IMK
Institut für Makroökonomie
und Konjunkturforschung
Macroeconomic Policy Institute

March 2016

165

Working Paper

Jörg Breitung and Sven Schreiber[1]

Assessing Causality and Delay within a Frequency Band

this version: February 2016

Abstract

We extend the frequency-specific Granger-causality test of Breitung et al. (2006) to a more general null hypothesis that allows causality testing at unknown frequencies within a prespecified range of frequencies. This setup corresponds better to empirical situations encountered in applied research and it is easily implemented in vector autoregressive models. We also provide tools for estimating the phase shift/delay at some prespecified frequency or frequency band. In an empirical application dealing with the dynamics of US temperatures and CO2 emissions we find that emissions cause temperature changes only at very low frequencies with more than 30 years of oscillation. Furthermore we analyze the indicator properties of new orders for German industrial production by assessing the delay at the frequencies of interest.

Keywords: Granger causality, frequency domain, filter gain

JEL codes: C32 (multivariate times series),
C53 (forecasting and prediction methods),
E32 (business fluctuations, cycles),
Q54 (global warming

[1]Jörg Breitung: University of Cologne. Sven Schreiber (corresponding author): Macroeconomic Policy Institute Düsseldorf (IMK), and Free University Berlin. Postal address: Hans Böckler Foundation, Hans-Böckler-Str. 39, D-40476 Düsseldorf. E-mail: mail@sven.schreiber.name

Assessing Causality and Delay within a Frequency Band*

Jörg Breitung and Sven Schreiber†

this version: February 2016

Abstract

We extend the frequency-specific Granger-causality test of Breitung and Candelon (2006) to a more general null hypothesis that allows causality testing at unknown frequencies within a prespecified range of frequencies. This setup corresponds better to empirical situations encountered in applied research and it is easily implemented in vector autoregressive models. We also provide tools for estimating the phase shift/delay at some prespecified frequency or frequency band. In an empirical application dealing with the dynamics of US temperatures and CO_2 emissions we find that emissions cause temperature changes only at very low frequencies with more than 30 years of oscillation. Furthermore we analyze the indicator properties of new orders for German industrial production by assessing the delay at the frequencies of interest.

Keywords: Granger causality, frequency domain, filter gain

JEL codes: C32 (multivariate times series), C53 (forecasting and prediction methods), E32 (business fluctuations, cycles), Q54 (global warming)

*We thank session and seminar participants at the German Economic Association (VfS) meeting, at the Computational and Financial Econometrics (CFE) meeting, and at the University of Bamberg for inspiring feedback. For computational resources we thank the high-performance computing service (Zedat-HPC) at Free University Berlin.

†Breitung: University of Cologne. Schreiber (corresponding author): Macroeconomic Policy Institute Düsseldorf (IMK), and Free University Berlin. Postal address: Hans Böckler Foundation, Hans-Böckler-Str. 39, D-40476 Düsseldorf. E-mail: mail@sven.schreiber.name

1 Introduction

The notion of empirical causality as predictive ability has a long history in science and was formalized by Granger (1969). It became very popular among practitioners due to the simplicity of its implementation in linear dynamic models, where a test for non-Granger-causality is equivalent to a joint exclusion test of lagged terms of the candidate variable. A generalization of this concept was later introduced by Geweke (1982), who noted that causal effects can vary between the different cycles of time series, where each cyclical component corresponds to a certain frequency of oscillation. However, the practical application of the test that Geweke's causality measure is zero at a certain frequency appeared to be quite difficult until Breitung and Candelon (2006, henceforth BC) noted that in the framework of a linear VAR the null hypothesis is equivalent to two linear restrictions that can be tested with a standard Wald test.

A drawback of the BC test is that the test is formulated in terms of a single frequency point that has to be specified a priori. In practice, however, many test statistics are calculated for a range of frequencies to gain insights into the relationship between the variables, although it is well known that the classical test approach does not allow a rigorous joint interpretation of these set of statistics. Furthermore, the underlying (economic) theory usually does not provide a hypothesis for only a single frequency. For example, consider the following implication of the expectations hypothesis of the term structure as noted by Shiller (1979, p.1190), pointing out Granger causality of short-term interest rates for long-term rates in a range of higher frequencies if the theory were failing:[1] "excess [short-run] volatility implies a kind of forecastability for long rates."

In order to better reflect the hypotheses that come naturally from underlying the-

[1] His precise definition of volatility is "variance of short-term holding yields on long-term bonds". We prefer to substitute the phrase "short-term" with "short-run" to avoid the double meaning of "term" in this context. This volatility is related to the short-run "percentage change in the long-term interest rate" (p. 1191) and thus to high-frequency fluctuations of long-term rates, but in a nonlinear way.

ory we extend the frequency-specific test for Granger non-causality by formulating a generalized null hypothesis for a frequency interval. As a side effect of this work we also present a different representation of the model under the null hypothesis of non-causality at some frequency, which turns out to be helpful for our present purpose.

The BC test was used to analyze Granger-causal effects of money on inflation in a series of papers by Assenmacher-Wesche and Gerlach (Assenmacher-Wesche and Gerlach, 2008a, 2007, 2008b). They noted some moderate size distortions and applied the bootstrap as a small-sample correction, but given the lack of other tools at the time, they were forced to use the point-wise tests even though they analyzed frequency bands. Another use for output forecasting was shown by Lemmens, Croux, and Dekimpe (2008), who concluded that the BC approach was the most efficient test among the ones considered. A more recent application is conducted by Wei (2015), and a concept which is closely related to frequency-specific Granger causality is "partial directed coherence", see Baccalá and Sameshima (2001), where inference is also carried out point-wise.

2 Setup and notation

Consider a standard vector autoregression (VAR) of order p in the two variables x_t and y_t:

$$A(L)\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \mathbf{c} + \begin{bmatrix} u_{x,t} \\ u_{y,t} \end{bmatrix}, \qquad t = p+1,...,T \tag{2.1}$$

where $\mathbf{u}_t = (u_{x,t}, u_{y,t})'$ are normally distributed white noise innovations with contemporaneous covariance matrix Ψ. We initially assume the polynomial $A(L)$ to be stable with roots outside the unit circle such that both x_t and y_t will be stationary. The extended case of unit roots will also be discussed below. Further deterministic terms such as linear trends or seasonal dummies could be easily added. Different lag lengths across or within the equations could be accommodated by setting some of the matrix elements to zero.

Let y_t be the potential target variable that is Granger-caused by x_t under the alternative. Using some obvious notation we can write the second equation of the system as follows:

$$y_t = c_y + \sum_{j=1}^{p} \alpha_j y_{t-j} + \sum_{k=1}^{p} \beta_k x_{t-k} + u_{y,t} \tag{2.2}$$

BC (2006) showed that the hypothesis of no Granger causality at frequency ω, or $M_{x\to y}(\omega) = 0$, can be imposed as two linear restrictions $R(\omega)\beta = 0$, where $\beta = (\beta_1, ..., \beta_p)'$ and

$$R(\omega) = \begin{bmatrix} \cos(\omega) & \cos(2\omega) & \cdots & \cos(p\omega) \\ \sin(\omega) & \sin(2\omega) & \cdots & \sin(p\omega) \end{bmatrix}$$

For a lag order of $p = 1$ or $p = 2$ there is only a trivial solution to this restriction, namely that $\beta_1 = \beta_2 = 0$. In these two cases therefore the hypothesis of Granger non-causality at a certain frequency $\omega \in (0, \pi)$ automatically implies the standard case of no Granger causality at any frequency. We therefore require a higher lag order, $p > 2$, in order to make a frequency-specific analysis interesting.

In practice the system (2.1) would often be augmented with further variables $\mathbf{z}_t$ to avoid spurious findings due to omitted variables, see BC for a discussion. Such an addition would lead to obvious augmentations of (2.2) with lagged (or in the case of exogenous variables, possibly contemporaneous and lagged) values $\mathbf{z}_t$, but would not affect our results in any other way. Therefore we focus on the bivariate case for ease of exposition.

3 A Beveridge-Nelson-type decomposition for specific frequencies

For our purposes it is useful to represent the null hypothesis of no Granger causality at frequency ω in a more convenient manner. Our representation is based on a decomposition that is similar to the well-known BN decomposition proposed by Beveridge and Nelson (1981) for the frequency $\omega = 0$. Let us first consider the test at frequency $\omega = 0$ (long-run causality). In this case the null hypothesis boils down to $\sum_{j=1}^{p} \beta_j = 0$. Following Dickey and Fuller (1979) we decompose the polynomial $\beta(L)$ as

$$\beta(L) = b_1^0 + (1-L)\gamma^0(L)$$

where $b_1^0 = \sum_{j=1}^{p} \beta_j$, $\gamma^0(L) = \gamma_0^0 + \gamma_1^0 L + \cdots + \gamma_{p-2}^0 L^{p-2}$ and $\gamma_j^0 = -\sum_{i=j+2}^{p} \beta_i$. Note that this decomposition is also used to obtain the Beveridge-Nelson decomposition. Accordingly, a test for causality at frequency $\omega = 0$ is equivalent to testing $b_1^0 = 0$ in the regression

$$y_t = c_y + \sum_{j=1}^{p} \alpha_j y_{t-j} + b_1^0 x_{t-1} + \sum_{k=1}^{p-1} \gamma_{k-1}^0 \Delta x_{t-k} + u_{y,t}. \tag{3.1}$$

In the following a similar approach is suggested for testing causality at frequencies $0 < \omega < \pi$. To this end we first present a suitable decomposition of the lag polynomial.

