The Stata Journal (2019)
19, Number 4, pp. 883–899 DOI: 10.1177/1536867X19893631

Vector autoregressive-based Granger causality test in the presence of instabilities

Barbara Rossi
University of Pompeu Fabra,
Barcelona Graduate School of Economics, and CREI
Barcelona, Spain
barbara.rossi@upf.edu

Yiru Wang
University Pompeu Fabra
Barcelona, Spain
yiru.wang@upf.edu

Abstract. In this article, we review Granger causality tests that are robust to the presence of instabilities in a vector autoregressive framework. We also introduce the `gcrobustvar` command, which illustrates the procedure in Stata. In the presence of instabilities, the Granger causality robust test is more powerful than the traditional Granger causality test.

Keywords: st0581, gcrobustvar, Granger causality, vector autoregressive, VAR, instability, structural breaks, local projections

1 Introduction

Vector autoregressive (VAR) models have played an important role in macroeconomic analysis since Sims (1980). A VAR is a multiequation, multivariable linear model where each variable is in turn explained by its own lagged values as well as current and past values of the remaining variables. Compared with a univariate autoregression, VARs provide both a systematic way to capture the rich dynamics in multiple time series and a coherent, credible approach to forecasting.

Granger (1969) causality is a useful tool for characterizing the dependence among time series in reduced-form VARs, and Granger causality test statistics are widely used to examine whether lagged values of one variable help to predict another variable—see Stock and Watson (2001).

However, VAR analyses in macroeconomic data face important practical challenges: economic time-series data are prone to instabilities (see Stock and Watson [1996, 1999, 2003, 2006]; Rossi [2013]; Clark and McCracken [2006]), and VAR estimates may also be prone to instabilities (see Boivin and Giannoni [2006], Kozicki and Tinsley [2001], and Cogley and Sargent [2001, 2005]).

Thus, given the widespread use of VARs and the evidence of instabilities, it is potentially important to allow for changes over time when doing VAR-based statistical inference. As demonstrated in Rossi (2005), because the traditional Granger causality test assumes stationarity, it is not reliable in the presence of instabilities and may lead to incorrect inference.

In this article, we present the **gcrobustvar** command, which illustrates how to test Granger causality in a way that is robust to the presence of instabilities. The test is based on methodologies developed by Rossi (2005) and includes the robust versions of the mean and exponential Wald tests (Andrews and Ploberger 1994), the Nyblom (1989) test, and the Quandt (1960) and Andrews (1993) quasilikelihood-ratio tests. In the presence of instabilities, the Granger causality robust tests are more powerful than the traditional Granger causality test. The tests can also be used to find the point in time in which Granger causality either appears or breaks down in the data. In addition, the test is valid for reduced-form VAR models and VAR-based direct multistep (VAR–LP) forecasting models. The former assume homoskedastic idiosyncratic shocks, while the latter are estimated via local projections (LP) (see Jordà [2005]) and hence assume heteroskedastic and serially correlated idiosyncratic shocks.

We first introduce the tests and then present the commands that implement them. Then, we illustrate the empirical implementation of the robust Granger causality tests using a three-variable (inflation, unemployment, and interest rate) VAR model with four lags as in Stock and Watson (2001), as well as a direct multistep VAR–LP forecasting model. Finally, we compare the results with those based on a traditional Granger causality test.

The remainder of this article is organized as follows. Section 2 describes the theoretical framework and the Granger causality robust tests. Section 3 introduces the **gcrobustvar** command, which implements the Granger causality robust tests in Stata. Section 4 applies the Granger causality robust tests in the three-variable VAR and compares the results with the traditional Granger causality test. Section 5 applies the Granger causality robust tests in the direct multistep VAR–LP forecasting model.

2 VAR-based Granger causality test in the presence of instabilities

2.1 Motivation

In the presence of instabilities, as shown in Rossi (2005), traditional Granger causality tests may have no power. Consider one of the equations in a two-variable VAR with one lag and fixed prediction horizon h, for example: $y_{t+h} = \beta_t x_{t-1} + \rho y_{t-1} + \varepsilon_{t+h}$, $t = 2, 3, \ldots, T$. For simplicity, we assume that $x_{t-1}, \varepsilon_{t+h} \overset{\text{i.i.d.}}{\sim} N(0,1)$ and x_{t-1}, y_{t-1}, ε_{t+h} are independent of each other. Suppose the parameter β_t changes through time as follows:

$$\beta_t = 2/3\,(t \leq T/3) - 1/3(t > T/3) \tag{1}$$

In this example, a traditional Granger causality test would be a t test applied on the full-sample ordinary least-squares (OLS) parameter estimator $\widehat{\beta}^{\text{OLS}}$, which, asymptotically, can be calculated as (because the regressors are independent)

