Testing for time-varying Granger causality

Christopher F Baum[1] Jesús Otero[2] Stan Hurn[3]

[1] Boston College, Chestnut Hill, MA, USA

[2] Universidad del Rosario, Bogotá, Colombia

[3] Queensland University of Technology, Brisbane, Australia

2021 Stata Economics Virtual Symposium

Introduction

Causal relationships in the econometric analysis of time series are typically based on the concept of predictability and are established by testing for Granger causality (Granger, 1969, 1988).

The popularity of Granger causality stems from the fact that it is identified using reduced-form VAR models, applicable to a set of potentially jointly determined variables.

Advantages:

- No need for normalization
- No need for guidance from economic theory

Introduction

Some important studies applying Granger causality in economics:

- Money and income (Friedman and Kuttner, 1993; Swanson, 1998; Shi, Hurn, and Phillips, 2020),
- GDP and energy consumption (Lee, 2006; Arora and Shi, 2016),
- CO_2 emissions and economic growth (Grossman and Krueger, 1995),
- economic growth and quality of health (Tapia Granados and Ionides, 2008),
- oil prices and output (Hamilton, 1983).

Introduction

Standard econometric software packages for the estimation/analysis of VAR models provide Granger causality tests.

However, just as with other aspects of structural stability, Granger causality may be supported over one time frame, but may be fragile when alternative periods are considered (see Thoma, 1994; Swanson, 1998; Psaradakis, Ravn, and Sola, 2005).

Introduction

Phillips, Wu, and Yu (2011) and Phillips, Shi, and Yu (2014, 2015a,b) derive theoretical results for testing for and date-stamping asset price bubbles.

Based on these results, Shi, Phillips, and Hurn (2018) and Shi, Hurn, and Phillips (2020) revisit the notion of time variation in testing for Granger causality. They establish that it is possible to assess the stability of causal relationships over time.

- Shi, Phillips, and Hurn (2018) study the stationary VAR model,
- Shi, Hurn, and Phillips (2020) study the lag-augmented VAR (LA-VAR) model, which allows for non-stationary variables.

Introduction

Although conceptually straightforward, these stability analysis algorithms offer significant challenges in terms of implementation in software.

- These methods produce large numbers of test statistics which must be efficiently stored and displayed for analysis.
- Analysis of the statistical significance of test results requires bootstrapping in order to ensure correct inference.

This presentation illustrates how the analysis can be accomplished using a new community-contributed Stata command, `tvgc`.

Introduction

Organization of the talk:

- briefly present the Granger causal framework,
- address recursive techniques for assessing time variation,
- present details of the `tvgc` command,
- provide an example focusing on key US macro series.

Granger causality

Consider without loss of generality the bivariate VAR(m) model given by

$$y_{1t} = \phi_0^{(1)} + \sum_{k=1}^{m} \phi_{1k}^{(1)} y_{1\,t-k} + \sum_{k=1}^{m} \phi_{2k}^{(1)} y_{2\,t-k} + \varepsilon_{1t} \tag{1}$$

$$y_{2t} = \phi_0^{(2)} + \sum_{k=1}^{m} \phi_{1k}^{(2)} y_{1\,t-k} + \sum_{k=1}^{m} \phi_{2k}^{(2)} y_{2\,t-k} + \varepsilon_{2t}, \tag{2}$$

where y_{1t} and y_{2t}, respectively, represent economic time series of interest. Variable y_1 is said to Granger cause variable y_2 if the past values of y_1 have predictive power for the current value of y_2, conditional on the past returns of y_2.

The null hypotheses of no Granger causality from y_1 to y_2 involves testing the joint significance of $\phi_{1k}^{(2)}$ $(k = 1, \cdots, m)$ by means of a Wald test.