Lemma 1. *Let* $\beta(L) = 1 + \beta_1 L + \cdots + \beta_{p-1} L^{p-1}$ *with* $p \geq 3$. *Then for* $0 < \omega < \pi$ *there exists a representation of the form*

$$\beta(L) = b_1^\omega + b_2^\omega L + \gamma^\omega(L) \nabla_\omega(L) \tag{3.2}$$

where $\nabla_\omega(L) = 1 - 2\cos(\omega)L + L^2$ *and* $\gamma^\omega(L) = \gamma_0^\omega + \gamma_1^\omega L + \cdots + \gamma_{p-3}^\omega L^{p-3}$. *The* *gain function* $|\beta\left(e^{i\omega}\right)|^2$ *is zero at frequency* ω *if and only if* $b_1^\omega = 0$ *and* $b_2^\omega = 0$.

Proof. Comparing the coefficients at different lags yields the system of equations

$$\begin{pmatrix} 1 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2\cos(\omega) & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & -2\cos(\omega) & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & \ddots & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & -2\cos(\omega) & 1 \end{pmatrix} \begin{pmatrix} b_1^\omega \\ b_2^\omega \\ \gamma_0^\omega \\ \vdots \\ \gamma_{p-3}^\omega \end{pmatrix}$$

For $0 < \omega < \pi$ this linear system can be solved to obtain b_1^ω, b_2^ω and $\gamma_0^\omega, \ldots, \gamma_{p-3}^\omega$. Since $\nabla_\omega\left(e^{i\omega}\right) = \nabla_\omega\left(e^{-i\omega}\right) = 0$, the gain function results as

$$\begin{aligned} |\beta\left(e^{i\omega}\right)|^2 &= \beta\left(e^{i\omega}\right)\beta\left(e^{-i\omega}\right) \\ &= (b_1^\omega)^2 + 2b_1^\omega b_2^\omega \cos(\omega) + (b_2^\omega)^2. \end{aligned}$$

It follows that $|\beta\left(e^{i\omega}\right)|^2 = 0$ if and only if $b_1^\omega = b_2^\omega = 0$. □

Accordingly, (2.2) can be re-written as

$$y_t = c_y + \sum_{j=1}^{p} \alpha_j y_{t-j} + b_1^\omega x_{t-1} + b_2^\omega x_{t-2} + \sum_{k=1}^{p-2} \gamma_{k-1}^\omega \nabla_\omega(L) x_{t-k} + u_{y,t}, \qquad (3.3)$$

for $0 < \omega < \pi$. Note that this representation requires a lag order of $p \geq 3$. From Lemma 1 it follows that the transfer function possesses a zero at frequency ω if and only if $b_1^\omega = 0$ and $b_2^\omega = 0$. Accordingly, the hypothesis that x_t is a Granger cause of y_t at frequency ω is equivalent to the joint null hypothesis $H_0 : b_1^\omega = 0$ and $b_2^\omega = 0$ in the representation (3.3).

The corresponding representation for frequency $\omega = \pi$ is given by $\nabla_\pi = 1 + L$ and causality at this frequency can be tested by replacing the difference operator Δ in (3.1) with ∇_π. It is clear from (3.1) that in these two special cases $\omega = 0$ and $\omega = \pi$ the test has only one degree of freedom.

It may be worthwhile to point out that it is impossible to impose non-causality for *all* frequencies within an interval in the framework of the linear VAR model (3.3), because a (lag) polynomial can only have a finite number of roots, hence $b_1^{\omega} = b_2^{\omega} = 0$ cannot hold for infinitely many ω. In fact, in order to factor out a second $\nabla_{\omega^{**}\neq\omega^*}$ polynomial from the $\gamma^{\omega^*}(L)$ polynomial with another non-causal frequency ω^{**}, $p \geq 5$ would be required, and in general the number of non-causal frequencies is bounded by $(p-1)/2$.

4 Testing when the frequency is unknown

In many applications it is reasonable to assume that the frequency for which x_t is not a Granger cause for y_t is unknown but it is assumed that the frequency lies within some prespecified interval $\omega \in \Omega_0 = [\omega_\ell, \omega_u]$. Thus the relevant null hypothesis is

$$H_0^u: \text{ There exists a frequency } \omega \in [\omega_\ell, \omega_u] \text{ such that } |\beta(e^{i\omega})|^2 = 0.$$

Notice that the actual non-causal frequency can be regarded as a nuisance parameter which is only present under the null hypothesis. For testing such a null hypothesis it is natural to employ the minimum of the sequence of (Wald/LR/LM) test statistics for all test statistics associated with the grid of frequencies

$$\omega \in \Omega_0^{\delta} = \{\omega_\ell,\, \omega_\ell + \delta,\, \omega_\ell + 2\delta, \ldots,\, \omega_u\} \tag{4.1}$$

where δ denotes the frequency increment, say $(\omega_u - \omega_l)/T$. Let λ_T^{ω} denote the BC test statistic at frequency ω. The next proposition shows that asymptotically the significance level of the test can be controlled by using the usual critical value of the χ_2^2-distribution of the test for a known frequency.

Proposition 1. *Let λ_T^{ω} denote the Wald/LM/LR test statistic for Granger causality at frequency ω and $\lambda_T^* = \inf\{\lambda_T^{\omega} | \omega \in \Omega_0^{\delta}\}$ with δ inversely proportional to T, e.g. $\delta =$*

$(\omega_u - \omega_l)/T$. The $(1-\alpha)$ quantile of the χ^2 distribution with d degrees of freedom is denoted by $\chi^2_{d,\alpha}$. Under the null hypothesis that there exists at least one frequency $\omega^* \in \Omega_0$, with $\omega_l > 0$, $\omega_u < \pi$, such that $|\beta(e^{i\omega^*})|^2 = 0$, it holds that

$$\lim_{T\to\infty} P(\lambda_T^* > \chi^2_{2,\alpha}) \leq \alpha.$$

Proof. As shown by BC the statistic $\lambda_T^{\omega^*}$ for the simple test at the frequency ω^* has a χ^2 limiting distribution with 2 degrees of freedom. Accordingly we have $\lim_{T\to\infty} P(\lambda_T^{\omega^*} > \chi^2_{2,\alpha}) = \alpha$. Since $\omega^* \in \lim\limits_{\delta\to 0} \Omega_0^\delta$, as $T \to \infty$ we have

$$\lambda_T^* = \inf\{\lambda_T^{\omega_\ell}, \lambda_T^{\omega_\ell+\delta}, \lambda_T^{\omega_\ell+2\delta}, \ldots, \lambda_T^{\omega_u}\} \leq \lambda_T^{\omega^*}$$

and, therefore, $\lim_{T\to\infty} P(\lambda_T^* > \chi^2_{2,\alpha}) \leq \alpha$.

For causal frequencies $\omega \neq \omega^*$ in the interval Ω_0 (with $|\beta(e^{i\omega})|^2 > 0$) we have $\lambda_T^\omega = |O_p(T)|$ (cf. BC). The grid Ω_0^δ must therefore remain dense enough in order to contain tested frequencies in a $\sqrt{T}$-neighborhood of ω^*, which is ensured by the convergence rate T of the frequency increment δ. □

It follows that the size of the test is controlled by using the minimal test statistic in the interval $[\omega_\ell, \omega_u]$, effectively applying the test at the associated frequency as if this frequency were known. By its nature of using always the minimal statistic, i.e. the one least favorable for the alternative hypothesis, the test is expected to be conservative in general. This is a common property of tests with a nuisance parameter under the null.

The following corollary clarifies the extension to the special frequencies 0 and π where only a single restriction is tested and hence the limiting distribution has only one degree of freedom. It is again an application of the principle of using the test configuration least favorable to the alternative hypothesis.

Corollary 1. *Let τ_T^0 and τ_T^π denote the corresponding t-statistics for the hypotheses*

$b_1^0 = 0$ in (3.1), and $b_1^\pi = 0$ in (3.1) with ∇_π instead of Δ, respectively. We construct adjusted test statistics as

$$\lambda_T^0 = (\tau_T^0)^2 \chi_{2,\alpha}^2 / \chi_{1,\alpha}^2$$
$$\lambda_T^\pi = (\tau_T^\pi)^2 \chi_{2,\alpha}^2 / \chi_{1,\alpha}^2$$

The test for the set of frequencies Ω_0 that includes either $\omega = 0$ or $\omega = \pi$ can be performed by letting $\omega_u = 0$ or $\omega_l = \pi$ in Ω_0^δ with

$$\lim_{T\to\infty} P(\inf\{\lambda_T^\omega \,|\, \omega \in \Omega_0^\delta\} > \chi_{2,\alpha}^2) \;\leq\; \alpha$$

where Ω_0^δ is constructed as in Proposition 1.