$$\begin{aligned}\widehat{\beta}^{\text{OLS}} &\approx \left(\sum_{t=2}^{T} x_{t-1}^2\right)^{-1} \sum_{t=2}^{T} x_{t-1} y_{t+h} \\ &= \left(T^{-1}\sum_{t=2}^{T} x_{t-1}^2\right)^{-1} T^{-1}\left\{\sum_{t=2}^{T/3} x_{t-1}^2\,(2/3) + \sum_{t=T/3+1}^{T} x_{t-1}^2\,(-1/3)\right\} \\ &\quad + \left(T^{-1}\sum_{t=2}^{T} x_{t-1}^2\right)^{-1} T^{-1}\sum_{t=2}^{T} x_{t-1} y_{t-1}\rho \\ &\quad + \left(T^{-1}\sum_{t=2}^{T} x_{t-1}^2\right)^{-1} T^{-1}\sum_{t=2}^{T} x_{t-1}\varepsilon_{t+h} \quad \underset{p}{\rightarrow} 0\end{aligned} \tag{2}$$

because $T^{-1}\sum_{t=2}^{T} x_{t-1}^2 \underset{p}{\rightarrow} E\left(x_t^2\right) = 1$, $T^{-1}\sum_{t=2}^{T} x_{t-1}y_{t-1} \underset{p}{\rightarrow} 0$ and $T^{-1}\sum_{t=2}^{T} x_{t-1}\varepsilon_{t+h} \underset{p}{\rightarrow} 0$.

Equation (2) implies that we do not reject the null hypothesis even if x_{t-1} does Granger-cause y_{t+h} in reality. This failure to reject results from the violation of the stationarity assumption underlying traditional Granger causality tests because the predictive ability is unstable across time. Thus, traditional Granger causality tests can be inconsistent if there are instabilities in the parameters. Without losing generality, this conclusion can be generalized to instabilities other than (1) by varying the time and the magnitude of the break. Note that this conclusion is empirically relevant because evidence shows that parameter estimates change substantially in sign and magnitude across time; see, for example, Welch and Goyal (2008) and Rossi (2005).

Considering the possibility of parameter instabilities, Rossi (2005) proposes tests to evaluate the predictive ability in the situation where the parameter might be time varying by testing jointly the significance of the predictors and their stability over time. More generally, let β_t change at some unknown point in time, τ: $\beta_t = \beta_1 \times 1\,(t \leq \tau) + \beta_2 \times 1(t > \tau)$.[1] Let $\widehat{\beta}_{1\tau}$ and $\widehat{\beta}_{2\tau}$ denote the OLS estimators before and after the break.[2]

1. Rossi (2005) considered various forms of instabilities, a more general case of testing possibly nonlinear restrictions in models fit with generalized method of moments, and tests on subsets of parameters.
2. That is, asymptotically, because the regressors are independent:

$$\begin{aligned}\widehat{\beta}_{1\tau} &\approx \left(\frac{1}{\tau}\sum_{t=1}^{\tau} x_{t-1}^2\right)^{-1}\left(\frac{1}{\tau}\sum_{t=1}^{\tau} x_{t-1}y_{t+h}\right) \\ \widehat{\beta}_{2\tau} &\approx \left(\frac{1}{T-\tau}\sum_{t=\tau+1}^{T} x_{t-1}^2\right)^{-1}\left(\frac{1}{T-\tau}\sum_{t=\tau+1}^{T} x_{t-1}y_{t+h}\right)\end{aligned}$$

With respect to the null hypothesis of no Granger causality at any point in time, that is, $H_0\colon \beta_t = \beta = 0$, the robust test builds on two components: $(\tau/T)\widehat{\beta}_{1\tau} + \{1 - (\tau/T)\}\widehat{\beta}_{2\tau}$ and $\widehat{\beta}_{1\tau} - \widehat{\beta}_{2\tau}$. A test on whether the first component (the full-sample estimate of the parameter)[3] is zero detects situations in which the parameter β_t is constant and different from zero. A test on whether the second component (the difference between the parameters estimated in the two subsamples) is zero detects situations in which the parameter changes, which detects situations in which the regressor Granger-causes the dependent variable in such a way that the parameter changes but the average estimate equals zero, as in (1). Rossi (2005) proposes several test statistics, including QLR^*_T, MeanW^*_T, and ExpW^*_T.[4] The corresponding critical values of the asymptotic distributions under the null are tabulated in Rossi's (2005) table B1.

Note that a test for structural breaks would not necessarily be the correct approach either. In fact, while in the previous example the researcher would identify a break, a structural break test is not sufficient or necessary for the existence of Granger causality. In fact, imagine that a variable has predictive content for another variable and the predictive ability is constant over time; that is, $\beta_t = \beta$. A structural break test is not necessary or sufficient to detect predictive ability. The approach taken in this article is to jointly test $\beta_t = \beta = 0$, which also avoids issues of multiple testing that one would incur when separately testing instability and Granger causality.