Granger causality

Recasting the bivariate VAR(m) in matrix notation:

$$\mathrm{y}_t = \Pi \mathrm{x}_t + \varepsilon_t. \quad (3)$$

where

$$\mathrm{y}_t = [\; y_{1t} \quad y_{2t} \;]',$$
$$\mathrm{x}_t = [\; 1 \quad \mathrm{y}'_{t-1} \quad \mathrm{y}'_{t-2} \quad \cdots \quad \mathrm{y}'_{t-k} \;]'$$
$$\Pi_{2\times(2m+1)} = [\; \Phi_0 \quad \Phi_1 \quad \cdots \quad \Phi_m \;]$$

with

$$\Phi_0 = [\; \phi_0^{(1)} \quad \phi_0^{(2)} \;]'$$

and

$$\Phi_k = \begin{bmatrix} \phi_{1k}^{(1)} & \phi_{2k}^{(1)} \\ \phi_{1k}^{(2)} & \phi_{2k}^{(2)} \end{bmatrix} \text{ for } k = 1, \ldots, m.$$

The null of no Granger causality from variable y_1 to y_2 is $\mathrm{R}_{1\to 2}\pi = 0$, where $\mathrm{R}_{1\to 2}$ is the coefficient restriction matrix and $\pi = \mathrm{vec}(\Pi)$ using row vectorization.

Granger causality

The heteroskedastic-consistent Wald statistic of the null hypothesis is denoted by $\mathcal{W}_{1\to 2}$ and is defined as

$$\mathcal{W}_{1\to 2} = T\left(\mathrm{R}_{1\to 2}\hat{\pi}\right)'\left[\mathrm{R}_{1\to 2}\left(\hat{\mathrm{V}}^{-1}\hat{\Sigma}\hat{\mathrm{V}}^{-1}\right)\mathrm{R}'_{1\to 2}\right]^{-1}\left(\mathrm{R}_{1\to 2}\hat{\pi}\right),$$

where $\hat{\mathrm{V}} = \mathrm{I}_n \otimes \hat{\mathrm{Q}}$ and $\hat{\mathrm{Q}} = T^{-1}\sum_t \mathrm{x}_t\mathrm{x}'_t$ and $\hat{\Sigma} = T^{-1}\sum_t \hat{\xi}_t\hat{\xi}'_t$ with $\hat{\xi}_t = \hat{\varepsilon}_t \otimes \mathrm{x}_t$ and $\hat{\varepsilon}_t = \mathrm{y}_t - \hat{\Pi}\mathrm{x}_t$.

The formulation of a test for Granger causality in a VAR(G) system, $G = 2, \ldots$ is straightforward.

This framework applies to testing for Granger causality in the context of a VAR model estimated using stationary variables.

Granger causality

To account for integrated variables, Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996) recommend estimating a LA-VAR model, which is the original VAR(m) model augmented with d lags for the possible maximum order of integration of the variables.

The resulting model is denoted VAR($m + d$).

To test for Granger causality in the LA-VAR model, one proceeds just as before.The coefficients associated to the additional d are not included in the testing restrictions.

Recursive testing algorithms

To allow for time variation in Granger causal orderings and to date-stamp the timing of the changes, recursive estimation methods are required.

There are three algorithms that generate a sequence of test statistics:

- the forward expanding (FE) window,
- the rolling (RO) window,
- the recursive evolving (RE) window.

Recursive testing algorithms

Consider a sample of $T + 1$ observations $\{y_0, y_1, \cdots, y_T\}$, a number r such that $0 < r < 1$ and consider $[Tr]$ to denote the integer part of the product.

Then $\mathcal{T}_{r_1,r}$ will be taken to denote a Wald test statistic computed over a subsample starting at $y_{[Tr_1]}$ and ending at $y_{[Tr]}$.

A schematic representation of the algorithms is given in what follows. Each of the arrows is representative of a subsample over which the relevant test statistic is computed.

The forward expanding window

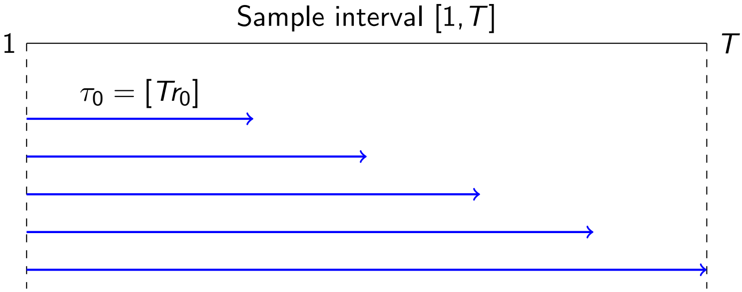
The FE algorithm (Thoma, 1994) is a standard forward recursion.