Proof. If $\omega^* = 0$ is the true non-causal frequency, it was already shown in BC that $\lim_{T\to\infty} P\left((\tau_T^0)^2 > \chi_{1,\alpha}^2\right) = \alpha$. Multiplying the inequality through by $\chi_{2,\alpha}^2/\chi_{1,\alpha}^2$ yields $P\left(\lambda_T^0 > \chi_{2,\alpha}^2\right) = P\left((\tau_T^0)^2 > \chi_{1,\alpha}^2\right)$. As $plim_{T\to\infty} \inf\{\lambda_T^*, \lambda_T^0\} = \lambda_T^0$ we have $\lim_{T\to\infty} P\left(\inf\{\lambda_T^*, \lambda_T^0\} > \chi_{2,\alpha}^2\right) = \alpha$, and the corollary holds with equality. If the non-causal frequency is $\omega^* > 0$, then $\lambda_T^0 = |O_p(T)|$, cf. BC again, and thus $plim_{T\to\infty} \inf\{\lambda_T^*, \lambda_T^0\} = \lambda_T^*$, referring to the case of Proposition 1. The proof for the second case $\omega = \pi$ follows by analogy. □

Remark 1. Note that even if the null hypothesis band for the present test were to include all possible frequencies, $\Omega_0 = [0, \pi]$, the hypothesis is quite different from the traditional test of Granger non-causality. The traditional test requires non-causality at *all* frequencies under the null, while our test posits non-causality only at *some* (possibly unknown) frequency.

So far we have assumed a stable VAR system with all roots outside the unit circle. Considering the possibility of unit roots at frequency zero (real and positive unit roots), the analysis is unaffected by this type of non-stationarity if the considered frequency band under the null does not contain 0, i.e. $\omega_l > 0$. The complete analysis also extends naturally to the case with $I(1)$ variables that are not cointegrated, by differencing the corresponding variables and proceeding as before. However, if the variables are cointegrated, an additional assumption must be made to ensure valid inference, which is summarized in the following corollary.

Corollary 2. *Suppose the VAR* (2.1) *is not stable but that $n-r<n$ roots of $A(L)$ are unity, and we assume a maximum integration order of one, such that the system is $I(1)$ with cointegration rank $r>0$, and $(1,\kappa)'$ is the normalized cointegration vector, $\kappa \neq 0$, such that $x_t+\kappa y_t = e_t \sim I(0)$.*

Under a null hypothesis containing $\omega_l = 0$ the test in Corollary 1 for non-causality of x_t with respect to the target y_t is valid if the number of unstable roots of $A(L)$ is not affected by setting $b_1^0 = 0$ in (3.1). *A necessary (but in general not sufficient) condition for this validity is $b_{x,1}^0 \neq 0$ in the analogously re-written first equation of the system,*

$$x_t = c_x + \sum_{j=1}^{p} \alpha_{x,j} x_{t-j} + b_{x,1}^0 y_{t-1} + \sum_{k=1}^{p-1} \gamma_{x,k-1}^0 \Delta y_{t-k} + u_{x,t}.$$

Proof. If $b_1^0 = b_{x,1}^0 = 0$, the error-correction terms would be removed from the system, contradicting the assumption of cointegration by virtue of the Engle-Granger representation theorem. More generally, if the long-run properties of the system hinge on the non-zero value of b_1^0, then a test of $b_1^0 = 0$ would amount to a test under the null of no cointegration, with the associated non-standard features. However, if cointegration holds for any value b_1^0, then the coefficient b_1^0 is effectively attached to the $I(0)$ error term e_t and inference is standard. □

Since a test of no long-run causality between two $I(1)$ variables in both directions would be equivalent to a test of the null of no cointegration, in bivariate systems it would be invalid to perform two tests along the lines of Corollary 1, first testing the influence of x_t on y_t, and then vice versa. In higher-order systems a generalized necessary condition for the test validity could be formulated in terms of an unchanged rank r of the matrix of loading coefficients.

5 Inference on the phase shift

After establishing the existence of causality in a certain frequency band, the natural next step is to analyze in more detail the timing of the causal effect. In particular it is interesting to assess the time delay between the cause and effect at some prespecified frequency. To this end we adapt the concept of a phase shift $\phi(\omega)$ associated with some frequency ω.

To fix ideas, assume that the input signal is a pure sine wave $x_t = \sin(\omega t)$ and we are interested in measuring the phase shift a lag polynomial $\rho(L)$ implies to the input signal, that is,

$$\begin{aligned} y_t &= \rho(L)x_t \\ &= |\rho(e^{\omega i})| \sin[\omega t + \phi_\rho(\omega)] \\ &= |\rho(e^{\omega i})| \sin[\omega(t + \phi^*_\rho(\omega))] \end{aligned}$$

where $|\rho(e^{\omega i})|$ is the gain of the filter, $\phi_\rho(\omega)$ is the phase shift involved and $\phi^*_\rho(\omega)$ is the time delay, that is the phase shift measured in the number of time periods. From the vector autoregressive representation (2.2) we obtain

$$y_t = \frac{\beta^*(L)}{\alpha(L)} x_{t-1} + v_t \tag{5.1}$$

$$= \rho(L)x_t + v_t \tag{5.2}$$

where $\rho(L) = \beta^*(L)L/\alpha(L)$, $\alpha(L) = 1 - \sum_{j=1}^{p} \alpha_j L^j = 1 - \alpha^*(L)$ and $v_t = \alpha(L)^{-1}u_{y,t}$. For convenience the constant is suppressed. In the following lemma we present some useful results for the phase shift induced by the filter $\rho(L) = \beta(L)/\alpha(L)$, where $\beta(L) = \sum_{j=1}^{p} \beta_j L^j$, such that the correspondence $\beta^*_j = \beta_{j+1}$, $j = 0...p-1$, holds with $\beta^*(L) = \sum_{j=0}^{p-1} \beta_j^* L^j$.

Lemma 2. *Let $F_\alpha(\omega)$ and $F_\beta(\omega)$ denote the Fourier transforms of the filters $\alpha(L) = 1 - \sum_{j=1}^{p} \alpha_j L^j$ and $\beta(L) = \sum_{j=1}^{p} \beta_j L^j$, respectively. Furthermore, define $F_\rho(\omega) = F_\beta(\omega)/F_\alpha(\omega) = c_\rho(\omega) + i s_\rho(\omega)$, where $c_\rho(\omega)$ and $s_\rho(\omega)$ are given in the appendix.*

If the gains $|F_\alpha(\omega)|$ and $|F_\beta(\omega)|$ are non-zero at frequency $\omega \in [0, \pi]$ the phase shift is given by

$$\phi_\rho(\omega) = \arctan^*(s_\rho(\omega)/c_\rho(\omega), sgn[s_\rho(\omega)], sgn[c_\rho(\omega)]), \tag{5.3}$$

where $\arctan^*$ *is the four-quadrant version of the* arctan *function with a range of* $(0; 2\pi]$.

Proof. See appendix. □

Remark 2. The full circle could also be described with the function range $(-\pi, \pi]$ for another variant of the arctan function, but negative phase shifts are not meaningful in our application of a one-sided (backward-oriented) filter, hence we use $(0, 2\pi]$. Variants of the function $\arctan^*$ are available in some programming languages, for example the Matlab routine "atan2".

Remark 3. If at some frequency ω the term $s_\rho(\omega)$ switches its sign while $c_\rho(\omega) > 0$, the resulting phase shift function will display a discontinuous jump down from (or up to) 2π to (or from) a value arbitrarily close to zero. The implied delay function will have a corresponding jump between $2\pi/\omega$ and zero. The reason is that the phase shift in principle is only identified up to adding integer multiples of 2π, but the standard definition in Lemma 2 maps all phases into the interval $(0, 2\pi]$. We remove these discontinuities in the phase shift function by adding or subtracting integer multiples of 2π where needed, which is sometimes referred to as "phase unwrapping". In the following we denote this adjusted measure as $\phi_{uw}(\omega)$, and the corresponding delay measure as $\phi^*_{uw}(\omega) = \phi_{uw}(\omega)/\omega$. However, the unwrapping procedure is independent of the estimation of the phase shift or delay, hence it does not affect our analysis of the sampling uncertainty of the locally identified measures.

Notice that $\omega = \pi$ implies $s_\rho(\pi) = 0$ and, therefore, the phase will be identical to 2π (i.e. the delay is identical to 2 periods) irrespective of the values of α_j or β_j.[2] Also, the case $c_\rho(\omega) = s_\rho(\omega) = 0$ is not addressed in Lemma 2 explicitly because in this case the phase shift is not defined due to vanishing gains.

The following proposition analyzes how the estimation error of the coefficients α_j and β_j affects the uncertainty of the estimated time delay.