Note also that the way the possible presence of instabilities is modeled here is via a one-time break; such an approach has been proven to be more powerful than cumulative sum (CUSUM) tests—see Andrews, Lee, and Ploberger (1996), who derived the optimal tests (the exponential averages of the Wald test statistics) for one or more change points at unknown times in a multiple linear regression model. They compare the power of the optimal exponential tests with that of other tests in the literature such as the likelihood-ratio or supF test; the CUSUM test in Brown, Durbin, and Evans (1975); and the midpoint F test considering a one-time break in parameter. They find that the optimal tests perform quite well in finite samples compared with the other tests considered while the CUSUM test performs poorly.

3. The first component is the full-sample estimate of the parameter, which, asymptotically, is equivalent to the following:

$$\frac{\tau}{T}\widehat{\beta}_{1\tau} + \left(1 - \frac{\tau}{T}\right)\widehat{\beta}_{2\tau} = \left(\frac{1}{T}\sum_{t=1}^{T} x_{t-1}^2\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} x_{t-1}y_{t+h}\right)$$

4. Please refer to Rossi (2005) for detailed expressions of these statistics.

2.2 Framework

We consider two types of VAR specifications. The first is a reduced-form VAR with time-varying parameters,

$$\begin{aligned} A_t(L)\mathbf{y}_t &= \mathbf{u}_t \\ A_t(L) &= \mathbf{I} - \mathbf{A}_{1,t}L - \mathbf{A}_{2,t}L^2 - \cdots - \mathbf{A}_{p,t}L^p \\ \mathbf{u}_t &\overset{\text{i.i.d}}{\sim} (\mathbf{O}, \boldsymbol{\Sigma}) \end{aligned} \tag{3}$$

where $\mathbf{y}_t = [y_{1,t}, y_{2,t}, \ldots, y_{n,t}]'$ is an $(n \times 1)$ vector and $\mathbf{A}_{j,t}, j = 1, \ldots, p$, are $(n \times n)$ time-varying coefficient matrices.

The second is a direct multistep VAR–LP forecasting model with time-varying parameters. By iterating (3), $\mathbf{y}_{t+h}$ can be projected onto the linear space generated by $(\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p})'$, specifically,

$$\mathbf{y}_{t+h} = \boldsymbol{\Phi}_{1,t}\mathbf{y}_{t-1} + \boldsymbol{\Phi}_{2,t}\mathbf{y}_{t-2} + \cdots + \boldsymbol{\Phi}_{p,t}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_{t+h} \tag{4}$$

where $\boldsymbol{\Phi}_{j,t}, j = 1, \ldots, p$ are functions of $\mathbf{A}_{j,t}, j = 1, \ldots, p$ in (3) and $\boldsymbol{\epsilon}_{t+h}$ is a moving average of the errors $\mathbf{u}$ from time t to $t + h$ in (3) and therefore uncorrelated with the regressors but serially correlated itself.[5] Note that $h = 0$ is a special case where (4) degenerates to (3). Thus, we focus on (4) from now on.

Let $\boldsymbol{\theta}_t$ be an appropriate subset of $\text{vec}(\boldsymbol{\Phi}_{1,t}, \boldsymbol{\Phi}_{2,t}, \ldots, \boldsymbol{\Phi}_{p,t})$. The null hypothesis of the Granger causality robust test is

$$H_0\colon \boldsymbol{\theta}_t = \mathbf{0} \qquad \forall t = 1, 2 \ldots T \tag{5}$$

The statistics to test H_0 in (5), following Rossi (2005), are ExpW* (the exponential Wald test), MeanW* (the mean Wald test), Nyblom* (the Nyblom test), and QLR* (the Quandt likelihood-ratio [QLR] test).[6]

The optimal ExpW* and the optimal MeanW* tests are based on the exponential test statistics proposed in Andrews and Ploberger (1994). The optimal MeanW* is designed for alternatives that are close to the null hypothesis, while the optimal ExpW* is designed for testing against more distant alternatives. The optimal Nyblom* is based on the Nyblom (1989) test, which is the locally most powerful invariant test for the constancy of the parameter process against the alternative that the parameters follow a random walk process. The optimal QLR* is based on Andrews's (1993) Sup-LR test (or the QLR test), which considers the supremum of the statistics over all possible break dates of the Chow statistic designed for a fixed break date.

5. See Jordà (2005) for more details on LP.
6. See Rossi (2005) for more details on constructing the statistics.

2.3 A special case: The traditional Granger causality test

The traditional Granger causality test is a special case where the parameters in (4) are time invariant; that is, for $j = 1, \ldots, p$, we replace $\mathbf{\Phi}_{j,t}$ with $\mathbf{\Phi}_j, t = 1, \ldots, T$. Thus, (4) becomes

$$\mathbf{y}_{t+h} = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_{t+h} \tag{6}$$