The Wald test statistic is computed first for a minimum window length, $\tau_0 = [Tr_0] > 0$, and the sample size then expands sequentially by one observation until the final test statistic is computed using the entire sample. The starting point of every subsample is the first data point.

This is what Stata's prefix command `rolling:` produces with the `recursive` option.

At the conclusion of the FE algorithm, a sequence of Wald test statistics, $\mathcal{T}_{r_1,r}$ with $r_1 = 0$ and $r \in [r_0, 1]$, is obtained.

The forward expanding window



(a) Forward expanding window

Figure 1: Forward expanding window (based on Phillips, Shi, and Yu, 2015a)

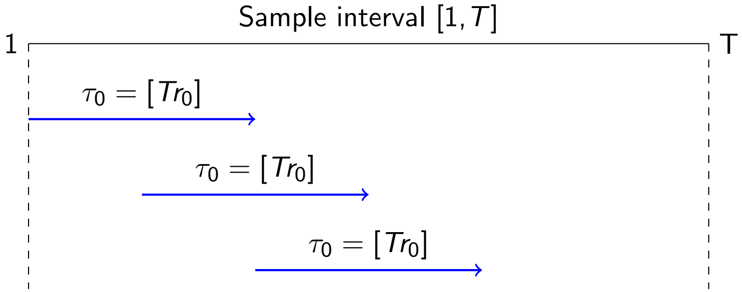
The rolling window

In the RO algorithm (Swanson, 1998; Arora and Shi, 2016) a window of size $[Tw]$ is rolled through the sample advancing one observation at a time and a Wald test statistic is computed for each window.

This is what Stata's prefix command `rolling:` produces by default.

The output from the RO algorithm is a sequence of test statistics $\mathcal{T}_{r_1,r}$ with $r_1 = r - w$ and $r \in [r_0, 1]$, where each test statistic is computed from a sample of the same size, $[Tw]$, with $0 < w < 1$.

The rolling window



(a) Rolling window

Figure 2: Rolling window (based on Phillips, Shi, and Yu, 2015a)

The recursive evolving window

In the RE algorithm, for a given observation of interest, the algorithm computes a test statistic for every possible subsample of size r_0 or larger with the observation of interest providing the common end point of all the subsamples.

This procedure is repeated taking the observation of interest to be every point in the sample, subject only to the minimum window size. Thus every observation in the sample byeond the first is associated with it a set of Wald test statistics. Phillips, Shi, and Yu (2015b) propose that inference be based on a sequence of supremum norms of these statistics.

The RE algorithm produces a sequence of test statistics $\mathcal{T}_{r_1,r}$ with $r_1 \in [0, r - r_0]$ and $r \in [r_0, 1]$ which are the sup norms of the Wald statistics at each observation.

The recursive evolving window

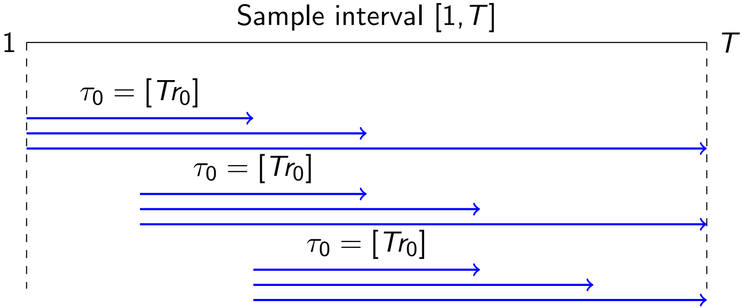
The RE algorithm encompasses both the FE and RO recursions as special cases.

For each observation in turn, a sequence of test statistics is defined that can be arranged in an upper triangular square matrix with column and row dimensions equal to the largest number of usable observations.

- the FE Wald statistic is the leading entry in each column,
- the RO Wald statistic is located on the main diagonal,
- the largest elements of each column are the RE statistics.

The information derived from these test statistics can be used over the full sample or analyzed through the period in order to focus on the timing of these time-varying phenomena via date stamping.