Proposition 2. *If the phase shift exists as given in Lemma 2, then the asymptotic distribution of the delay estimate $\hat{\phi}^*_\rho(\omega)$ at a frequency $\omega \in (0, \pi)$ is given as:*

$$\sqrt{T}\left(\hat{\phi}^*_\rho(\omega) - \phi^*_\rho(\omega)\right) \xrightarrow{d} \mathsf{N}(0, \omega^{-2} J_\rho(\omega)' V_r J_\rho(\omega)), \tag{5.4}$$

where

$$J_\rho(\omega) = \left[\frac{\mathbf{v}'_{s,p}(\omega)c_\beta(\omega) - \mathbf{v}'_{c,p}(\omega)s_\beta(\omega)}{|F_\beta(\omega)|^2}, \frac{\mathbf{v}'_{s,p}(\omega)(1 - c_{\alpha^*}(\omega)) + \mathbf{v}'_{c,p}(\omega)s_{\alpha^*}(\omega)}{|F_\alpha(\omega)|^2}\right]',$$

and $\mathbf{v}_{s,p}(\omega)$, $\mathbf{v}_{c,p}(\omega)$, $c_\beta(\omega)$, $s_\beta(\omega)$, $c_{\alpha^*}(\omega)$, and $s_{\alpha^*}(\omega)$ are given in the appendix. The variance-covariance matrix of the consistent estimate of the $2p$ coefficient vector $\mathbf{r} = (\beta', \alpha')' = (\beta_1, ..., \beta_p, \alpha_1, ..., \alpha_p)'$ is denoted by V_r. In practice, $J_\rho(\omega)$ and V_r can be replaced by consistent estimates $\hat{J}_\rho(\omega)$ and $\hat{V}_r$.

Proof. See appendix. □

The limit distribution given in Proposition 2 can be used to construct approximate and pointwise confidence intervals for the estimated delay at various frequen-

[2] Again we resolve the global identification deficiency of the phase shift explicitly, by attributing the case $s_\rho(\omega) = 0$ to a phase shift $\phi(\omega) = 2\pi$ rather than $\phi(\omega) = 0$. The reason for this choice is that the underlying ARDL model is purely backward-oriented without contemporaneous terms.

cies. In practical applications we recommend to check numerically that $|\hat{F}_\alpha(w)|^2$ and $|\hat{F}_\beta(w)|^2$ are bounded away from zero in a neighborhood $w \approx \omega$.

6 Monte Carlo Simulations

First we assess the empirical characteristics of the frequency domain causality tests by means of Monte Carlo experiments. The regressor is generated by a univariate exogenous AR(1) process,

$$x_t = \alpha_{1,x} x_{t-1} + u_{x,t},$$

with two degrees of persistence, $\alpha_{1,x} \in \{0, 0.8\}$. To analyze the size properties x does not cause y at a frequency $\omega \in \{0, 0.39, \pi/2\}$ in the model

$$y_t = \alpha_{1,y} y_{t-1} + \gamma_0 \left((x_{t-1} - x_{t-2}) + \frac{1}{2}(x_{t-2} - x_{t-3}) \right) + u_{y,t} \quad \text{for } \omega^* = 0$$
$$y_t = \alpha_{1,y} y_{t-1} + \gamma_0 (x_{t-1} - 2\cos(\omega^*) x_{t-2} + x_{t-3}) + u_{y,t} \quad \text{for } \omega^* > 0$$

where again $\alpha_{1,y} \in \{0, 0.8\}$ and $\gamma_0 \in \{-1, 0.5, 10\}$. The innovations are uncorrelated Gaussian white noise with normalized variance, $u_t \sim NID(0, I_2)$.

In table 1 we report the frequency bands that are considered as null hypotheses for the simulation of the size of the test. For the grid of tested frequencies we evenly distribute T points from 0 to π and test at all points that lie inside the null band. For the analysis of the power of the test we confine ourselves to situations where x is still non-causal at some frequency, but this frequency now lies outside the null band. In addition one could specify a true DGP without any non-causality. In the third column of table 1 we report the analyzed frequency bands for the simulated power of the test.

In table 2 we have collected the simulation results. It is apparent that the test is quite conservative in general and the empirical rejection frequencies under the null do not attain the nominal significance level of 5% for the considered sample sizes

Table 1: Specified frequency bands for the simulation of size and power

True non-causal frequency ω^*	Bands for H_0^u (size)	Bands for H_0^u (power)
0	$[0, 0.2]$	$[0.2, 0.79]$, $[0.79, \pi]$
0.39	$[0.2, 0.79]$	$[0, 0.2]$, $[0.79, \pi]$
$\pi/2$	$[0.79, \pi]$	$[0, 0.2]$, $[0.2, 0.79]$

Notes: The frequency 0.39 corresponds to approximately 16 periods wavelength, i.e. 4 years for quarterly data, frequency 0.20 means approximately 32 periods (8 years), and frequency 0.79 translates into 8 periods (2 years).

of $T = 200$ and $T = 5000$, except when the true non-causal frequency is 0. On the other hand the size distortions are not dramatic, with the empirical size remaining above 1% in all cases. This is also reflected in the satisfactory power characteristics of the test. The only problem occurs when the null hypothesis band is specified as $[0, 0.2]$ while the true non-causal frequency is close by at 0.39, the data are noisy ($\alpha_{1,x} = \alpha_{1,y} = 0$), and the impact is limited ($|\gamma_0| \leq 1$). In this quite extreme case the power may drop below 10% for a sample size of 200, and converges towards unity only slowly. All in all, however, the performance of the test is good.

7 Empirical Illustration

7.1 Assessing the Greenhouse effect

First we consider an empirical example from environmental science. Specifically we apply the Granger causality test to the annual time series of greenhouse gas emissions and US temperature from 1895 to 2013. The temperature data are from the US National Climatic Data Center (Climate at a Glance), while the CO2 emissions data are the "Total" series from Boden, Marland, and Andres (2014) ranging from 1751 to 2010. The jointly available sample is thus 1895-2010 ($T = 116$) and

Table 2: Simulation results

	$T = 200$					
	$\alpha_{1,x} = \alpha_{1,y} = 0$			$\alpha_{1,x} = \alpha_{1,y} = 0.8$		
Frequency band under H_0	$\gamma_0 = -1$	$\gamma_0 = 0.5$	$\gamma_0 = 10$	$\gamma_0 = -1$	$\gamma_0 = 0.5$	$\gamma_0 = 10$
	True frequency $\omega^* = 0$					
[0, 0.2]	.058	.047	.049	.060	.062	.054
[0.2, 0.79]	1	1	1	1	1	1
[0.79, π]	1	.953	1	.994	.705	1
	True frequency $\omega^* = 0.39$					
[0.2, 0.79]	.015	.025	.017	.013	.020	.018
[0, 0.2]	.088	.044	1	.519	.162	1
[0.79, π]	.959	.469	1	1	.980	1
	True frequency $\omega^* = \pi/2$					
[0.79, π]	.015	.014	.023	.013	.013	.023
[0, 0.2]	1	1	1	1	1	1
[0.2, 0.79]	1	1	1	1	1	1

	$T = 5000$					
	$\alpha_{1,x} = \alpha_{1,y} = 0$			$\alpha_{1,x} = \alpha_{1,y} = 0.8$		
Frequency band under H_0	$\gamma_0 = -1$	$\gamma_0 = 0.5$	$\gamma_0 = 10$	$\gamma_0 = -1$	$\gamma_0 = 0.5$	$\gamma_0 = 10$
	True frequency $\omega^* = 0$					
[0, 0.2]	.051	.049	.053	.051	.052	.050
[0.2, 0.79]	1	1	1	1	1	1
[0.79, π]	1	1	1	1	1	1
	True frequency $\omega^* = 0.39$					
[0.2, 0.79]	.013	.015	.014	.012	.014	.014
[0, 0.2]	.977	.486	1	1	.999	1
[0.79, π]	1	1	1	1	1	1
	True frequency $\omega^* = \pi/2$					
[0.79, π]	.013	.015	.013	.013	.018	.017
[0, 0.2]	1	1	1	1	1	1
[0.2, 0.79]	1	1	1	1	1	1

Notes: Empirical rejection frequencies, nominal significance level 0.05, 5000 replications. The value "1" means unity up to a precision of six decimal digits. Power is raw (not size-adjusted). Number of frequencies in the respective grid is (excluding special cases 0 and π) $\frac{\omega_u - \omega_l}{\pi} T$. The case $\omega^* = 0$ is created as $\gamma_0(1 - 0.5L - 0.5L^2) = \gamma_0(1 - L + 0.5(L - L^2))$, see the text.

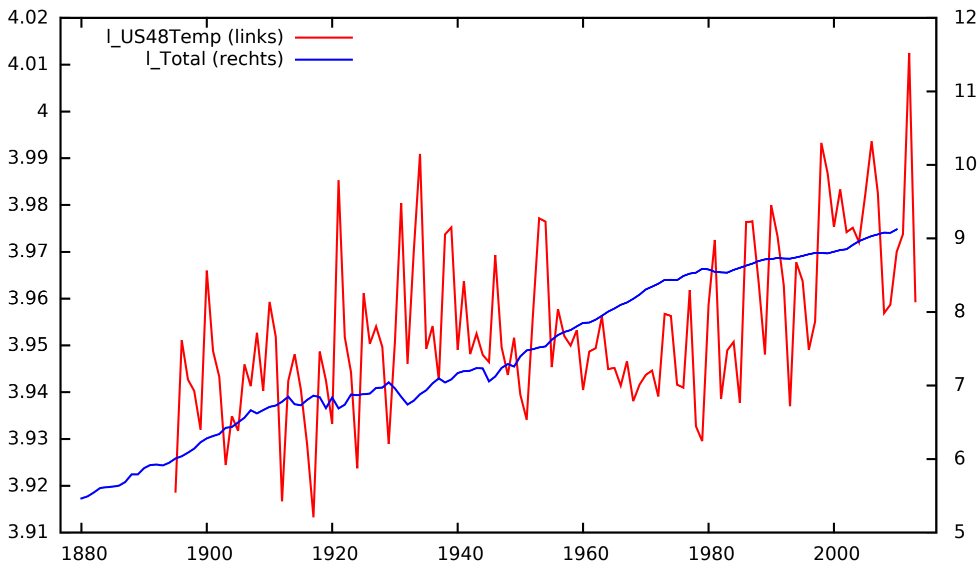

Figure 7.1: Time series of US continental (48 states) temperatures (log degrees Fahrenheit) and total CO2 emissions (log millions of metric tons).

we use log transforms.

