	30	13	13	14	15	23	41	24	29	25	26	14	17	16	36	27	21	27	17	10	43	26	32	27	28	0					
VA	39	32	32	29	30	22	21	28	28	31	25	25	28	24	27	28	29	30	32	27	52	29	34	30	34	28	32	32	14	31	28	28	26	27	55	34	19	21	34	28	0				
WA	53	35	12	11	21	25	39	42	18	18	20	13	17	34	41	18	22	21	12	17	46	23	16	14	13	20	23	22	46	17	15	9	14	20	50	21	38	30	17	26	29	0			
WV	43	21	19	23	17	18	29	36	22	22	14	18	17	17	32	15	19	19	20	24	50	15	20	21	22	16	19	19	30	20	16	19	18	18	52	21	30	19	22	22	20	24	0		
WI	50	41	12	18	17	22	35	45	23	15	20	12	18	32	37	17	18	19	13	10	51	24	22	14	17	19	18	18	48	15	13	18	12	18	53	19	39	31	21	22	27	14	23	0	
WY	49	20	15	13	14	20	30	31	11	17	14	12	11	26	33	13	20	17	14	17	47	16	8	10	14	13	16	17	30	11	10	11	18	16	54	15	26	22	13	22	30	14	13	17	0

Table 2: Dissimilarity Index Matrix

Depending on how demanding we are to define a cluster, one can identify several clusters. If we partition the tree in three clusters, then we have a big cluster of several states whose cycles are very similar. Then there is a smaller second cluster (Arkansas, Tennessee, Kentucky, North Carolina, Virginia, Florida, Louisiana and Texas), and finally, a third cluster of states with a very asynchronous electoral cycle (Alabama, Georgia, Mississippi and South Carolina). In sum, the two clusters of states that exhibit greater dissimilarity from the 'national' cycle ■ and from the 'core' states that are closest to that cycle ■ include all the states from the South, plus a 'border South' state (Kentucky)."

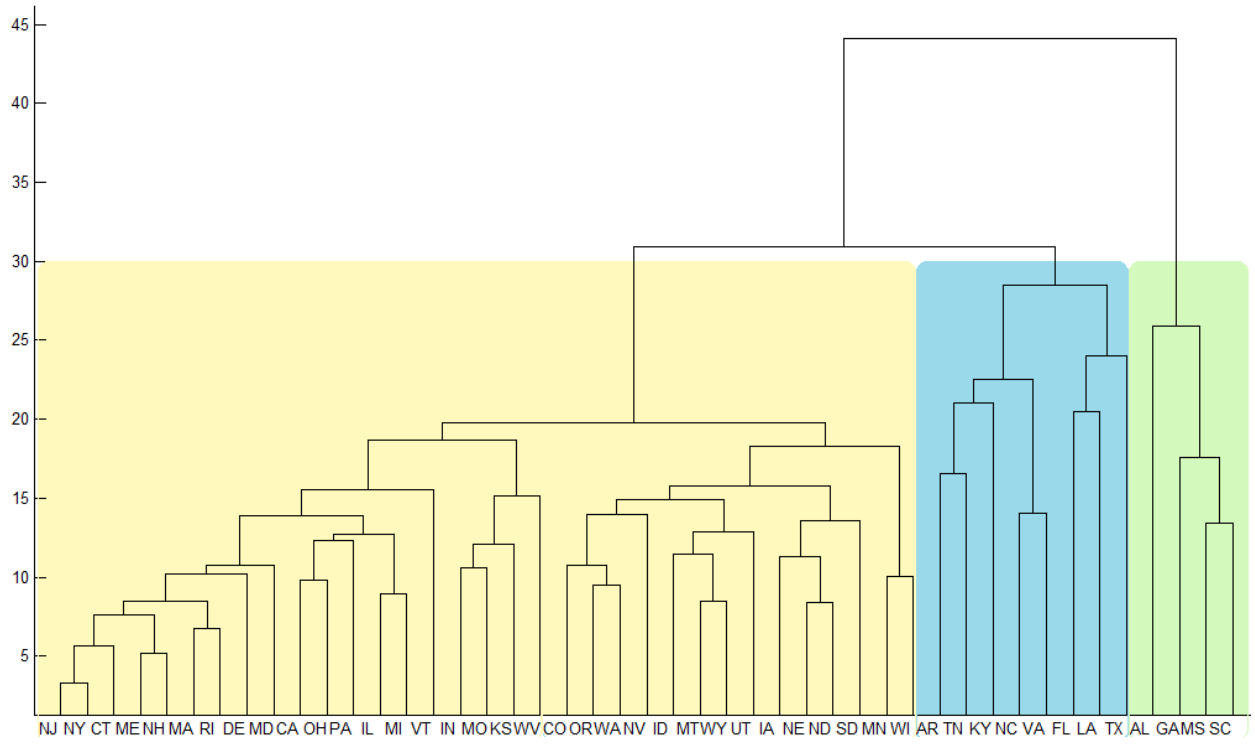


Figure 8: Hierarchical Tree Clusters

Although suggestive, the clustering tree has some limitations that may distort the analysis, because each state is solely linked to one other state (or cluster of states), one may loose site of the whole picture. An alternative approach is to use the dissimilarity matrix as a distance matrix and map the states in a two axis system. The idea is to reduce the dissimilarity matrix to a two column matrix. This new matrix, the configuration matrix, contains the position of each country in two orthogonal axis. Therefore we can position each state on a two dimensional map. This cannot be performed with perfect accuracy because the dissimilarity matrix is nonmetric. i.e. does not represent euclidean distances. Its interpretation should be ordinal. Therefore, the goal is not to reproduce the "distances" given by Table 2 on a map, but to find a map, with pairwise distances that reproduce, as much as possible, the ordering of Table 2. We use Kruskal (1964a and 1964ab)'s stress algorithm and minimize the square differences between the distances in the map and the true "distances" given in Table 2.

Again, although the precise frontiers are, naturally, a little bit arbitrary, in Figure 9 it is possible to identify three clusters of states that almost coincide with the information we had extracted from the clustering tree.