The recursive evolving window



(a) Recursive evolving window

Figure 3: Sample sequences and window widths (based on Phillips, Shi, and Yu, 2015a)

Full sample analysis

If the null hypothesis of interest is whether a particular variable does not Granger causes another variable at any time during the sample, with the alternative that there is Granger causality at some time, a single test statistic is required.

- The maximal FE statistic is the largest element of the first row of the upper triangular matrix of test statistics.
- The maximal RO statistic is the largest element of the main diagonal of the matrix.
- The maximal RE statistic is the largest element of the entire upper triangular matrix.

Date stamping

Beyond these summary measures for the full sample, the sequence of FE, RO and RE statistics can be graphed and compared with the bootstrap percentiles derived by methods described in Shi, Hurn, and Phillips (2020), section 3 and Shi, Phillips, and Hurn (2018), section 4.1.

These estimates can then be used to identify periods in which the potential Granger-casual relationships vary significantly.

The estimated origination date of a change is determined as the first instance at which the test statistic exceeds its critical value. Subsequent changes are then identified in a similar fashion.

The `tvgc` command

Syntax

The `tvgc` command tests whether the first variable in the *varlist* is Granger-caused by the remaining variables.

Before using the `tvgc` command, it is necessary to `tsset` the data. The *varlist* cannot contain gaps, but can contain time-series operators. `tvgc` does not support the `by:` prefix. It may be applied to one unit of a panel dataset.

`tvgc` *varlist* [*if*] [*in*] [, `prefix`*(string)* `p`(*integer*) `d`*(string)* `robust` `trend` `matrix` `window`(*integer*) `boot`(*integer*) `seed`(*integer*) `sizecontrol`(*integer*) `graph` `eps` `pdf` `notitle` `restab`]

Ben Jann's `moremata` package is required: `ssc install moremata` for the latest version.

The tvgc command supports the following options:

- prefix can be used to provide a 'stub' with which variables created in tvgc will be named. If this option is given, three Stata variables will be created for the appropriate range of dates: *prefix*_forward_*varname*, *prefix*_rolling_*varname*, and *prefix*_recursive_*varname*. These variables contain the sequences of these three test statistics over the sample period. The prefix option must be specified to enable the graph option.
- p sets the number of lags to be included in the VAR model, with a default of 2.
- d sets the number of lags to be included in the lag-augmented part of the VAR model, with a default of 1. This option must be used when there are integrated variables in the *varlist*. Set d=0 if no augmented lags are needed.

The tvgc command
Options

- `robust` specifies that heteroskedasticity-robust test statistics should be computed.
- `trend` specifies that a linear trend should be included in the VAR.
- `matrix` specifies that the T x T matrices of test statistics should be returned. They are named `r_`*rhsvar* for each of the test variables.
- `window` specifies the number of observations to be included in the rolling windows. If not specified, 20% of the sample is used.
- `boot` sets the number of replications to perform the bootstrap advocated by Phillips and Shi (2020). The default is 199.
- `seed` sets the initial seed for random number generation in bootstrapping.

- sizecontrol specifies the number of observations to be included in the bootstrap computations in order to control the empirical size, with a default of 12.

- graph specifies that the timeseries of the three test statistics should be graphed along with their 90% and 95% critical values. The graphs will be saved with names specified by the prefix() option as *prefix*_forward, *prefix*_rolling and *prefix*_rolling. The eps and pdf options specify the format in which the graphs are saved. The notitle option suppresses the graph titles.

- restab specifies that a LaTeX table containing the test statistics and their 95th and 99th percentile values should be written to restab.tex. The file will be replaced if it exists. When including this fragment in a LaTeX document, use the booktabs package.

The tvgc command

Implementation

Although reference is made to the rolling-window algorithms provided by Stata's `rolling:` prefix, all computations are performed in Mata rather than the ado-file language in order to produce results with acceptable speed. Even without bootstrapping, computation of the upper triangular matrix of test statistics requires $(T)(T-1)/2$ estimates of the VAR or LA-VAR model for each right-hand variable where T is the number of usable observations.

Empirical application

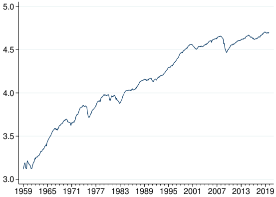
The tvgc command is illustrated using a three-variable VAR specification for monthly US data.