First we determine the lag order of the bivariate VAR in log-levels. The Akaike information criterion suggests just two lags, but the third lag is also significant at the 10% level, and as explained above at least three lags are needed in order to distinguish causality at different frequencies. Hence we choose $p = 3$. In principle it would be possible to consider a more complicated lag structure, restricting some of the intermediate lag coefficients in certain equations to zero, but we do not pursue this strategy in this illustration.

For these data it is natural to suspect cointegration so we run the Johansen test, with an unrestricted constant to deal with the trending data. The highest eigenvalue is 0.16 and the p-value of the trace test of no cointegration yields 0.0098, such that there is evidence for cointegration at the nominal 1% significance level. Notice that the error correction term is insignificant in the emissions equation (p-value of 0.12), hence emissions do not seem to be caused by temperatures at frequency zero. This is plausible, but it also means that we cannot test the restriction of no long-run causality running in the other direction –from emissions on temperatures– without affecting the cointegration property of the system, as discussed above. Therefore the causality at frequency zero is already established, and our interpretation focuses on non-zero frequencies. Figure 7.2 shows the frequency-wise test results.

For any frequency band up to roughly 0.2, corresponding to wavelengths down to roughly 31 periods (years) the minimal test statistics would exceed the critical value, and hence for those frequency bands we would reject the null hypothesis that there exists a frequency without Granger causality. In the plot we have included a null hypothesis band that extends down to a wavelength of 40 (years); given the inverse relationship between frequency and wavelength small variations of the frequency mean relatively large absolute changes of the implied cycle lengths. Of course the frequency band below 0.2 is very close to zero, and with this effective sample of $T = 113$ it is very difficult to distinguish cycles of 30 periods from even lower frequencies. Hence some leakage from the zero frequency is expected. For any frequency bands containing higher frequencies (shorter wavelengths) we would not be able to reject the corresponding null hypothesis. As an example we have included a second possible null hypothesis band, covering the frequencies that correspond to 5 to 10 years of oscillation. The overall conclusion is thus that Granger causality from emissions to temperatures varies across frequencies.

7.2 The leading indicator properties of new orders

The next empirical example that we consider are the Granger-causal effects on German industrial production growth gIP_t originating in (growth of) new orders received by German firms from abroad, gAA_t. The monthly time index of these series refers to the real-time date of actual publication.[3] The sample runs from 1995m10 to 2012m12 ($T = 207$), and we fit an ARDL(4,4) model to these data (t-ratios below

[3]For example, gIP_t has a publication lag of one month, so refers to activity in period $t-1$. This aspect is not important for the present analysis, however, which deals only with predictive content of the available information in real time.

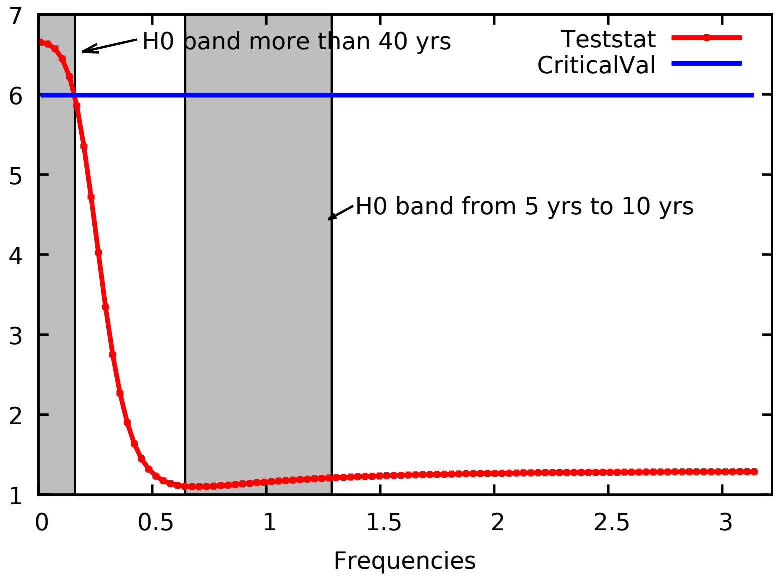

Figure 7.2: Frequency-wise causality test from log total emissions on log US continental temperatures. System with 3 lags. The horizontal line is the critical value of the χ^2 distribution with two degrees of freedom at the 5% level. The lowest tested frequency here is 0.01, see the text for the zero frequency.

point estimates):

$$\begin{aligned} gIP_t \quad &= \quad \underset{-1.44}{-0.0017} + (\underset{3.72}{0.12}L + \underset{3.80}{0.13}L^2 + \underset{4.81}{0.17}L^3 + \underset{3.86}{0.13}L^4)gAA_t + \\ & \quad (\underset{-3.33}{-0.25}L - \underset{-2.44}{0.19}L^2 - \underset{-0.92}{0.07}L^3 - \underset{-2.35}{0.16}L^4)gIP_t + \hat{u}_t \\ \bar{R}^2 &= 0.15 \quad DW = 2.02 \end{aligned}$$

As a preliminary step we establish that there actually exists G-causality at some frequencies, with the clearcut result shown in Figure 7.3. However, the fact of weak causality roughly between frequencies 1.5 and 2.2 means that any results about the lead-lag relationship in that band should be interpreted with caution.

Next we proceed to the point estimates of the frequency-specific time delays in Figure 7.4, calculated from the estimated polynomials $\hat{\alpha}(L)$ and $\hat{\beta}(L)$. The delay of industrial production (or lead of foreign orders) is almost two months at the long-run frequencies and is rising to roughly three months around frequency $\pi/2$ (wavelength four months). At higher frequencies the delay measures are even somewhat

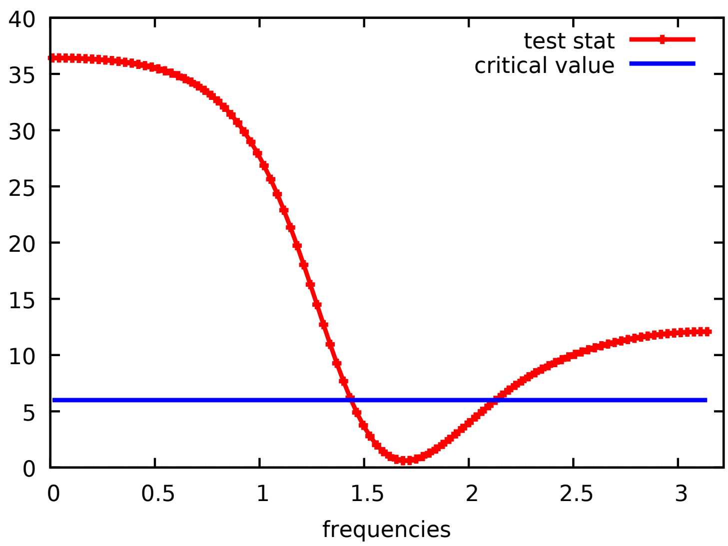

Figure 7.3: Sequence of BC tests on (growth of) German industrial production (target) and new foreign orders (cause), without including the endpoints 0 and π. The critical value refers to two degrees of freedom and is appropriate for the interior case $\omega \in (0, \pi)$, cf. section 4.

higher, at least after the unwrapping procedure described in section 5.

As we discussed in section 5, the phase is not well-defined everywhere and the squared gain of the involved filters should be checked numerically whether or where it vanishes. This is done for the present example in Figure 7.5. It can be seen that roughly up to frequency 1.3 both sequences of squared absolute values are reasonably far away from zero. Also, we already saw in Figure 7.3 that in the range of roughly 1.5 to 2.2 the BC test statistics are below the critical value. This finding is directly reflected in the gain function $|\hat{F}_\beta(\omega)|^2$ which is extremely small in this range. Therefore we expect the delta method to become problematic for frequencies higher than perhaps 1.3, and to break down completely between 1.5 and 2.2 in this case. Given that $\omega = 1.3$ corresponds to a wavelength of roughly five months, the frequency ranges most interesting for business-cycle analysts are not affected by these problems here.