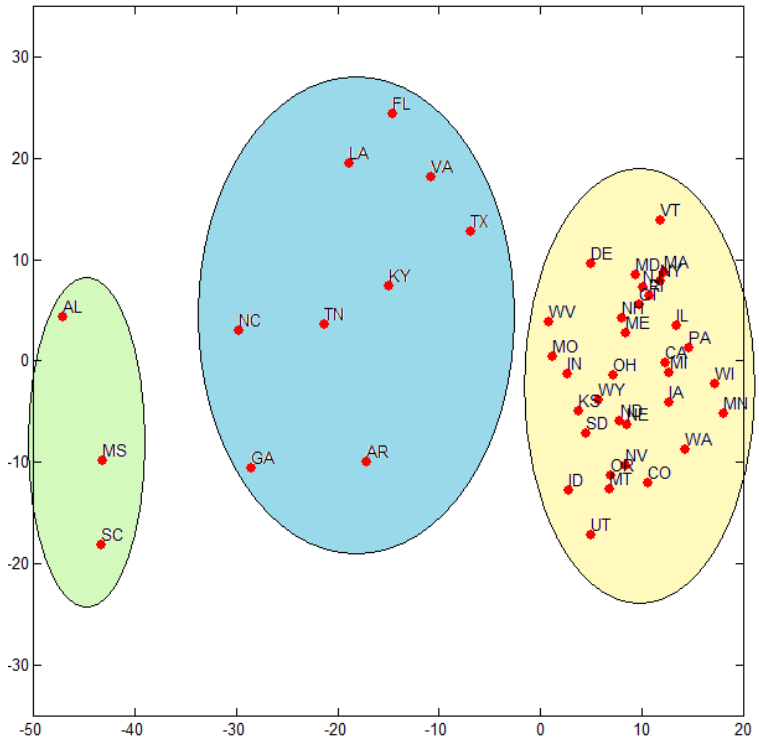


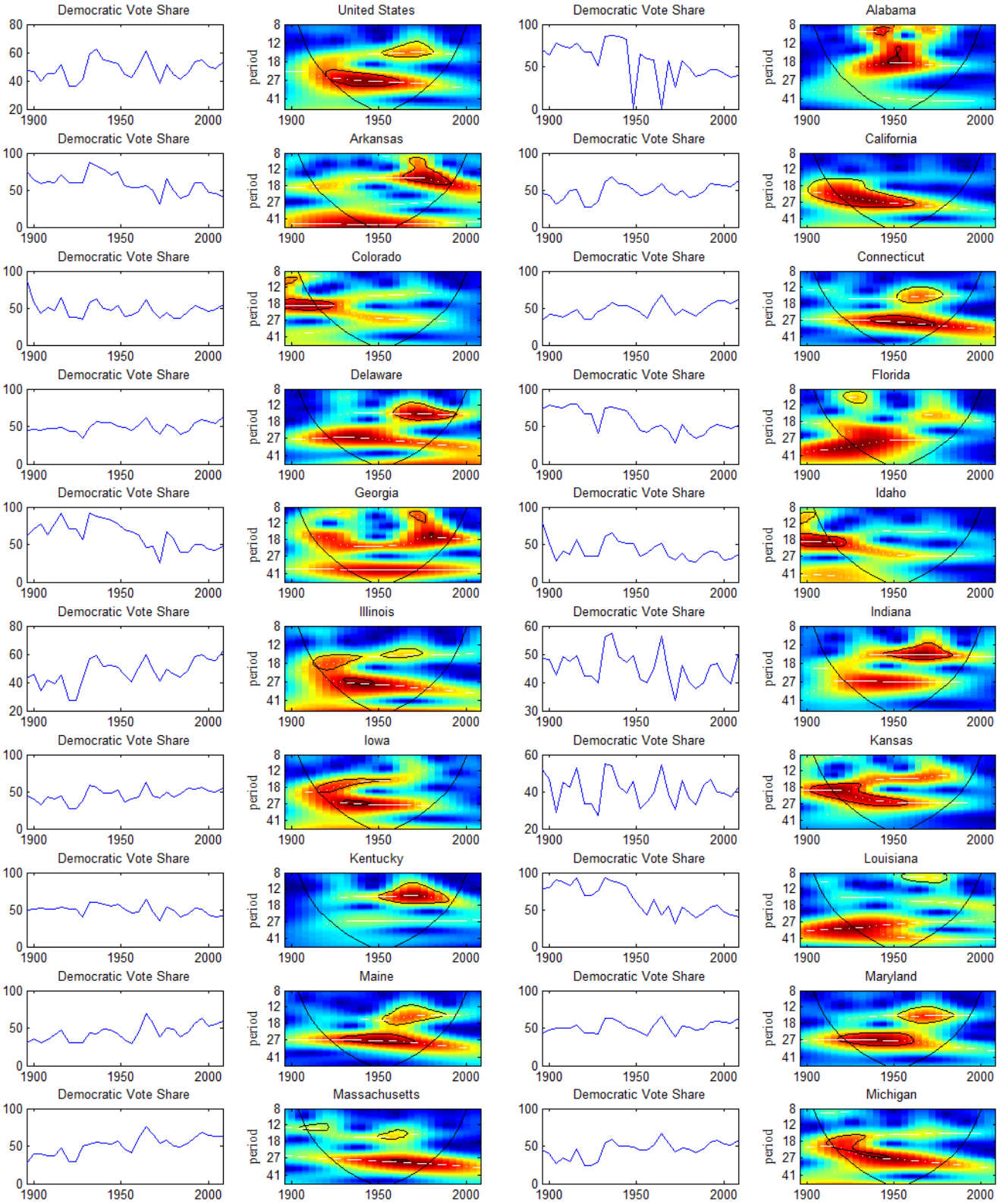
Figure 9: Multidimensional Scaling Map

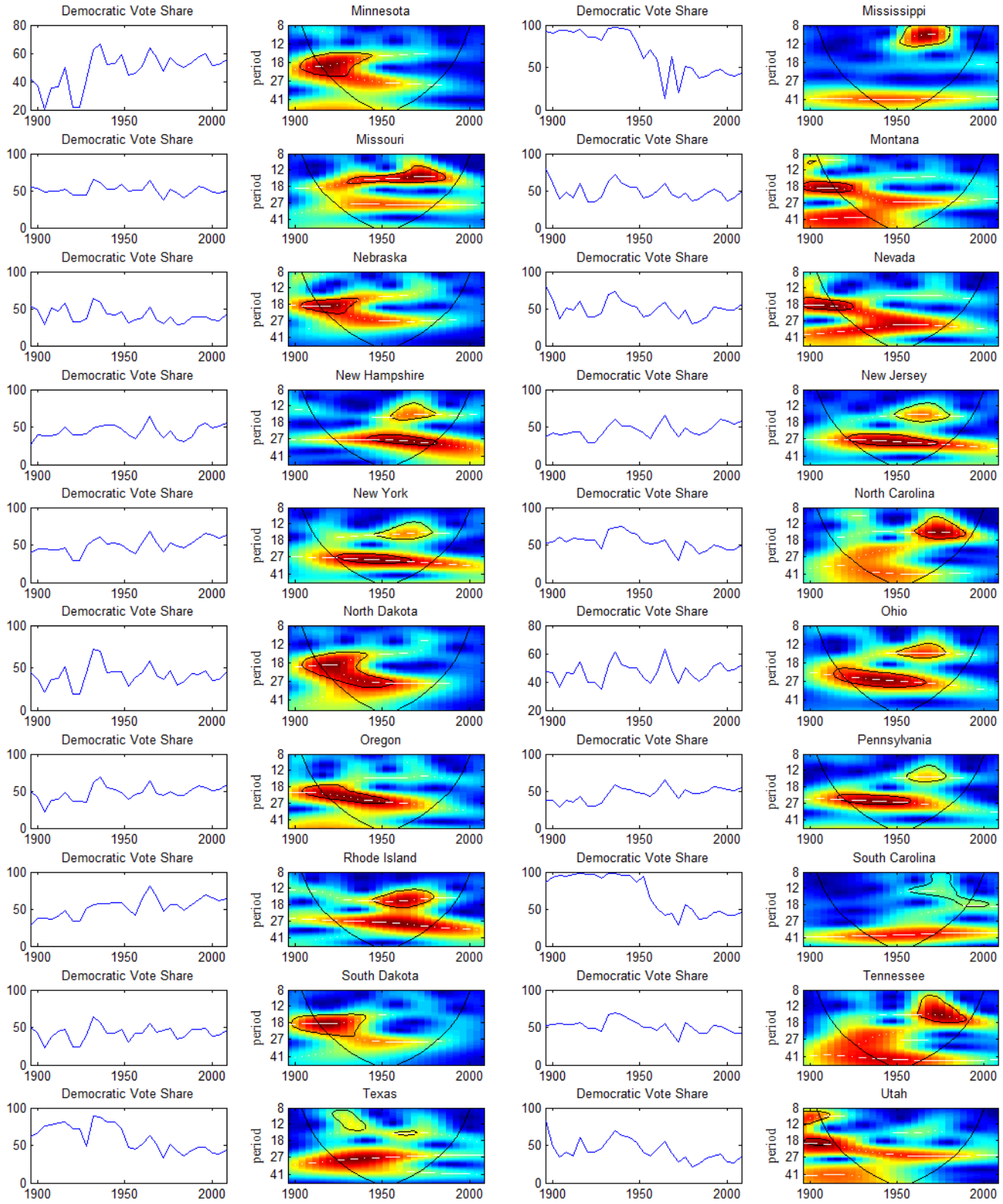
4.2.1 Wavelet Power Spectrum

In Figure 10, we can see the continuous wavelet power spectra of the electoral democratic share for several states. We assess the statistical significance against the null hypothesis of an AR(1).Looking at the time-frequency decomposition some interesting facts are revealed. First, as we have already argued, looking at the United Sates as a whole, we identify a persistent 26 year cycle, like Merrill *et al.* (2008), but we also identify a transient 15 year cycle between early 1950s and 1980. The same pattern is found in some states, like Virginia, Maine, Ohio, Maryland, New Hampshire, New York, Pennsylvania, among some others. But not all states are alike. For example, in Washington and Wisconsin most of the action occurred until 1950, at several frequencies. In Utah, the 16 year cycle is not apparent. In Tennessee, a $10 \sim 14$ year cycle is very strong between 1960 and 1990. In Texas one can find a cycle at these same frequencies before 1950, etc.

Although suggestive, the wavelet power spectrum is not the best tool to analyze business cycle synchronization, as no information is revealed about the phase. Therefore, even if two states share a similar high power region, one cannot infer that their electoral cycles look alike. It is possible that while in one state the democratic share tends to increase in the other state the reverse is

happening. We now move to the cross wavelet tools.




Democratic Vote Share
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New York
North Carolina
North Dakota
Ohio
Oregon
Pennsylvania
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
period
1900
1950
2000