- logarithm of industrial production ($\ln i$),
- unemployment rate (u),
- logarithm of the price of crude oil ($\ln o$).

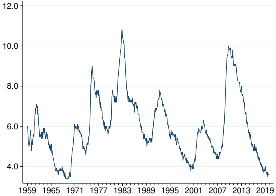
These variables are a subset of those used by Hamilton (1983) to study the relationship between oil and the macroeconomy using Granger causality tests.

The sample period is January 1959 to December 2019, with 732 observations accessed from FRED. The task of downloading these series is simplified with the Stata command freduse; see Drukker (2006). The Stata dataset used in this illustration can be downloaded using the command bcuse us_outoil; install bcuse from SSC if needed.

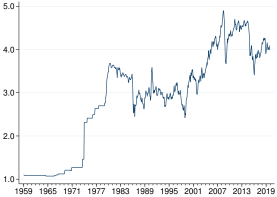
Empirical application



(a) $\ln i$



(b) u



(c) $\ln o$

Figure 4: Variables in levels

Empirical application

It is apparent that both $\ln i$ and $\ln o$ are trending.

The order of integration of the variables is assessed with the ADF_{max} and DFGLS unit root tests of Leybourne (1995) and Elliott, Rothenberg, and Stock (1996).

Using the Stata commands `adfmaxur` (Otero and Baum, 2018) and `ersur` (Otero and Baum, 2017), the results suggest:

- presence of a unit root in $\ln i$ and $\ln o$,
- stationarity of the unemployment rate

As there are I(1) variables in the VAR model, our analysis proceeds in the context of a LA-VAR model where $d = 1$.

Empirical application

The optimal lag order of the VAR model is chosen using the Stata command `varsoc`.

The command is applied to a model that includes a linear trend, which enters as an exogenous variable. The maximum number of lags is set to 12 as the data are monthly.

The Schwartz lag-order selection statistic recommends $p = 2$ lags, while Akaike favors $p = 6$ lags. The more parsimonious choice of $p = 2$ is adopted here.

Empirical application

Letting $x \xrightarrow{GC?} y$ to denote that the direction of Granger causality being tested runs from x to y, the following relationships are tested:

- $u \xrightarrow{GC?} \ln i$ and $\ln o \xrightarrow{GC?} \ln i$;
- $\ln i \xrightarrow{GC?} u$ and $\ln o \xrightarrow{GC?} u$;
- $\ln i \xrightarrow{GC?} \ln o$ and $u \xrightarrow{GC?} \ln o$.

The required commands are respectively:

Empirical application

```
. tvgc li u lo, trend win(72) sizecontrol(12) p(2) d(1) robust

. tvgc u li lo, trend win(72) sizecontrol(12) p(2) d(1) robust

. tvgc lo li u, trend win(72) sizecontrol(12) p(2) d(1) robust
```

In all three cases the chosen options indicate

- presence of a linear trend as an exogenous variable, `trend`,
- initial estimation window of 72 observations, `win(72)`,
- size of the tests controlled over a one-year period, `sizecontrol(12)`
- two lags in the VAR, `p(2)`,
- one lag in the lag-augmented part of the VAR, `d(1)`,
- tests robust to heteroskedasticity, `robust`.

To produce the output shown here, the options `boot(499)`, `seed(123)`, `prefix`, `graph` and `eps` were also used.

Empirical application

The results for the full sample show that we fail to reject the null hypothesis of no Granger causality from income and unemployment to the price of oil when applying the FE window.

In all other cases there is evidence of Granger causality when the test statistics are compared with the 95th percentiles of the empirical distribution of the bootstrap test statistics.