To consider a more concrete example, Stock and Watson (2001) study a three-variable VAR with four lags ($p = 4$) and $h = 0$. The variables included are inflation (π_t), unemployment (u_t), and interest rate (R_t). Their reduced-form VAR is

$$\begin{bmatrix} \pi_t \\ u_t \\ R_t \end{bmatrix} = \mathbf{\Phi}_1 \begin{bmatrix} \pi_{t-1} \\ u_{t-1} \\ R_{t-1} \end{bmatrix} + \mathbf{\Phi}_2 \begin{bmatrix} \pi_{t-2} \\ u_{t-2} \\ R_{t-2} \end{bmatrix} + \mathbf{\Phi}_3 \begin{bmatrix} \pi_{t-3} \\ u_{t-3} \\ R_{t-3} \end{bmatrix} + \mathbf{\Phi}_4 \begin{bmatrix} \pi_{t-4} \\ u_{t-4} \\ R_{t-4} \end{bmatrix} + \begin{bmatrix} \epsilon_t^\pi \\ \epsilon_t^u \\ \epsilon_t^R \end{bmatrix}$$

$$\mathbf{\Phi}_j = \begin{bmatrix} \phi_j^{\pi,\pi} & \phi_j^{\pi,u} & \phi_j^{\pi,R} \\ \phi_j^{u,\pi} & \phi_j^{u,u} & \phi_j^{u,R} \\ \phi_j^{R,\pi} & \phi_j^{R,u} & \phi_j^{R,R} \end{bmatrix}, \qquad j = 1, \ldots, 4$$

Thus, in Stock and Watson (2001), the reduced-form VAR involves three equations: current unemployment as a function of past values of unemployment, inflation, and the interest rate; current inflation as a function of past values of inflation, unemployment, and the interest rate; and current interest rate as a function of past values of inflation, unemployment, and the interest rate. Stock and Watson (2001) consider traditional Granger causality tests in each equation where the null hypothesis is $H_0^*: \boldsymbol{\theta} = \mathbf{0}$, where $\boldsymbol{\theta}$ is the appropriate subset of $\text{vec}(\Phi_1, \Phi_2, \ldots, \Phi_p)$. For example, unemployment does not Granger-cause inflation if

$$\phi_1^{\pi,u} = \phi_2^{\pi,u} = \phi_3^{\pi,u} = \phi_4^{\pi,u} = 0$$

If unemployment does not Granger-cause inflation, then lagged values of unemployment are not useful for predicting inflation.

3 The gcrobustvar command

3.1 Syntax

The `gcrobustvar` command implements the VAR-based Granger causality robust test. The general syntax of the `gcrobustvar` command is

`gcrobustvar` *depvarlist*`, pos(`#`, `#`)` [`nocons horizon(`#`) lags(`*numlist*`)` `trimming(`*level*`)`]

depvarlist is a list of dependent variables, that is, all the variables in $\mathbf{y}_t$ in the notation in (6).

3.2 Options

`pos(#, #)` is a numeric list (that is, *numlist* in Stata) including two integers indicating the positions of the targeted dependent variable and restricted regressor, respectively. For example, if we are testing whether the second variable, $y_{2,t}$, Granger-causes the first variable, $y_{1,t}$, in the presence of instabilities, then we assign the numeric list as `pos(1,2)`, where the integer 1 refers to the position of the targeted dependent variable in the VAR (that is, $y_{1,t}$ in this example) and the integer 2 refers to the position of the targeted restricted regressor in the VAR (that is, $y_{2,t}$ in this example). `pos()` is required.

`nocons` suppresses the constant term. The default regression includes the constant term.

`horizon(#)` specifies the targeted horizon, that is, h in the notation in (4). The default refers to a reduced-form VAR assuming homoskedastic idiosyncratic shocks. When `horizon(`*h*`)` ($h \geq 0$) is specified, the command assumes heteroskedastic and serially correlated idiosyncratic shocks and chooses the truncation lag used in the estimation of the long run variance. The truncation lag is automatically determined using Newey and West (1994) optimal lag-selection algorithm. Note that `horizon(0)` refers to a reduced-form VAR assuming heteroskedastic and serially correlated idiosyncratic shocks, and `horizon(`*h*`)` ($h > 0$) refers to the $(h + 1)$-step-ahead forecasting model; see (4). For example, in a one-year-ahead VAR–LP forecasting model with quarterly data, `horizon(3)` should be specified.

`lags(`*numlist*`)` specifies the lags included in the VAR. The default is `lags(1 2)`. This option takes a *numlist* and not simply an integer for the maximum lag. For example, `lags(2)` would include only the second lag in the model, whereas `lags(1 2)` would include both the first and second lags in the model. The shorthand to indicate the range follows *numlist* in Stata.

`trimming(`*level*`)` is the trimming parameter. As is standard in the structural break literature, the possible break dates are usually trimmed to exclude the beginning and end of the sample period. If we specify `trimming(`μ`)`, the range where we search for instabilities is set to be $[\mu T, (1 - \mu)T]$, where T is the number of total periods. The default is `trimming(0.15)`, which is recommended in the structural break literature and commonly used in practice.