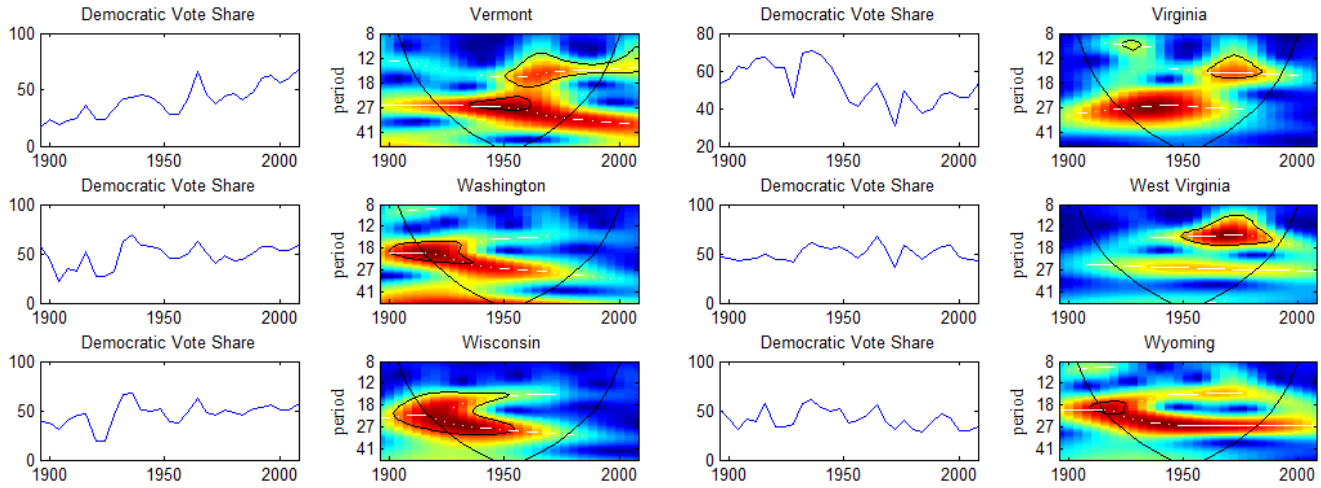
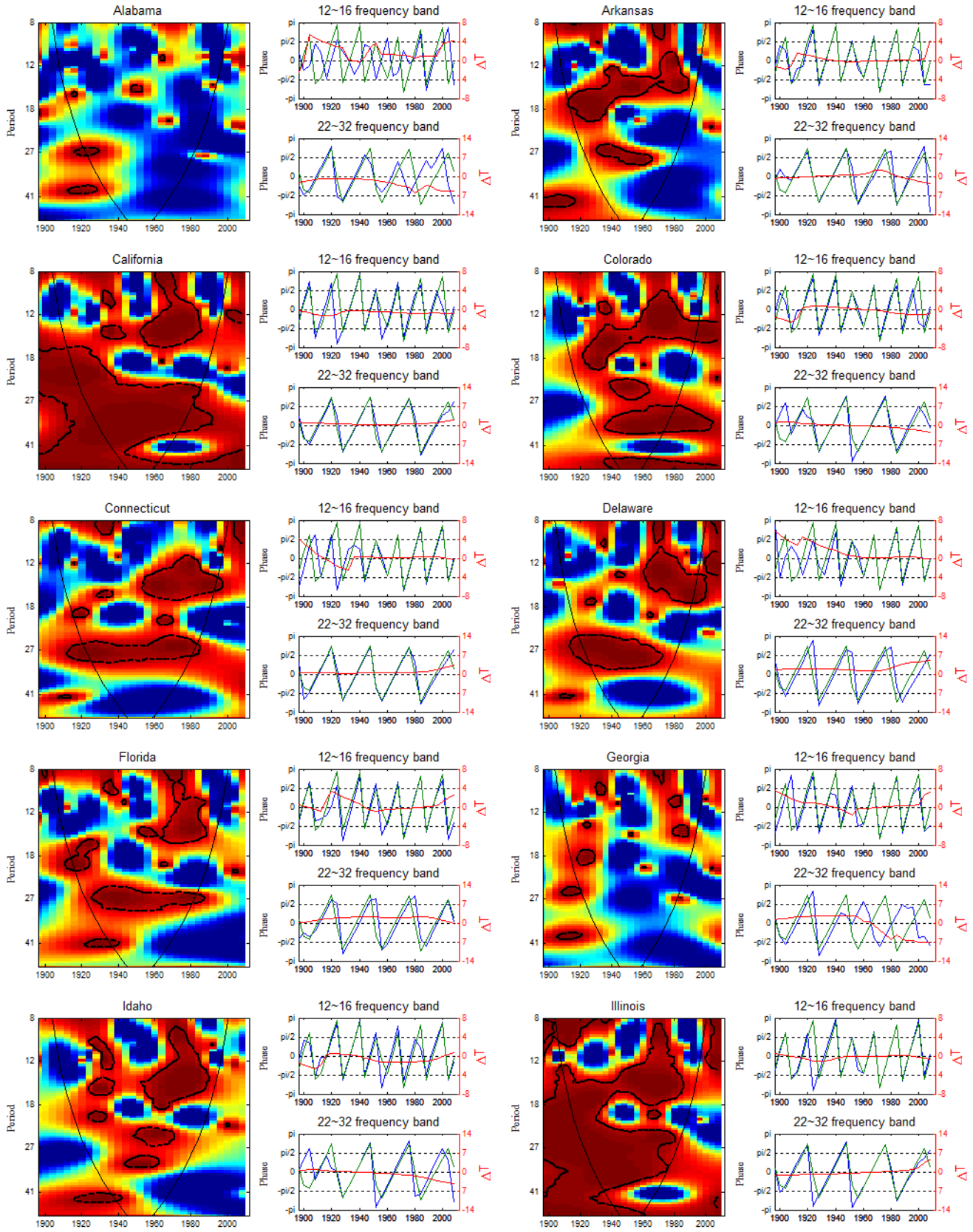
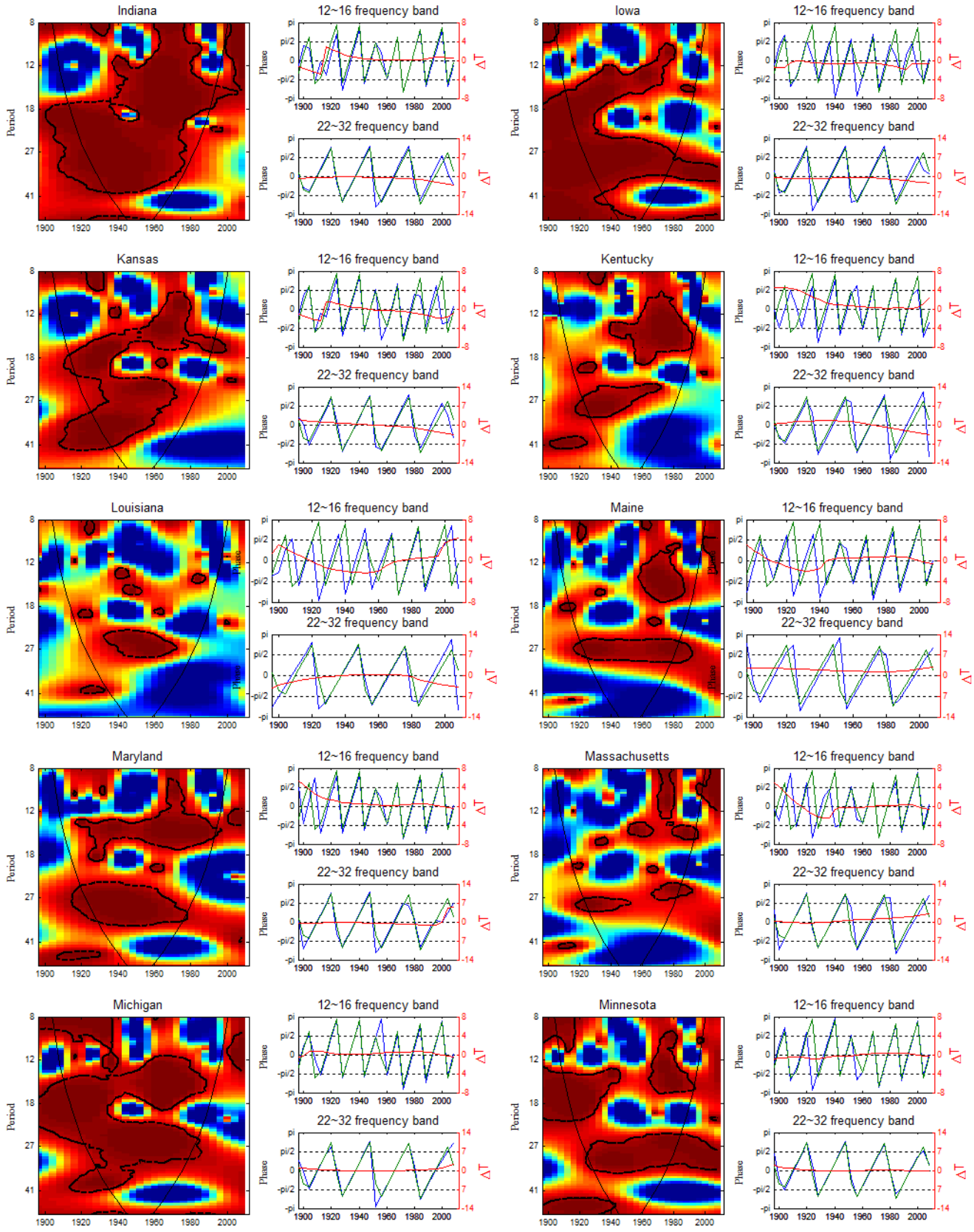