Empirical application

Table 1: Wald tests of Granger causality

Direction of causality	Max Wald FE	Max Wald RO	Max Wald RE
$u \xrightarrow{GC?} \ln i$	20.524 (10.283) [15.751]	31.073 (10.355) [15.110]	38.806 (10.775) [16.131]
$\ln o \xrightarrow{GC?} \ln i$	12.037 (8.709) [12.459]	28.322 (8.970) [13.526]	31.689 (9.324) [14.389]
$\ln i \xrightarrow{GC?} u$	70.205 (10.360) [15.850]	68.762 (10.469) [17.025]	75.290 (10.544) [17.892]
$\ln o \xrightarrow{GC?} u$	46.355 (9.673) [13.607]	42.252 (9.807) [13.607]	64.877 (10.118) [13.962]
$\ln i \xrightarrow{GC?} \ln o$	4.349 (7.913) [14.964]	25.639 (8.565) [14.956]	30.328 (9.344) [14.964]
$u \xrightarrow{GC?} \ln o$	3.440 (9.333) [15.417]	17.229 (9.333) [14.775]	17.253 (10.121) [15.417]

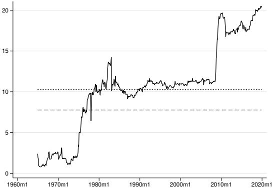
The 95th and 99th percentiles of the empirical distribution of the bootstrap statistics are in parentheses and brackets, respectively.

Empirical application

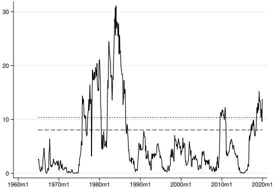
Figures 5 to 10 display the time-varying Granger causality test results. In general, these plots all support the conclusion that Granger causal relationships are extremely dynamic and that the patterns of causation found in the data depend on the type of recursive algorithm used.

These plots display the 90th and 95th percentiles of the empirical distribution of the bootstrap statistics, to be compared with the sequence of FE, RO, and RE test statistics.

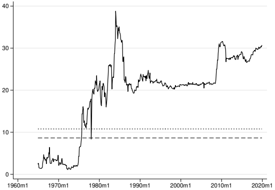
Empirical application



(a) Forward: $u \xrightarrow{GC?} \ln i$



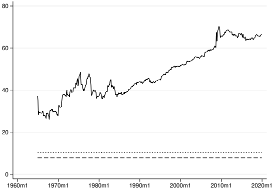
(b) Rolling: $u \xrightarrow{GC?} \ln i$



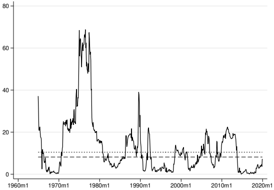
(c) Recursive: $u \xrightarrow{GC?} \ln i$

Figure 5: Time-varying Granger causality tests between $\ln i$ and u

Empirical application



(a) Forward: $\ln i \xrightarrow{GC?} u$

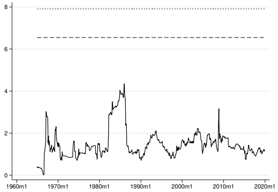


(b) Rolling: $\ln i \xrightarrow{GC?} u$

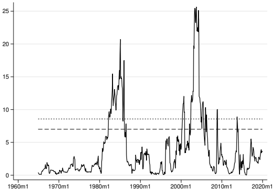
(c) Recursive: $\ln i \xrightarrow{GC?} u$

Figure 6: Time-varying Granger causality tests between $\ln i$ and u

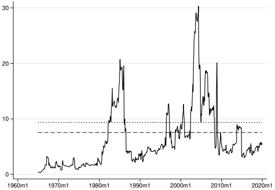
Empirical application



(a) Forward: $\ln i \xrightarrow{GC?} \ln o$



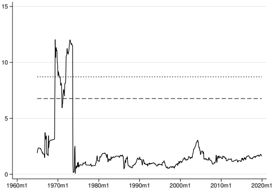
(b) Rolling: $\ln i \xrightarrow{GC?} \ln o$



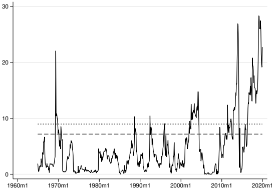
(c) Recursive: $\ln i \xrightarrow{GC?} \ln o$