3.3 Stored results

gcrobustvar stores the following in r():

Macros

r(cmd)	gcrobustvar
r(cmdline)	command as type

Matrices

r(result_stat)	4-by-1 matrix containing four statistics: ExpW*, MeanW*, Nyblom*, and SupLR*
r(result_pv)	4-by-1 matrix containing four p-values, corresponding, respectively, to ExpW*, MeanW*, Nyblom*, and SupLR*
r(result_wald)	a column vector containing Wald statistics across time, the supremum of which is the optimal QLR test statistic (QLR*)

3.4 Empirical example of practical implementation in Stata

In what follows, we illustrate how to use the gcrobustvar command to implement the robust Granger causality test in Stata. The data (GCdata.xlsx, provided with the article files) include quarterly U.S. data on the rate of price inflation (π_t), the unemployment rate (u_t), and the interest rate (R_t, specifically, the federal funds rate) from 1959:I–2000:IV. These are the same variables used in Stock and Watson (2001). Inflation is computed as $\pi_t = 400 \times \ln(P_t/P_{t-1})$, where P_t is the chain-weighted gross domestic product price index. Quarterly data on u_t and R_t are quarterly averages of their monthly values.

Consider the inflation equation in (3):

$$\pi_t = c_t^{\pi} + \Phi_t^{\pi,\pi}(L)\pi_t + \Phi_t^{\pi,u}(L)u_t + \Phi_t^{\pi,R}(L)R_t + \epsilon_t^{\pi}$$
$$\text{where} \quad \Phi_t^{\cdot,\cdot}(L) = \phi_{1,t}^{\cdot,\cdot}L + \phi_{2,t}^{\cdot,\cdot}L^2 + \phi_{3,t}^{\cdot,\cdot}L^3 + \phi_{4,t}^{\cdot,\cdot}L^4$$

Suppose we are interested in testing whether unemployment (u) Granger-causes inflation (π), and we want the test to be robust to instabilities over time. That is, we want to test whether the coefficients of lagged values of unemployment (u) are zero across time:

$$H_0\colon \phi_{j,t}^{\pi,u} = 0 \qquad \forall j = 1,2,3,4 \quad \forall t = 1,2 \ldots T$$

Implementing the Granger causality tests in the presence of instabilities

The following scripts implement the Granger causality robust test. We first import the data, claim the data to be time series, and import the p-value tables needed for the tests:

```
. * import data
. import excel gcdata.xlsx, sheet(SW2001) firstrow clear
(4 vars, 168 obs)
. * time-series settings
. generate year = int(pdate)
. generate quarter = (pdate - int(pdate))*4 + 1
```

```
. generate tq = yq(year, quarter)
. format tq %tq
. tsset tq
        time variable:  tq, 1959q1 to 2000q4
                delta:  1 quarter
. * import p-value table
. mata:
--------------------------------------------- mata (type end to exit) -----------
: mata clear
: mata matuse pvtable,replace
(loading pvap0opt[34,21], pvapiopt[34,21], pvnybopt[34,21], pvqlropt[34,21])
: st_matrix("r(pvap0opt)",pvap0opt)
: st_matrix("r(pvapiopt)",pvapiopt)
: st_matrix("r(pvnybopt)",pvnybopt)
: st_matrix("r(pvqlropt)",pvqlropt)
: end
----------------------------------------------------------------------------------
. matrix pvap0opt = r(pvap0opt)
. matrix pvapiopt = r(pvapiopt)
. matrix pvnybopt = r(pvnybopt)
. matrix pvqlropt = r(pvqlropt)
```

Then, we run the Granger causality robust test using the `gcrobustvar` command. When we run the `gcrobustvar` command, important information (variables, lags, etc.) will be displayed:

```
. * run gcrobustvar test for a VAR
. gcrobustvar pi u R, pos(1,2) lags(1/4)
Running the Granger Causality Robust Test...
Setting:
 Variables in VAR: pi u R
 Lags in VAR:1 2 3 4
 h is 0 (reduced-form VAR).
 Trimming parameter is .15
 Constant is included.
 Assuming homoskedasticity in idiosyncratic shocks.
```

The results are displayed in the following script. The `gcrobustvar` command provides the four optimal test statistics (ExpW*, MeanW*, Nyblom*, QLR*) and their corresponding p-values:

```
Results of Granger Causality Robust Test: Lags of u Granger cause pi
ExpW*,MeanW*,Nyblom*,QLR* -- and their p-values below
                       ExpW      MeanW     Nyblom      SupLR
statistics(pi:u)  9.1974593  17.047538   4.689049  23.187923
   p-value(pi:u)  .07383772  .05866136  .07939046  .07069996
```

Here is how we get all the inputs of the `gcrobustvar` command. *depvarlist* lists the variables included in the VAR, that is, π, u, and R in this order. Because we are testing whether lags of the second variable, u_t, Granger-cause the first variable, π_t, in the presence of instabilities, we assign the positions `pos(1,2)`. As for the options, we include the constant term and include four lags, that is, `lags(1/4)`, as in Stock and Watson (2001). In addition, we assume homoskedasticity and choose the standard trimming parameter 0.15.