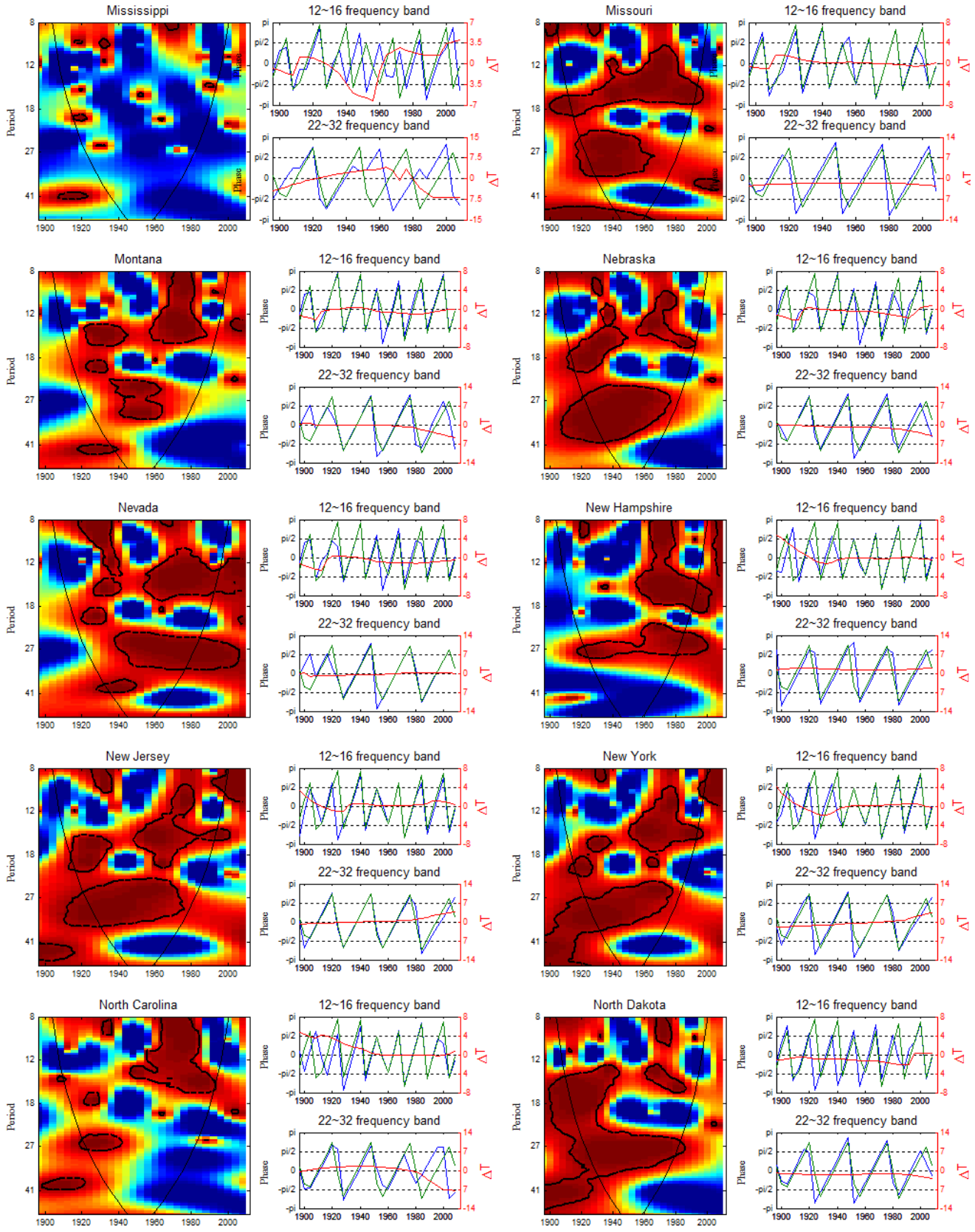
Figure 10: Figure 10: On the left, the Democratic Share by State in each Presidential election since 1896. On the right, the Wavelet Power Spectrum — The black contour designates the 5% significance level against an AR(1) null. The cone of influence, which indicates the region affected by edge effects, is shown with a thin black line. The color code for power ranges from blue (low power) to red (high power). The white lines show the maxima of the undulations of the wavelet power spectrum.

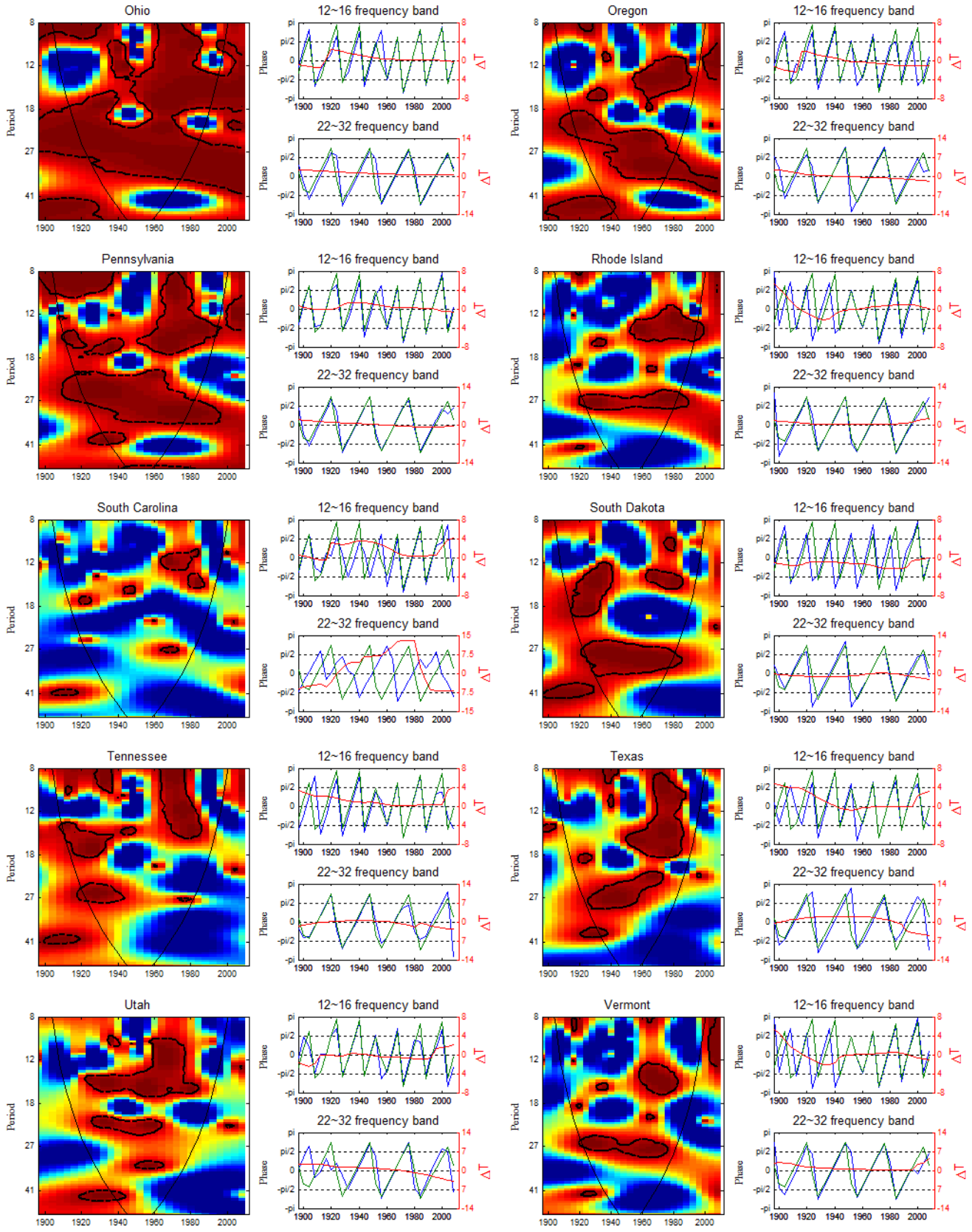
4.2.2 Phase-difference and cross-wavelets

To perform the cross-wavelet analysis we will focus on the wavelet coherency, instead of the wavelet cross spectrum, because there is some redundancy between both measures and the wavelet coherency has the advantage of being normalized by the power spectrum of the two time-series. Regions of high coherency between two countries are synonym of strong local (both in time and frequency) correlation.