Figure 7: Time-varying Granger causality tests between $\ln i$ and $\ln o$

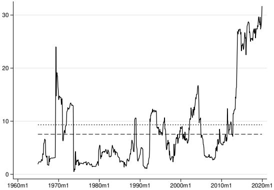
Empirical application



(a) Forward: $\ln o \xrightarrow{GC?} \ln i$



(b) Rolling: $\ln o \xrightarrow{GC?} \ln i$

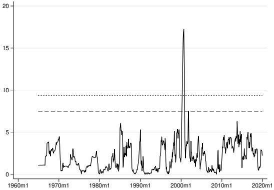


(c) Recursive: $\ln o \xrightarrow{GC?} \ln i$

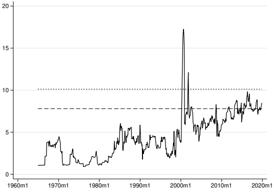
Figure 8: Time-varying Granger causality tests between $\ln i$ and $\ln o$

Empirical application

(a) Forward: $u \xrightarrow{GC?} \ln o$



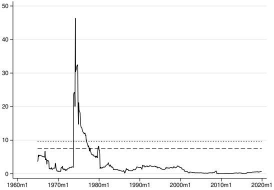
(b) Rolling: $u \xrightarrow{GC?} \ln o$



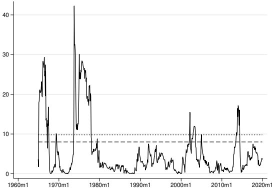
(c) Recursive: $u \xrightarrow{GC?} \ln o$

Figure 9: Time-varying Granger causality tests between u and $\ln o$

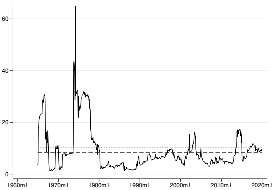
Empirical application



(a) Forward: $\ln o \xrightarrow{GC?} u$



(b) Rolling: $\ln o \xrightarrow{GC?} u$



(c) Recursive: $\ln o \xrightarrow{GC?} u$

Figure 10: Time-varying Granger causality tests between u and $\ln o$

Summary of findings

Estimation using the FE and RE windows indicate that during most of the study period there is evidence of Granger causality from unemployment to income; see Figures 5(a) and 5(c), and vice versa, see Figures 6(a) and 6(c). These results provide strong support for the intuition that these two measures of economic activity are closely related.

FE estimation also shows that the price of oil Granger-causes income in the late 1960s and early 1970s; see Figure 8(a).

Summary of findings

In contrast, strong evidence of Granger causality from income to the price of oil is apparent in the 1980s and 2000s with the RO and RE windows: see Figures 7(b) and 7(c), respectively.

The fact that the FE window fails to pick up the opening of this causal channel confirms a well-known problem with the FE algorithm: namely, that it is not sensitive to changes late in the sample period.

Summary of findings

A particularly strong illustration of the effects of the first oil shock of October 1973 is to be found in Figures 10(a), 10(b) and 10(c) showing Granger causality from the oil price to unemployment.All the algorithms identify a period of strong causality which starts at the time of the first oil shock and lasts until the second oil shock in 1979.

Interestingly, although the causal channel from the oil price to unemployment is active at times in the latter half of the sample period, the channel is not evident during the great recession of 2008–2009.

Concluding remarks

Evaluation of Granger-causal relationships among macroeconomic aggregates is an important component of macroeconometric modeling, as it is key to formally assessing the temporal stability of these relationships.

Concluding remarks

Evaluation of Granger-causal relationships among macroeconomic aggregates is an important component of macroeconometric modeling, as it is key to formally assessing the temporal stability of these relationships.

We present the `tvgc` command to test for time-varying Granger causality. The command produces full-sample test statistics as well as date-stamping of the periods in which there are significant findings of Granger-causal relationships.

Concluding remarks

Evaluation of Granger-causal relationships among macroeconomic aggregates is an important component of macroeconometric modeling, as it is key to formally assessing the temporal stability of these relationships.

We present the `tvgc` command to test for time-varying Granger causality. The command produces full-sample test statistics as well as date-stamping of the periods in which there are significant findings of Granger-causal relationships.

Using US monthly data on industrial production, unemployment and oil prices, we find support for the conclusion that causal relationships can change dramatically over any given sample period.

Concluding remarks

Evaluation of Granger-causal relationships among macroeconomic aggregates is an important component of macroeconometric modeling, as it is key to formally assessing the temporal stability of these relationships.

We present the `tvgc` command to test for time-varying Granger causality. The command produces full-sample test statistics as well as date-stamping of the periods in which there are significant findings of Granger-causal relationships.