Here is how to interpret the results. Let's take the ExpW* statistics as an example. The value of ExpW* is 9.20, and the p-value is 0.07. Thus, the test rejects the null hypothesis that unemployment (u) does not Granger-cause inflation (π) for all t at the 10% significance level.

4 Comparison with the traditional Granger causality test

In this section, we compare the robust Granger causality tests with the traditional Granger causality test in the three-variable VAR model in Stock and Watson (2001). The VAR includes a constant term and four lags and assumes homoskedastic idiosyncratic shocks.

Table 1 reports the p-values of the traditional Granger causality Wald statistics. The results show that π Granger-causes R, u Granger-causes both π and R, and R Granger-causes u at the 5% significance level.

Table 1. Traditional reduced-form VAR-based Granger causality tests

	Dependent variable		
Restricted regressors	π	u	R
π	0.00	0.25	0.00
u	0.01	0.00	0.00
R	0.22	0.00	0.00

NOTE: This table reports p-values of the Wald statistics of the traditional Granger causality test. $h = 0$ (that is, the reduced-form VAR model), lags $= (1, 2, 3, 4)$, assuming homoskedastic idiosyncratic shocks.

Table 2 reports the p-values of the robust Granger causality test statistics (for ExpW*, MeanW*, Nyblom*, and QLR*, respectively). We are testing whether the restricted regressor Granger-causes the dependent variable in the presence of instabilities. For example, if we consider the dependent variable π and the restricted regressor R, we are testing whether R Granger-causes π in a way robust to instabilities across time, that is, whether the coefficients of lags of R are constant and equal to zero over time. The p-value of the ExpW* statistics in panel A in table 2 is 0.01, so the test does reject the null at the 5% significance level. Hence, R does Granger-cause π.

Table 2. Robust Granger causality tests in the reduced-form VAR

Panel A—ExpW*

	Dependent variable		
Restricted regressors	π	u	R
π	0.00	0.20	0.00
u	0.07	0.00	0.00
R	0.01	0.00	0.00

Panel B—MeanW*

	Dependent variable		
Restricted regressors	π	u	R
π	0.00	0.44	0.00
u	0.06	0.00	0.00
R	0.20	0.01	0.00

Panel C—Nyblom*

	Dependent variable		
Restricted regressors	π	u	R
π	0.00	0.22	0.00
u	0.08	0.00	0.00
R	0.03	0.02	0.00

Panel D—QLR*

	Dependent variable		
Restricted regressors	π	u	R
π	0.00	0.08	0.00
u	0.07	0.00	0.00
R	0.00	0.00	0.00

NOTE: This table reports p-values of the statistics of the Granger causality robust test. $h = 0$ (that is, the reduced-form VAR model), lags $= (1, 2, 3, 4)$, the trimming parameter $\mu = 0.15$, assuming homoskedastic idiosyncratic shocks.

Comparing tables 1 and 2, we find the empirical conclusions differ if a researcher uses the Granger causality robust test instead of the traditional Granger causality test. In fact, R does not Granger-cause π at the 5% significance level in the traditional Granger causality test, but R does Granger-cause π at the 5% significance level in the Granger causality robust test according to the ExpW*, Nyblom*, and SupLR* test statistics. Hence, there is empirical evidence that lagged values of R can predict π, but the predictive ability shows up only sporadically over time, which is the reason why the traditional Granger causality test does not detect it.

`gcrobustvar` also returns a graph showing the whole sequence of the Wald statistics across time, which gives more information on when the Granger causality occurs. In fact, the optimal QLR* is the supremum of the sequence of Wald statistics testing whether the parameters are zero at each point in time against the alternative that the parameters change at a given break date at time `tq`. Figure 1 plots the sequence of the Wald statistics over the possible break dates, reported on the x axis. Consider as an example the test of whether unemployment (u) Granger-causes inflation (π); figure 1 documents the whole sequence of the Wald statistics testing whether unemployment (u) Granger-causes inflation (π). The sequence of the Wald statistic (depicted by a continuous line in figure 1) is above the 10% critical value line (depicted by the dashed lines) around 1970q1 and 1980q1. The figure is saved as `gcrobustvar_pi_u`.