Alabama
Arkansas
California
Colorado
Connecticut
Delaware
Florida
Georgia
Idaho
Illinois
12~16 frequency band
22~32 frequency band
Period
Phase
ΔT
8
12
18
27
41
pi
pi/2
0
-pi/2
-pi
1900
1920
1940
1960
1980
2000
14
7
-14


Indiana
12~16 frequency band
22~32 frequency band
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Period
Phase
ΔT


Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New York
North Carolina
North Dakota
12~16 frequency band
22~32 frequency band
Period
Phase
ΔT


Ohio
12~16 frequency band
22~32 frequency band
Oregon
12~16 frequency band
22~32 frequency band
Pennsylvania
12~16 frequency band
22~32 frequency band
Rhode Island
12~16 frequency band
22~32 frequency band
South Carolina
12~16 frequency band
22~32 frequency band
South Dakota
12~16 frequency band
22~32 frequency band
Tennessee
12~16 frequency band
22~32 frequency band
Texas
12~16 frequency band
22~32 frequency band
Utah
12~16 frequency band
22~32 frequency band
Vermont
12~16 frequency band
22~32 frequency band
Period
Phase
ΔT
1900 1920 1940 1960 1980 2000

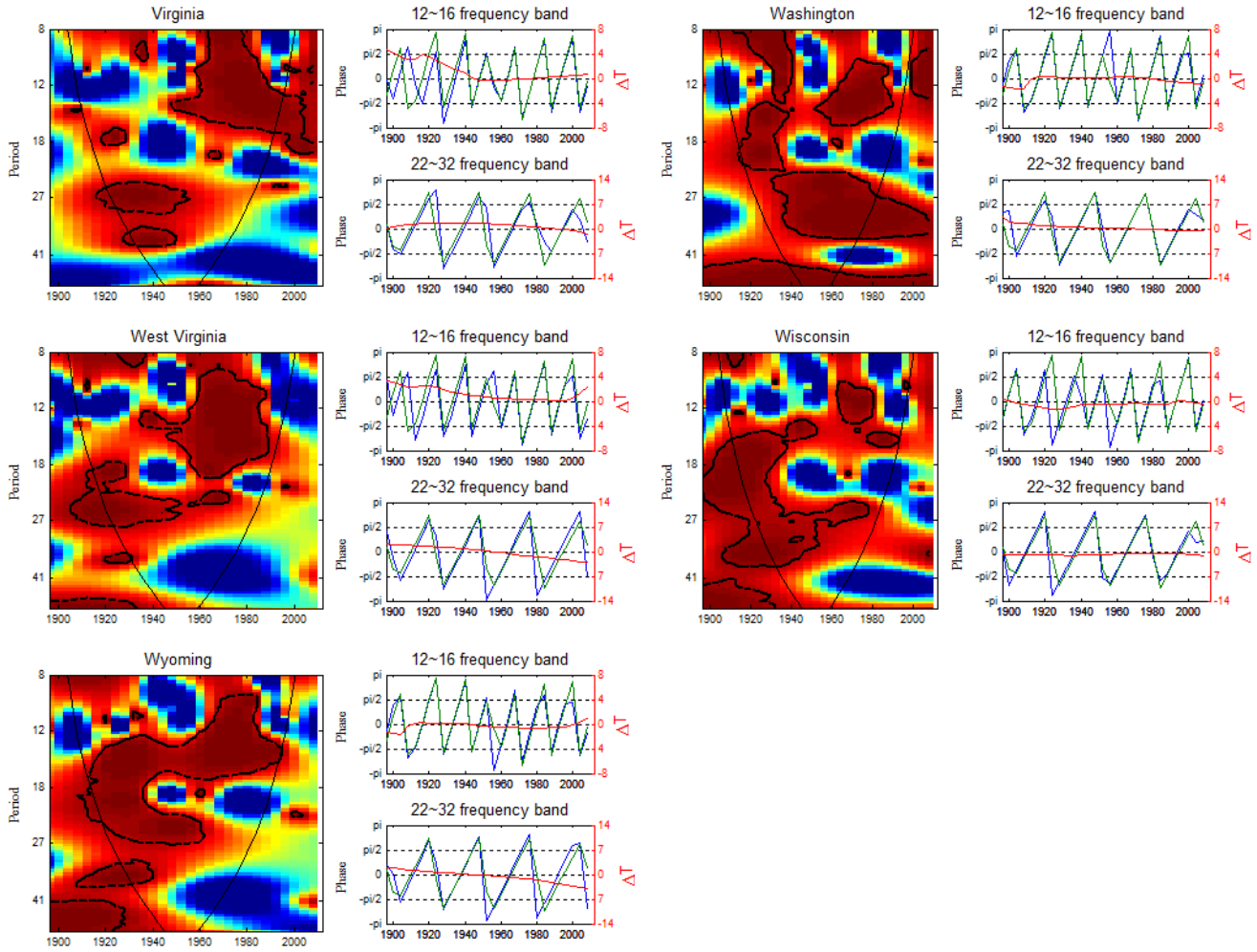


Figure 11: On the Left:Cross-Wavelet Coherency − The black thick contour designates the 5% significance level estimated by Monte Carlo simulations against an AR(1) null. The color code for coherency ranges from blue (low coherency – close to zero) to red (high coherency – close to one). On the right: Phase and instantaneous time lag between the two series. The green line represents the National phase, and the blue line represents the state's phase. The red line gives us the instantaneous time lag between the two series.

The phase-difference gives us information on the delay, or synchronization, between oscillations of the two time-series for a given frequency. In Figure 11, for each state, we estimate the coherency between the national electoral democratic share and the state share. We also estimate the phase of the oscillations at the national and state level, as well as their phase-difference. Given that we identified two main cycles, one at the 14 year frequency and the other at the 27 years frequency, we focus our phase difference analysis on these. So for each state, we calculate the average phase and

phase-difference for the $12 \sim 16$ and for the $22 \sim 32$ frequency bands. For ease of interpretation, we have the instantaneous time lag on the right axis.

We can appreciate some interesting dynamics. For example, the more peripherical states (according to Table 1), like South Carolina and Mississippi, do not show many regions of high coherence and the phase difference shows that their cycles not only are not aligned with the rest of the country but also show no definite pattern. Other states, like Michigan, Pennsylvania, Washington, New Jersey, Minnesota, etc., show many areas of strong coherency and show oscillations that are very much aligned with the national oscillations. Ohio, that according to table 1 is the most aligned state, shows many regions of high coherency (probably the most coherent state), but their phases reveal that its cycles have been slightly lagging the national cycle (although it is also clear that even on that regard there has been a strong convergent since mid-century. In the case of New York, we observe the opposite dynamics: New York seems to have led the national cycle (on both frequencies) until mid-century, after which it converged to the national cycle. Massachusetts also displays a change in the long run behavior. In the first half of the sample, in the $22 \sim 32$ frequency band, the phases are very much aligned, but after 1950 this long cycle is lagging the national cycle.