Using US monthly data on industrial production, unemployment and oil prices, we find support for the conclusion that causal relationships can change dramatically over any given sample period.

Arbitrarily choosing the sample period over which to conduct causality tests may lead to misleading inference compared with a strategy that allows data-driven identification of change points.

Concluding remarks

An extended version of this presentation, with additional details on inference, is available from Baum, Hurn, and Otero (2021).

References I

Arora, V. and S. Shi (2016). Energy consumption and economic growth in the United States. *Applied Economics 48*(39), 3763–3773.

Baum, C. F., S. Hurn, and J. Otero (2021). The dynamics of U.S. industrial production: A time-varying Granger causality perspective. *Econometrics and Statistics*, in press.

Dolado, J. J. and H. Lütkepohl (1996). Making Wald tests work for cointegrated VAR systems. *Econometric Reviews 15*(4), 369–386.

Drukker, D. M. (2006). Importing Federal Reserve economic data. *The Stata Journal 6*(3), 384–386.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*(4), 813–836.

Friedman, B. M. and K. N. Kuttner (1993). Another look at the evidence on money-income causality. *Journal of Econometrics 57*(1-3), 189–203.

References II

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

Granger, C. W. J. (1988). Some recent development in a concept of causality. *Journal of Econometrics 39*(1-2), 199–211.

Grossman, G. M. and A. B. Krueger (1995). Economic growth and the environment. *Quarterly Journal of Economics 110*(2), 353–377.

Hamilton, J. D. (1983). Oil and the macroeconomy since World War II. *Journal of Political Economy 91*(2), 228–248.

Lee, C.-C. (2006). The causality relationship between energy consumption and GDP in G-11 countries revisited. *Energy Policy 34*(9), 1086–1093.

Leybourne, S. (1995). Testing for unit roots using forward and reverse Dickey-Fuller regressions. *Oxford Bulletin of Economics and Statistics 57*(4), 559–571.

References III

Otero, J. and C. F. Baum (2017). Response surface models for the Elliott, Rothenberg, and Stock unit-root test. *The Stata Journal 17*(4), 985–1002.

Otero, J. and C. F. Baum (2018). Unit-root tests based on forward and reverse Dickey-Fuller regressions. *The Stata Journal 18*(1), 22–28.

Phillips, P. C. B. and S. Shi (2020). Real time monitoring of asset markets: bubbles and crises. In H. D. Vinod and C. R. Rao (Eds.), *Handbook of Statistics: Financial, Macro and Micro Econometrics Using R*, Volume 42, pp. 61–80. Amsterdam: Elsevier.

Phillips, P. C. B., S. Shi, and J. Yu (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics 76*(3), 315–333.

Phillips, P. C. B., S. Shi, and J. Yu (2015a). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review 56*(4), 1043–1077.

References IV

Phillips, P. C. B., S. Shi, and J. Yu (2015b). Testing for multiple bubbles: Limit theory ofreal-time detectors. *International Economic Review 56*(4), 1079–1134.

Phillips, P. C. B., Y. Wu, and J. Yu (2011). Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? *International Economic Review 52*(1), 201–226.

Psaradakis, Z., M. O. Ravn, and M. Sola (2005). Markov switching causality and the money-output relationship. *Journal of Applied Econometrics 20*(5), 665–683.

Shi, S., S. Hurn, and P. C. B. Phillips (2020). Causal change detection in possibly integrated systems: Revisiting the money-income relationship. *Journal of Financial Econometrics 18*(1), 158–180.

Shi, S., P. C. B. Phillips, and S. Hurn (2018). Change detection and the causal impact of the yield curve. *Journal of Time Series Analysis 39*(6), 966–987.

References V

Swanson, N. R. (1998). Money and output viewed through a rolling window. *Journal of Monetary Economics 41*(3), 455–474.

Tapia Granados, J. A. and E. L. Ionides (2008). The reversal of the relation between economic growth and health progress: Sweden in the 19th and 20th centuries. *Journal of Health Economics 27*(3), 544–563.

Thoma, M. A. (1994). Subsample instability and asymmetries in money-income causality. *Journal of Econometrics 64*(1-2), 279–306.

Toda, H. Y. and T. Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics 66*(1-2), 225–250.