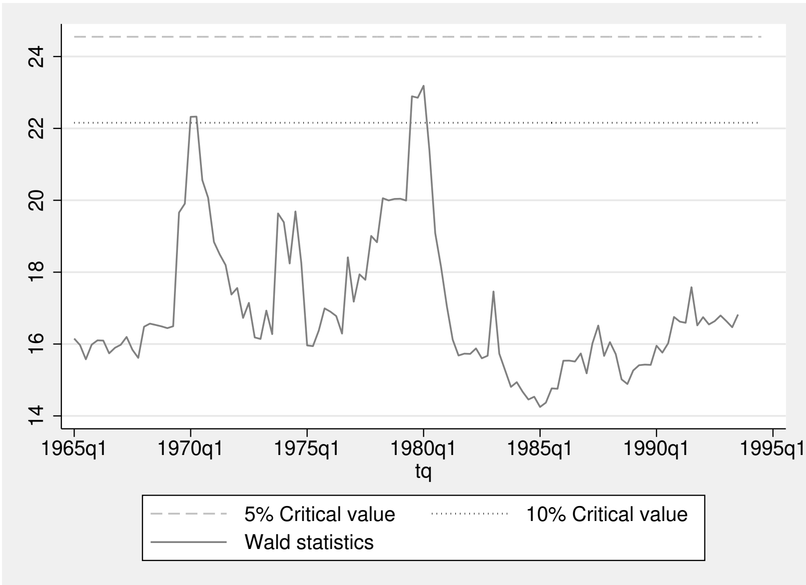

Figure 1. Wald statistics testing whether unemployment (u) Granger-causes inflation (π) against the alternative of a break in Granger causality at time `tq` (reported on the x axis)

5 Robust Granger causality tests in LP

Section 4 considers the reduced-form VAR assuming homoskedastic idiosyncratic shocks. In this section, we extend the VAR analysis to Jordà's (2005) LP by implementing the direct multistep VAR–LP forecasting model in (6) and assuming heteroskedastic and serially correlated idiosyncratic errors. Allowing for heteroskedasticity and serial correlation is important when the researcher extends the VAR analysis to LP, where the error terms in (6) can be both heteroskedastic and serially correlated.

We consider the one-year-ahead VAR–LP forecasting model with a constant term and four lags. The setting is similar to section 4 except that we specify $h = 3$ and relax the homoskedasticity assumption.

The following is the command to implement the robust Granger causality test to investigate whether the coefficients on $R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}$ are zero across time in the one-year-ahead VAR–LP forecasting model in the equation where the dependent variable is π_{t+3}. To test other coefficients, we implement the command similarly, except for adjusting the input of `pos(#, #)`.

```
. gcrobustvar pi u R, pos(1,3) lags(1/4) horizon(3)
Running the Granger Causality Robust Test...
Setting:
 Variables in VAR: pi u R
 Lags in VAR:1 2 3 4
 h is 3 (4-step-ahead VAR-LP forecasting model).
 Trimming parameter is .15
 Constant is included.
 Assuming heteroskedasticity and serial correlation in idiosyncratic shocks.
  (output omitted)
```

Table 3 reports the p-values of the robust Granger causality test statistics (the ExpW*, MeanW*, Nyblom* and QLR* statistics, respectively). The results show that lags of inflation (π) can significantly forecast the one-year-ahead unemployment (u) and interest rate (R), lags of unemployment can significantly forecast the one-year-ahead inflation and interest rate, and lags of interest rate can significantly forecast the one-year-ahead inflation and unemployment.

Table 3. Robust Granger causality tests in the direct multistep VAR–LP forecasting model

Panel A—ExpW*

Restricted regressors	Dependent variable π	u	R
π	0.00	0.00	0.00
u	0.00	0.00	0.00
R	0.00	0.00	0.00

Panel B—MeanW*

Restricted regressors	Dependent variable π	u	R
π	0.00	0.00	0.00
u	0.00	0.00	0.00
R	0.00	0.00	0.00

Panel C—Nyblom*

Restricted regressors	Dependent variable π	u	R
π	0.00	0.00	0.00
u	0.00	0.00	0.00
R	0.00	0.00	0.00

Panel D—QLR*

Restricted regressors	Dependent variable π	u	R
π	0.00	0.00	0.00
u	0.00	0.00	0.00
R	0.00	0.00	0.00

NOTE: This table reports p-values of the statistics of the Granger causality robust test. $h = 3$ (that is, the one-year-ahead VAR–LP forecasting model), lags $= (1, 2, 3, 4)$, the trimming parameter $\mu = 0.15$, assuming heteroskedastic and serially correlated idiosyncratic shocks.

6 Acknowledgment

The project received funding from the European Research Council (ERC). The Barcelona Graduate School of Economics acknowledges financial support from the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563).