We can also identify some states that are very much synchronized for some periods and some frequencies, but not for others. North Carolina is one such example. In the first half of last century, there is a region of high coherency at the 27 years frequency, while in the second half the high coherency shifts to the $12 \sim 18$ year frequency. In this latter case, one can also see that the phases are perfectly aligned with the national phase

If one had to choose the leader state, that choice would fall on North Dakota (also Illinois, but not as strongly), whose cycles have persistently been leading the national cycles, on both frequency bands.

5 Conclusions

We have claimed that wavelet analysis can naturally be applied to the study of political and electoral cycles (given its periodic nature), specially when one is interested in estimating the spectrum as a function of time, revealing how the different periodic components of the time-series change over time. The main advantage of the wavelet approach over spectral analysis is the ability to analyze transient dynamics, both for single time-series or for the association between two time-

series. We used three tools that, to our knowledge, have not yet been used by political scientists: the wavelet power spectrum, the cross-wavelet coherency and the phase-difference. While the wavelet power spectrum 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 possible to study synchronization and delays between two time-series across different frequencies or timescales.

We have also developed a new metric to compare different power spectra. We focused on the common high power time-frequency regions, extracting the components of the covariance matrix of the wavelet spectra pairs using the 'Singular Value Decomposition'. Because the wavelet is complex, we had to define the distance between complex vectors, which led us to use Hermitian angles. Given this, we proposed a metric to measure the distances between wavelet spectra and build a dissimilarity matrix, to which clustering and multidimensional scaling techniques can be applied.

Using aggregate data for the US presidential elections since 1856, we aimed at answering the questions of whether there are cycles in presidential electoral politics and whether such cycles are common to all states. Like previous works using either spectral analysis or autoregressive models, we were able to identify a 26 year cycle. However, wavelet analysis also allowed us to determine that such cycle dissipates since the late 1960s, and that a transitional 15 year cycle, initiated in the late 1950s, has prevailed since then. Furthermore, using state level data for the presidential elections since 1896, we found several interesting facts. There are several states, like Ohio, Maine, New Hampshire, New Jersey, California, Wyoming, Iowa, etc., whose electoral cycle is highly synchronized with the national electoral cycle. At the same time there are two clusters of states that display asynchronous electoral cycles, i.e., all the states in South, particularly Alabama, Georgia, Mississippi and South Carolina.

Finally, some peculiar dynamics were revealed with the use of the wavelet coherency and the phase difference. We also saw that asynchronous states, like South Carolina and Mississippi, do not show many regions of high coherence and the phase difference shows that their cycles not only are not aligned with the rest of the country but also show no definite pattern. On the other extreme, Ohio, which is the most synchronous state, has cycles that have been slightly lagging the national cycle, although this lag has been disappearing with time. In other states, like New York, we observe

the opposite dynamics: New York seems to have led the national cycle (on both frequencies) until mid-century, after which it converged to the national cycle. We could also identify some states that are very much synchronized for some periods and some frequencies, but not for others. For example, North Carolina shows strong coherency at the 27 years frequency in the first half of the 2009 century, while in the second half the high coherency shifts to the $12 \sim 18$ year frequency, with phases that are perfectly synchronized with the national phase. If one had to name a state that has consistently led the national cycle, that choice would fall on North Dakota, whose cycles have persistently been leading the national cycles, at all frequency bands.

In sum, this paper's main contribution to the literature is to clearly demonstrate the utility of wavelets and cross-wavelets for the analysis of political time-series and to illustrate how relationships between variables can change over time and across different frequencies. We applied these wavelet tools to electoral data from the U.S. Presidential elections and we were able to detect transient effects which would be very difficult to detect using spectral analysis or classical time-series techniques. Such findings have important implications for the discussion of the existence of cycles in American politics.

References

[1] Adams, H. (1919), "The rule of phase applied to history." In H.Adams, *The degradation of the democratic dogma*, New York, p. 267–311.

[2] Aguiar-Conraria, L., Azevedo, N. and Soares, M. J. (2008) "Using Wavelets to Decompose the Time-Frequency Effects of Monetary Policy", *Physica A: Statistical Mechanics and its Applications*, 387, 2863–2878.

[3] Bartels, L. (1998), "Electoral continuity and change, 1868-1996." *Electoral Studies,* 17(3), 301-326.

[4] Beck, N., (1991), "The Illusion of Cycles in International Relations", *International Studies Quarterly*, 35(4), 455-476.

[5] Bloomfield, D., McAteer, R., Lites, B., Judge, P., Mathioudakis, M. and Keena, F. (2004), "Wavelet Phase Coherence Analysis: Application to a Quiet-Sun Magnetic Element", *The Astrophysical Journal*, 617, 623–632.

[6] Brown, R., and Bruce, J. (2008), "Partisan-Ideological Divergence and Changing Party Fortunes in the States, 1968–2003: A Federal Perspective", *Political Research Quarterly,* 61(4), 585-597.

[7] Burnham, W. D. (1970), *Critical elections and the mainsprings of American politics*. New York: Norton.

[8] Burnham, W. D. (1967), "Party systems and the political process." In *The American Party Systems: Stages of Political Development*, New York: Oxford University Press, 277–307.

[9] Campbell, J. E. (2006), "Party Systems and Realignments in the United States, 1868-2004." *Social Science History,* 30(3), 359–386.

[10] Carmines, E. G, and J. A Stimson. (1989), *Issue evolution*. Princeton: Princeton University Press.

[11] Cazelles, B., Chavez, M., de Magny, G. C., Guégan, J-F and Hales, S. (2007), "Time-Dependent Spectral Analysis of Epidemiological Time-Series with Wavelets", *Journal of the Royal Society Interface*, 4, 625–36.

[12] Crowley, P. (2007), "A Guide to Wavelets for Economists", *Journal of Economic Surveys*, 21 (2), 207–267.

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

[14] Enders, W., and Sandler, T., (2000), "Is transnational terrorism becoming more threatening? A time-series investigation", *Journal of Conflict Resolution*, 44(3), 307–332.

[15] Enders, W., and Sandler, T., (2002), "Patterns of transnational terrorism, 1970-1999: Alternative time-series estimates", *International Studies Quarterly*, 46(2), 145-165.

[16] Enders, W., and Sandler, T., (2006), *The Political Economy of Terrorism*, Cambridge University Press.

[17] Engle, R. (2002), "Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH", *Journal of Business and Economics Statistics,* 20, 339-50.

[18] Foufoula-Georgiou, E. and Kumar, P. (1994), *Wavelets in Geophysics*, volume 4 of *Wavelet Analysis and Its Applications*. Academic Press, Boston.