Our results for the sequence of confidence intervals constructed with the delta method are reported in Figure 7.6. For illustrative purposes we calculate the un-

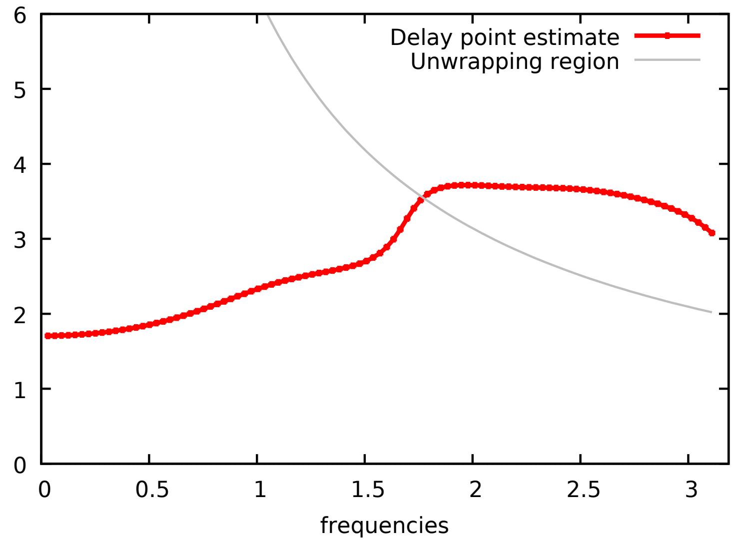

Figure 7.4: Time delay point estimates for the effect of gAA_t on gIP_t. Delays north-east of the thin line are uniquely identified with "unwrapping" the phase shift, i.e. avoiding discontinuities of the delay curve. This corresponds to $\phi^*_{uw}(\omega)$ in Remark 3.

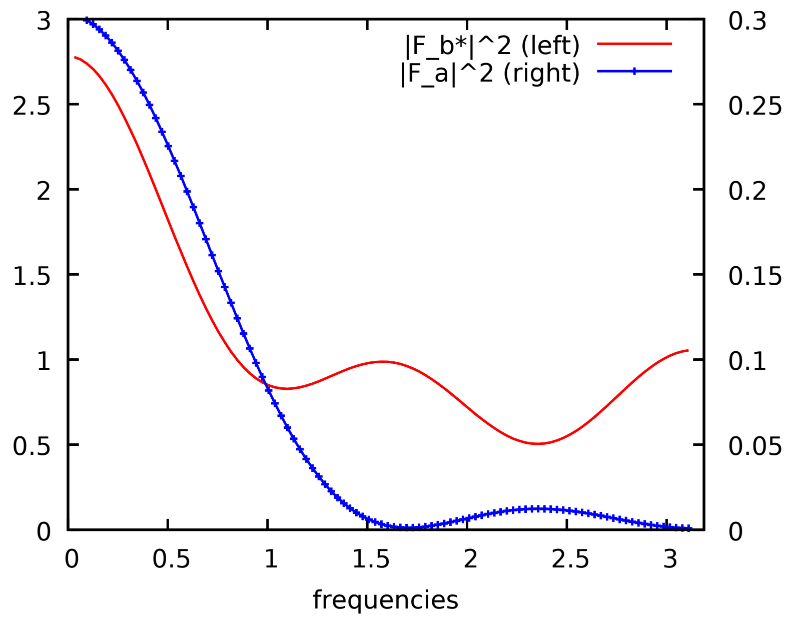

Figure 7.5: Numerical check of whether $|\hat{F}_\alpha(\omega)|^2 = 0$ and/or $|\hat{F}_\beta(\omega)|^2 = 0$. (Here "|F_b*|^2" refers to $|\hat{F}_\alpha(\omega)|^2$ and "|F_a|^2" corresponds to $|\hat{F}_\beta(\omega)|^2$.)

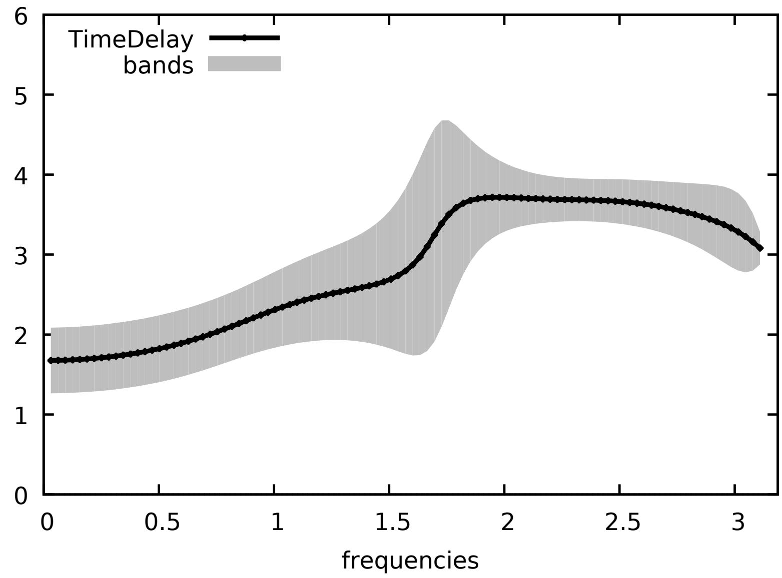

Figure 7.6: Delta-method confidence intervals for the time delay (pointwise, nominal 95% coverage). Y-axis in months. All frequencies $\omega \in (0, \pi)$ are shown, but see the text for regions of near-nonexistence.

certainty measures for all frequencies, bearing in mind our previous checks that suggested problems of near nonexistence for $1.3 < \omega < 2.2$. The point estimates are the same as in Figure 7.4. The time delays are significantly larger than one month throughout. The width of the (pointwise) confidence intervals does in fact not depend much on the frequencies down to a wavelength of roughly six months ($\omega \approx 1$). In contrast, for the problematic frequency range in the neighborhood of 1.7 the mechanically calculated sampling uncertainty almost explodes, highlighting the importance of the numerical pre-checks.

Finally we want to assess whether the method from Proposition 2 performs well in terms of the empirical coverage of the constructed confidence intervals. To this end we run a small simulation study based on the current illustration. That is, we assume a DGP that is essentially based on the estimated ARDL(4,4) model, holding the gAA_t variables fixed as exogenous.[4] The estimated residual variance is also directly used to complete the DGP specification, and we assume the innovations to

[4]The only change with respect to the estimated model is that the DGP equation does not contain a constant term, since that was insignificant.

be Gaussian white noise. We then draw many times randomly from this assumed distribution of innovations and simulate the dynamic ARDL equation forward each time. In each simulation run we re-estimate the equation on the simulated data (including a constant) and use the result from Proposition 2 to construct confidence intervals as described above. Each time we record whether the confidence intervals cover the true delays implied by the DGP. The results are shown in Figure 7.7. Again it can be verified that the method is problematic in the neighborhood of frequencies where the numerical pre-checks fail, i.e. where any of the functions shown in Figure 7.5 is close to zero such that the phase shift is close to being non-existent. This also applies to the neighborhood of the frequency π; first of all it turned out that $|\hat{F}_\beta(\omega \approx \pi)|^2 \approx 0$, and secondly we had mentioned that at $\omega = \pi$ the estimated phase will necessarily have a degenerate distribution and be identical to 2π. However, in the interesting frequency range (roughly below 1.3) the coverage is close to its nominal value of 95%.

We conclude that this method of measuring the sampling uncertainty of the time delay estimates can be recommended for applied work, provided the caveat is borne in mind that the confidence bands may not exist everywhere, and thus the described pre-checks should be viewed as an integral part of any application.

8 Conclusion

In this paper we have shown that tests of Granger non-causality can also be specified in terms of frequency bands or intervals instead of single frequency points. We propose a rigorous framework enabling standard inference that circumvents the ad-hoc procedures of joint testing with unknown statistical properties. The implementation is easy because in practice the relevant test statistic is just the minimum over a pre-specified frequency band, apart from a special but straightforward treatment of the frequencies 0 and π. In a simulation study the test performed satisfactorily albeit

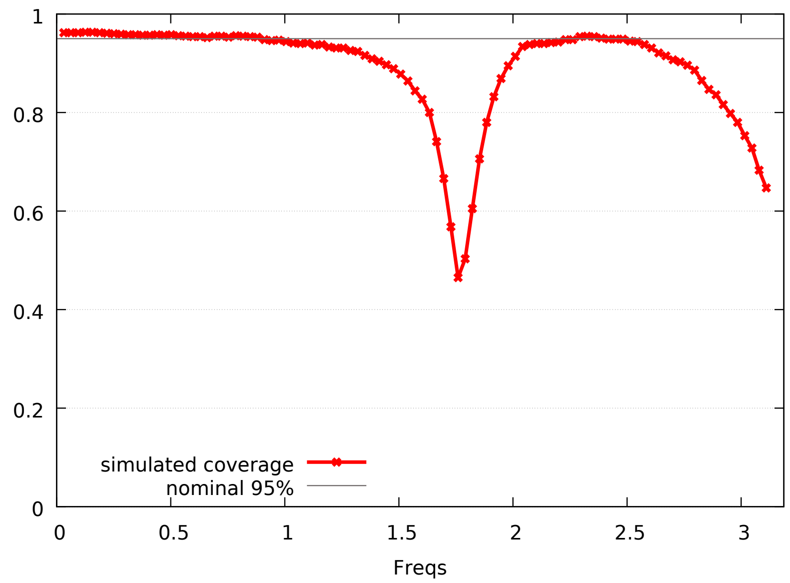

Figure 7.7: Simulation of the coverage of the confidence intervals based on the delta method (nominal coverage 95%). Fraction of 1000 simulation runs. All frequencies $\omega \in (0, \pi)$ are shown, but see the text for regions of near-nonexistence.

slightly conservatively.