7 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 19-4
. net install st0581      (to install program files, if available)
. net get st0581          (to install ancillary files, if available)
```

8 References

Andrews, D. W. K. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61: 821–856.

Andrews, D. W. K., I. Lee, and W. Ploberger. 1996. Optimal changepoint tests for normal linear regression. *Journal of Econometrics* 70: 9–38.

Andrews, D. W. K., and W. Ploberger. 1994. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62: 1383–1414.

Boivin, J., and M. P. Giannoni. 2006. Has monetary policy become more effective? *Review of Economics and Statistics* 88: 445–462.

Brown, R. L., J. Durbin, and J. M. Evans. 1975. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society, Series B* 37: 149–192.

Clark, T. E., and M. W. McCracken. 2006. The predictive content of the output gap for inflation: Resolving in-sample and out-of-sample evidence. *Journal of Money, Credit and Banking* 38: 1127–1148.

Cogley, T., and T. J. Sargent. 2001. Evolving post-World War II U.S. inflation dynamics. *NBER Macroeconomics Annual* 16: 331–373.

———. 2005. Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics* 8: 262–302.

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

Jordà, Ò. 2005. Estimation and inference of impulse responses by local projections. *American Economic Review* 95: 161–182.

Kozicki, S., and P. A. Tinsley. 2001. Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics* 47: 613–652.

Newey, W. K., and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61: 631–653.

Nyblom, J. 1989. Testing for the constancy of parameters over time. *Journal of the American Statistical Association* 84: 223–230.

Quandt, R. E. 1960. Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55: 324–330.

Rossi, B. 2005. Optimal tests for nested model selection with underlying parameter instability. *Econometric Theory* 21: 962–990.

———. 2013. Advances in forecasting under instability. In *Handbook of Economic Forecasting, vol. 2B*, ed. G. Elliott and A. Timmermann, 1203–1324. Amsterdam: Elsevier.

Sims, C. A. 1980. Macroeconomics and reality. *Econometrica* 48: 1–48.

Stock, J. H., and M. W. Watson. 1996. Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics* 14: 11–30.

———. 1999. Forecasting inflation. *Journal of Monetary Economics* 44: 293–335.

———. 2001. Vector autoregressions. *Journal of Economic Perspectives* 15: 101–115.

———. 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 41: 788–829.

———. 2006. Forecasting with many predictors. In *Handbook of Economic Forecasting, vol. 1*, ed. G. Elliott, C. W. J. Granger, and A. Timmermann, 515–554. Amsterdam: Elsevier.

Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21: 1455–1508.

About the authors

Barbara Rossi is an ICREA Professor of Economics at Pompeu Fabra University. She previously has been an associate professor with tenure at the Department of Economics at Duke University after earning her PhD from Princeton University. She is a CEPR Fellow, a member of the CEPR Business Cycle Dating Committee, and a Director of the International Association of Applied Econometrics.

Professor Rossi specializes in the fields of time-series econometrics and applied international finance and macroeconomics. Her current research focuses on forecasting and macroeconometrics. Professor Rossi has published her research findings in the *Review of Economic Studies*, the *Quarterly Journal of Economics*, the *Journal of Business and Economic Statistics*, the *International Economic Review*, *Econometric Theory*, the *Journal of Applied Econometrics*, the

Journal of Money, Credit and Banking, *Journal of Econometrics*, the *Review of Economics and Statistics*, the *Journal of International Money*, and *Finance and Macroeconomic Dynamics*. She also wrote the chapter "Advances in forecasting under model instabilities" for the *Handbook of Economic Forecasting* (Elsevier–North Holland eds.), the chapter "Forecasting in macroeconomics" for the *Handbook of Research Methods and Applications in Empirical Macroeconomics*, and an article for the *Journal of Economic Literature*.

Along with her teaching and research responsibilities, Professor Rossi holds various other professional positions. She serves as the editor of the *Journal of Applied Econometrics*, a coeditor of the *International Journal of Central Banking*, and an associate editor of *Quantitative Economics*, and she has served as associate editor for the *Journal of Business and Economic Statistics* and the *Journal of Economic Dynamics and Control*. She was the Program Chair for the 2016 Econometric Society European Summer Meetings and the 2014 International Association of Applied Econometrics Conference. Professor Rossi's research has been supported by the National Science Foundation through Grants NSF #0647627 and #1022125, the Marie Curie Fellowship program, the Spanish Ministry of Research, and the ERC through Grant 615608.

Professor Rossi has presented her findings at a variety of professional conferences and meetings, including the Econometric Society Meetings, the SED meetings, the Joint Statistical Meetings, the NBER-NSF Time Series Conference, the NBER, the IAAE, and the AEA meetings. She has held a visiting researcher position at the University of California–Berkeley, the University of Montreal in Canada, UC San Diego, the Federal Reserve Banks of Atlanta and Philadelphia, Norges Bank, Bank of France, and ENSAE–CREST in France.

Yiru Wang is a PhD candidate in economics at Pompeu Fabra University. She specializes in the fields of time-series econometrics and applied macroeconomics. Her current research focuses on forecasting and macroeconometrics. Wang's research has been supported by the Spanish Ministerio de Economa y Competitividad under FPI grant BES-2016-076868.

Wang has presented her findings at a variety of professional conferences and meetings, including the Econometric Society Meetings, the fourth Vienna Workshop on High Dimensional Time Series in Macroeconomics and Finance, the BigNOMICS Workshop on Big Data and Economic Forecasting in Ispra, the IXt Workshop in Time Series Econometrics in Zaragoza, and the Barcelona GSE Summer Forum.