[19] Gabor, D., (1946), "Theory of Communication", *Journal of the Institute of Electrical Engineers*, 93, 429-457.

[20] Gallegati, M. (2008), "Wavelet analysis of stock returns and aggregate economic activity", *Computational Statistics and Data Analysis*, 52, 3061—3074.

[21] Gans, D. J. (1985), "Persistence of Party Success in American Presidential Elections." *Journal of Interdisciplinary History* 16(2): 221-237.

[22] Gençay, R., Selçuk, F. and Witcher, B. (2005), "Multiscale Systematic Risk", *Journal of International Money and Finance,* 24, 55–70.

[23] Gerace, M., (2002), "US - Military Expenditures and Economic Growth: Some Evidence from Spectral Methods", *Defence and Peace Economics,* 13(1), 1-11.

[24] Goodhart, C., and Bhansali, R. (1970), "Political Economy", *Political Studies,* 18(1), 43-106.

[25] Goldstein, J. (1988), *Long Cycles: Prosperity and War in the Modern Age*, Yale University Press.

[26] Goupillaud, P., A. Grossman and Morlet, J. (1984), "Cycle-Octave and Related Transforms in Seismic Signal Analysis", *Geoexploration*, 23, 85–102.

[27] Grinsted, A., Moore, J. and Jevrejeva S. (2004), "Application of the Cross Wavelet Transform and Wavelet Coherence to Geophysical Time Series", *Nonlinear Processes in Geophysics*, 11(5/6), 561-566.

[28] Grossmann, A. and Morlet, J. (1984), "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape", *SIAM Journal on Mathematical Analysis,* 15, 723-736.

[29] Howlett, M., (1998), "Predictable and unpredictable policy windows: institutional and exogenous correlates of Canadian federal agenda-setting", *Canadian Journal of Political Science/Revue canadienne de science politique*, 31(3), 495-524.

[30] Hubbard, B. (1996), *The world According to Wavelets: The Story of a Mathematical Technique in the Making*, A K Peters, Wellesley Massachusetts.

[31] Hudgins, L., Friehe, C. and Mayer, M. (1993) "Wavelet Transforms and Atmospheric Turbulence" *Physical Review Letters*, 71:20, 3279-82.

[32] Jevrejeva, S., Moore, J. and Grinsted, A. (2003), "Influence of the Arctic Oscillation and El Niño-Southern Oscillation (ENSO) on Ice Conditions in the Baltic Sea: The wavelet approach", *Journal of Geophysical Research*, 108.

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

[34] Kelly, B., Hughes, P., Aller, H. and Aller, M. (2003), "The Cross-Wavelet Transform and Analysis of Quasi-Periodic Behavior in the Pearson-Readhead VLBI Survey Sources", *The Astrophysical Journal*, 591, 695–713.

[35] Kruskal, J. B. (1964a), "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis", *Psychometrika*, 29, 1-27.

[36] Kruskal, J. B. (1964b), "Nonmetric multidimensional scaling: A numerical method", *Psychometrika*, 29, 115-129.

[37] Lebo, M. and Box-Steffensmeier, J. (2008), "Dynamic Conditional Correlations in Political Science", *American Journal of Political Science*, 52:3,688-704.

[38] Lin, T.-M., and Guillén, M., (1998), "The Rising Hazards of Party Incumbency: A Discrete Renewal Analysis", *Political Analysis,* 7(1), 31-57.

[39] Mayhew, D. R. (2000), "Electoral Realignments." Annual Review of Political Science 3(1): 449–474.

[40] Mayhew, D. R. (2002), *Electoral Realignments: A Critique of an American Genre*. New Haven: Yale University Press.

[41] Merrill, S., B. Grofman, and T. L Brunell (2008), "Cycles in American National Electoral Politics, 1854–2006: Statistical Evidence and an Explanatory Model", *American Political Science Review,* 102(1), 1–17.

[42] Miller, W. and Mackie, M. (1973), "The Electoral Cycle and the asymmetry of Government and opposition Popularity: an Alternative model of the relationship between Economic Conditions and Political Popularity", *Political Studies*, 21(3), 263-279.

[43] Midlarsky, M. I. (1984), "Political Stability of Two-Party and Multiparty Systems: Probabilistic Bases for the Comparison of Party Systems." *The American Political Science Review,* 78(4), 929-951.

[44] Nerlove, M. (1964), "Spectral Analysis of Seasonal Adjustments Procedures", *Econometrica*, 32 (3), 241–286.

[45] Norpoth, H. (1995), "Is Clinton Doomed? An Early Forecast for 1996", *PS: Political Science and Politics,* 28(2), 201-207.

[46] Norpoth, H. (2002), "On a Short Leash: Term Limits and the Economic Voter." In Dorussen, H. and M. Taylor (Eds.), *Economic voting*. Routledge.

[47] Nardulli, P. F. (1995), "The concept of a critical realignment, electoral behavior, and political change", *American Political Science Review*, 89, 10-22.

[48] Ramsey, J. and Lampart, C. (1998), "Decomposition of Economic Relationships by Time Scale Using Wavelets: Money and Income", *Macroeconomic Dynamics*, 2, 49–71.

[49] Richards, D., (1992), “Spatial Correlation Test for Chaotic Dynamics in Political Science”, *American Journal of Political Science,* 36 (4), 1047–1069.

[50] Rouyer, T., Fromentin, J.-M., Stenseth, N. and Cazelles, B. (2008) "Analysing multiple time series and extending significance testing in wavelet analysis", *Marine Ecology Progress Series*, 359, 11-23.

[51] Scharnhorst, K. (2001) "Angles in Complex Vector Spaces", *Acta Applicandae Mathematicae*, 69, 95–103.

[52] Stokes, D. E., and G, R. Iversen. (1962), “On the Existence of Forces Restoring Party Competition”, *Public Opinion Quarterly*, 26(2), 159-171.

[53] Shafer, B. E. (1991), *The End of realignment?* Madison: University of Wisconsin Press.

[54] Schlesinger, A. M. (1949), *Paths to the Present.* New York: Macmillan.

[55] Torrence, C. and Compo, G. P. (1998), "A Practical Guide to Wavelet Analysis", *Bulletin of the American Meteorological Society*, 79, 605–618.

[56] Wen, Y. (2002), "The Business Cycle Effects of Christmas", *Journal of Monetary Economics*, 49 (6), 289–1314.

[57] Wen Y. (2005), "Understanding the Inventory Cycle", *Journal of Monetary Economics*, 52 (8), 1533–1555.

[58] Williams, J. and McGinnis, M., (1992), “The Dimensions of Superpower Rivalry: A dynamic factor analysis,” *Journal of Conflict Resolution,* 36(1), 86-118.

[59] Zhan,Y., Halliday, D., Jiang, P., Liu, X. and Feng, J. (2006), "Detecting Time-Dependent Coherence between Non-Stationary Electrophysiological Signals—A Combined Statistical and Time–Frequency Approach", *Journal of Neuroscience Methods*, 156, 322–332.