Given that strict non-causality over a range of frequencies is impossible in this (linear) framework except if there is no causality at all, accepting the null hypothesis still means that some causality exists in the band of the null hypothesis. For practical purposes it may therefore be advisable to keep the specified frequency band reasonably short.

As a complementary piece of information concerning the frequencies with non-vanishing causality we have proposed additional tools to analyze the time delay of the target variable relative to the cause, based on the standard cross-spectral phase shift analysis. In particular we used our parametric framework to construct confidence intervals for the estimated delay measures with the delta method. These asymptotic confidence intervals may not be well-defined for a certain (finite) number of frequency points, depending on the coefficients of the underlying model. In a finite-sample setting the neighborhood of these frequency points is likely to be affected as well, resulting in a lower quality of the confidence bands (in terms of actual coverage probabilities). However, in practice it is straightforward to check for and find the location of these neuralgic regions prior to constructing the confidence intervals in the remaining frequency regions.

Our first empirical application with long time series of CO2 emissions and earth surface temperatures demonstrated that varying degrees of Granger causality in the frequency domain are of practical relevance. In another application we demonstrated the sampling uncertainty of the delays of German industrial production growth with respect to the foreign orders indicator, across frequencies. In the introduction we already mentioned the case of the term structure of interest rates where such varying connections are also expected. In addition, according to the economic hypothesis of consumption smoothing a similar result about differing impacts of short- versus long-term fluctuations might hold between income and consumption. We believe that many more potential applications in economics and perhaps other disciplines are likely to exist.

References

ASSENMACHER-WESCHE, K., AND S. GERLACH (2007): "Money at Low Frequencies," *Journal of the European Economic Association*, 5(2-3), 534–542.

——— (2008a): "Interpreting euro area inflation at high and low frequencies," *European Economic Review*, 52, 964–986.

——— (2008b): "Money growth, output gaps and inflation at low and high frequency: Spectral estimates for Switzerland," *Journal of Economic Dynamics and Control*, 32, 411–435.

BACCALÁ, L. A., AND K. SAMESHIMA (2001): "Partial directed coherence: a new concept in neural structure determination," *Biological Cybernetics*, 84, 463–474.

BEVERIDGE, S., AND C. R. NELSON (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics*, 7, 151–174.

BODEN, T. A., G. MARLAND, AND R. J. ANDRES (2014): "Global, Regional, and National Fossil-Fuel CO2 Emissions," .

BREITUNG, J., AND B. CANDELON (2006): "Testing for short- and long-run causality: A frequency-domain approach," *Journal of Econometrics*, 132, 363–378.

DICKEY, D. A., AND W. A. FULLER (1979): "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 74, 427–431.

GEWEKE, J. (1982): "Measurement of Linear Dependence and Feedback Between Multiple Time Series," *Journal of the American Statistical Association*, 77, 304–324.

GRANGER, C. W. J. (1969): "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, 37, 424–438.

LEMMENS, A., C. CROUX, AND M. G. DEKIMPE (2008): "Measuring and testing Granger causality over the spectrum: An application to European production expectation surveys," *International Journal of Forecasting*, 24, 414–431.

SHILLER, R. J. (1979): "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," *Journal of Political Economy*, 87(6), 1190–1219.

WEI, Y. (2015): "The informational role of commodity prices in formulating monetary policy: a reexamination under the frequency domain," *Empirical Economics*, 49(2), 537–549.

A Proofs

For easier readability in this appendix we suppress the dependence of the various functions on the frequency ω, which should be clear from the definitions in the main text.

A.1 Proof of Lemma 2

Lemma 2 mainly translates various results from established signal processing analysis to our setting. The causal recursive filter implied by the ARDL model (5.1) is nothing else than the following rational polynomial in the lag operator:

$$\rho(L) = \frac{\beta(L)}{1-\alpha^*(L)} = \frac{\sum_{j=1}^{p} \beta_j L^j}{1-\sum_{j=1}^{p} \alpha_j L^j} \tag{A.1}$$

The frequency-specific response function is given by the division of the Fourier transforms of the component filters:

$$F_\rho = \frac{F_\beta}{1-F_{\alpha^*}} = \frac{\sum_{j=1}^{p} \beta_j e^{i\omega j}}{1-\sum_{j=1}^{p} \alpha_j e^{i\omega j}}, \tag{A.2}$$

Note that we can also write this with $\alpha(L) = 1-\alpha^*(L)$, so $\alpha_0 = 1$ and $\alpha_{j>0} = -\alpha_j^*$, and correspondingly $F_\rho = F_\beta / F_\alpha$. With Euler's formula ($e^{i\theta} = \cos\theta + i\sin\theta$) we have $F_\beta = |F_\beta|\left(\cos(\phi_\beta) + i\sin(\phi_\beta)\right)$, or alternatively

$$F_\beta = \sum_{j=1}^{p} \beta_j \cos(\omega j) + i\sum_{j=1}^{p} \beta_j \sin(\omega j) \equiv c_\beta + i s_\beta,$$

such that it also holds that $c_\beta = |F_\beta|\cos\left(\phi_\beta\right)$ and $s_\beta = |F_\beta|\sin\left(\phi_\beta\right)$.

We can represent the denominator filter part in an analogous way:

$$F_\alpha = 1-\sum_{j=1}^{n} \alpha_j \cos(\omega j) - i\sum_{j=1}^{n} \alpha_j \sin(\omega j) \equiv 1 - c_{\alpha^*} - i s_{\alpha^*}$$

And generically: $F_\alpha = |F_\alpha|\left(\cos(\phi_\alpha) + i\sin(\phi_\alpha)\right)$. So we have:

$$\begin{aligned} |F_\alpha|\cos(\phi_\alpha) &= 1 - c_{\alpha^*} \\ |F_\alpha|\sin(\phi_\alpha) &= -s_{\alpha^*} \end{aligned}$$

For F_ρ at each frequency we can therefore write

$$\frac{c_\beta + i s_\beta}{1 - c_{\alpha^*} - i s_{\alpha^*}},$$

provided the denominator does not vanish. The complex-number division yields:

$$F_\rho = |F_\alpha|^{-2}(c_\rho + i\,s_\rho), \tag{A.3}$$

where $c_\rho \equiv c_\beta(1-c_{\alpha^*}) - s_\beta s_{\alpha^*}$ and $s_\rho \equiv s_\beta(1-c_{\alpha^*}) + c_\beta s_{\alpha^*}$.

The imaginary part of this frequency response function is $|F_\alpha|^{-2}s_\rho$, and the real part $|F_\alpha|^{-2}c_\rho$. The phase shift is given by the angle of this complex number in polar-coordinate form, thus $\tan\phi_\rho = s_\rho/c_\rho$ for $c_\rho \neq 0$. The four-quadrant refined arctan* function provides the necessary information: For example, if $c_\rho < 0$ the number $c_\rho + i\,s_\rho$ lies in the upper-left or lower-left quadrants and therefore the phase shift would be between $\frac{\pi}{2}$ and $\frac{3}{2}\pi$. Accordingly the correct arctan* value in these cases would be obtained as $\arctan(s_\rho/c_\rho) + \pi$. If instead $c_\rho > 0$ together with $s_\rho \leq 0$, then the lower-right quadrant is concerned with $\phi_\rho \in (\frac{3}{2}\pi, 2\pi]$, and we obtain the shift value as $\arctan(s_\rho/c_\rho) + 2\pi$. In the upper-right quadrant with $s_\rho > 0$, $c_\rho > 0$ the standard calculation by $\arctan(s_\rho/c_\rho) \in (0, \frac{\pi}{2})$ remains unchanged. In addition, when $c_\rho = 0$ one could define the phase shift as $\frac{\pi}{2}$ or $3\frac{\pi}{2}$ (depending on the sign of s_ρ) to close the gap in the domain of the arctan function. Figure A.1 displays the resulting function graph. It can be seen that whenever the phase is wrapped from 0 to 2π, i.e. when $s_\rho(\omega) = 0$ together with $c_\rho(\omega) > 0$, the phase shift function is not continuous and therefore not differentiable per se. However, after unwrapping the phase this discontinuity vanishes, and since the slope of the function is implicitly given by the derivative of the standard arctan function at point 0, the gradient of the unwrapped phase is finite and continuous.

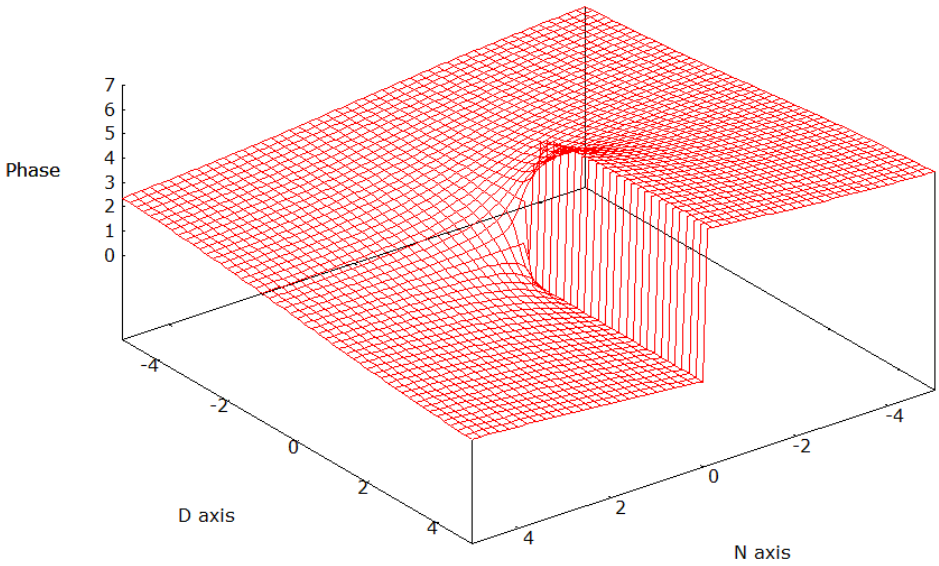

Figure A.1: Phase shift as a function of the terms $s_\rho(\omega)$ ("N axis") and $c_\rho(\omega)$ ("D axis"), without "unwrapping".

A.2 Proof of Proposition 2

For any integer z we define

$$\mathbf{v}_{s,z} = (\sin(\omega), ..., \sin(z\omega))'$$
$$\mathbf{v}_{c,z} = (\cos(\omega), ..., \cos(z\omega))',$$

We have the $2p$ coefficient vector $\mathbf{r} = (\beta', \alpha')'$ such that $\beta = (\mathbf{I}, \mathbf{0})\mathbf{r}$ and $\alpha = (\mathbf{0}, \mathbf{I})\mathbf{r}$, where the zero matrix $\mathbf{0}$ is of dimension $p \times p$ and $\mathbf{I} = \mathbf{I}_p$.[5] The estimate $\hat{\mathbf{r}}$ comes with a (consistently estimated) variance-covariance matrix V_r.

Assuming an interior solution the arctan* function behaves like the standard arctan function, therefore the variance of the estimate $\hat{\phi}_\rho$ can be inferred from V_r with the delta method by using the derivative $\partial \arctan(s_\rho / c_\rho)/\partial \mathbf{r}'$ as the relevant Jacobian J_ρ. It turns out that $c_\rho = 0$ is automatically accommodated, such that as

[5] In a generalized case with differing lag lengths p_x and p_y for the x- and y-terms the dimensions can be easily adjusted.

a formal workaround for the non-existence of $\arctan(s_\rho/c_\rho)$ in that case we can instead take the limit $J_\rho = \lim_{\varepsilon\to 0} \partial \arctan(s_\rho/(c_\rho+\varepsilon))/\partial \mathbf{r}'$.

Differentiating $\arctan(s_\rho/c_\rho)$ yields:

$$J_\rho = \left(1+\frac{s_\rho^2}{c_\rho^2}\right)^{-1} c_\rho^{-2}\left(c_\rho \frac{\partial s_\rho}{\partial \mathbf{r}} - s_\rho \frac{\partial c_\rho}{\partial \mathbf{r}}\right),$$

with

$$\begin{aligned} \frac{\partial s_\rho}{\partial \mathbf{r}'} &= (\mathbf{I},\mathbf{0})'\mathbf{v}_{s,p}(1-c_{\alpha^*}) - s_\beta(\mathbf{0},\mathbf{I})'\mathbf{v}_{c,n} + (\mathbf{I},\mathbf{0})'\mathbf{v}_{c,p}s_{\alpha^*} + c_\beta(\mathbf{0},\mathbf{I})'\mathbf{v}_{s,p} \\ &= (\mathbf{I},\mathbf{0})'\left(\mathbf{v}_{s,p}(1-c_{\alpha^*}) + \mathbf{v}_{c,p}s_{\alpha^*}\right) + (\mathbf{0},\mathbf{I})'\left(\mathbf{v}_{s,p}c_\beta - \mathbf{v}_{c,p}s_\beta\right), \end{aligned}$$

$$\begin{aligned} \frac{\partial c_\rho}{\partial \mathbf{r}'} &= (\mathbf{I},\mathbf{0})'\mathbf{v}_{c,p}(1-c_{\alpha^*}) - c_{\beta^*}(\mathbf{0},\mathbf{I})'\mathbf{v}_{c,p} - \left((\mathbf{I},\mathbf{0})'\mathbf{v}_{s,p}s_{\alpha^*} + s_\beta(\mathbf{0},\mathbf{I})'\mathbf{v}_{s,p}\right) \\ &= (\mathbf{I},\mathbf{0})'\left(\mathbf{v}_{c,p}(1-c_{\alpha^*}) - \mathbf{v}_{s,p}s_{\alpha^*}\right) - (\mathbf{0},\mathbf{I})'\left(\mathbf{v}_{c,p}c_\beta + \mathbf{v}_{s,p}s_\beta\right). \end{aligned}$$

Thus:

$$\begin{aligned} J_\rho &= \frac{1}{c_\rho^2+s_\rho^2}\times \\ &\quad \{(\mathbf{I},\mathbf{0})'\left((\mathbf{v}_{s,p}c_\rho - \mathbf{v}_{c,p}s_\rho)(1-c_{\alpha^*}) + (\mathbf{v}_{c,p}c_\rho + \mathbf{v}_{s,p}s_\rho)s_{\alpha^*}\right) + \\ &\quad (\mathbf{0},\mathbf{I})'\left((\mathbf{v}_{s,p}c_\beta - \mathbf{v}_{c,p}s_\beta)c_\rho + (\mathbf{v}_{c,p}c_\beta + \mathbf{v}_{s,p}s_\beta)s_\rho\right)\} \end{aligned} \tag{A.4}$$

Notice that $c_\rho^2 + s_\rho^2 > 0$ is guaranteed by the assumption that the frequency responses F_α and F_β do not vanish. Further manipulation yields:

$$
\begin{aligned}
J_\rho &= \left[(s_\beta^2+c_\beta^2)\left((1-c_{\alpha^*})^2+s_{\alpha^*}^2\right)\right]^{-1}\times \\
&\quad \{(\mathbf{I},\mathbf{0})'\left(\mathbf{v}_{s,p}(c_\rho(1-c_{\alpha^*})+s_\rho s_{\alpha^*})+\mathbf{v}_{c,p}(c_\rho s_{\alpha^*}-s_\rho(1-c_{\alpha^*})\right)+ \\
&\quad (\mathbf{0},\mathbf{I})'\left(\mathbf{v}_{s,p}(c_\rho c_\beta+s_\rho s_\beta)+\mathbf{v}_{c,p}(s_\rho c_\beta-c_\rho s_\beta)\right)\} \\
&= \left[(s_\beta^2+c_\beta^2)\left(s_{\alpha^*}^2+(1-c_{\alpha^*})^2\right)\right]^{-1}\times \\
&\quad \{(\mathbf{I},\mathbf{0})'\left(\mathbf{v}_{s,p}c_\beta-\mathbf{v}_{c,p}s_\beta\right)\left(s_{\alpha^*}^2+(1-c_{\alpha^*})^2\right)+ \\
&\quad (\mathbf{0},\mathbf{I})'\left(\mathbf{v}_{s,p}(1-c_{\alpha^*})+\mathbf{v}_{c,p}s_{\alpha^*}\right)\left(s_\beta^2+c_\beta^2\right)\} \\
&= (\mathbf{I},\mathbf{0})'\frac{\mathbf{v}_{s,p}c_\beta-\mathbf{v}_{c,p}s_\beta}{s_\beta^2+c_\beta^2}+(\mathbf{0},\mathbf{I})'\frac{\mathbf{v}_{s,p}(1-c_{\alpha^*})+\mathbf{v}_{c,p}s_{\alpha^*}}{s_{\alpha^*}^2+(1-c_{\alpha^*})^2} \\
&= \left[\frac{\mathbf{v}'_{s,p}c_\beta-\mathbf{v}'_{c,p}s_\beta}{|F_\beta|^2},\frac{\mathbf{v}'_{s,p}(1-c_{\alpha^*})+\mathbf{v}'_{c,p}s_{\alpha^*}}{|F_\alpha|^2}\right]' \qquad \text{(A.5)}
\end{aligned}
$$

The (asymptotic) variance of the time delay $\phi^* = \phi/\omega$ is accordingly given by

$$
V\left(\hat{\phi}_\rho^*\right) = \omega^{-2} J_\rho' V_r J_\rho, \qquad \text{(A.6)}
$$

at all frequencies where J_ρ exists. Asymptotic normality follows from standard central limit theorems.

Publisher: Hans-Böckler-Stiftung, Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
Phone: +49-211-7778-331, IMK@boeckler.de, http://www.imk-boeckler.de

IMK Working Paper is an online publication series available at:
http://www.boeckler.de/imk_5016.htm

ISSN: 1861-2199

Hans Böckler Stiftung

Fakten für eine faire Arbeitswelt.