Testing Instantaneous Causality in Presence of Nonconstant Unconditional Covariance

Quentin Giai Gianetto and Hamdi Raïssi
IRMAR-INSA, F-35708 Rennes, France (quentin2g@yahoo.fr)

This article investigates the problem of testing instantaneous causality between vector autoregressive (VAR) variables with time-varying unconditional covariance. It is underlined that the standard test does not control the Type I errors, while the tests with White and heteroscedastic autocorrelation consistent (HAC) corrections can suffer from a severe loss of power when the covariance is not constant. Consequently, we propose a modified test based on a bootstrap procedure. We illustrate the relevance of the modified test through a simulation study. The tests considered in this article are also compared by investigating the instantaneous causality relations between U.S. macroeconomic variables.

KEY WORDS: Unconditionally heteroscedastic errors; VAR model; Wild bootstrap.

1. INTRODUCTION

The concept of Granger causality (Granger 1969) has evolved as a standard tool to analyze cause-and-effect relationships between variables. Seminal papers relying on this concept are Sims (1972), Ashenfelter and Card (1982), Hamilton (1983), Lee (1992), and Hiemstra and Jones (1994) among others. In particular, it is said that there exists an instantaneous causality relation between $X_1_t$ and $X_2_t$ if the prediction of $X_2_t$ given the past values of $X_1_t = (X_1'_t, X_2'_t)'$ is improved by adding the available current information of the variables in $X_1_t$ (see Lütkepohl 2005, p. 42). When $X_t$ is assumed to follow a VAR process, a standard Wald test for zero restrictions on the errors covariance matrix is usually used for testing the instantaneous causality. Such a test, available in most specialized software, is based on the assumption of iid Gaussian errors (see Lütkepohl and Krätzig 2004). In nonstandard stationary cases, White (White 1980) or HAC corrections are considered (see Den Haan and Levin 1997 for the HAC estimation). Such situations arise when the error process is non-Gaussian or dependent. For instance nonlinear models are often adjusted to residuals (see, e.g., Bauwens, Laurent, and Rombouts 2006; Andrews, Davis, and Jay Breidt 2006; Amendola and Francq 2009 for nonlinear processes).

The Wald test and its corrections assume a constant unconditional covariance structure that has been challenged through many statistical studies applied to real data. For instance, Sensier and Van Dijk (2004) found that most of the 214 U.S. macroeconomic variables they investigated exhibit a break in their unconditional variance. Ramey and Vine (2006) highlighted a declining variance of the U.S. automobile production and sales. McConnell and Perez-Quiros (2000) documented a break in the variance driven by a Markov chain. Engle and Rangel (2008) and Hafner and Linton (2010) proposed detect instantaneous changes in the variance driven by a Markov chain. Engle and Rangel (2008) and Hafner and Linton (2010) proposed deterministic changes of the unconditional variance combined with a GARCH structure. Kokoszka and Leipus (2000) and Dahlhaus and Rao (2006) studied ARCH processes with time-varying parameters. Hansen (1995) assumed that the volatility is driven by a near integrated process. In the above contributions, the nonconstant unconditional variance is subject to stochastic effects. Reference can be also made to Azrak and Melard (2006) who investigated the estimation of ARMA models with possible nonconstant deterministic unconditional variance. Zhao and Li (2013) considered a deterministic general form of nonconstant variance which allows for nonlocally stationary processes. In this article we focus on the test of zero restrictions on time-varying covariance structure of the VAR innovations. Various forms of nonconstant unconditional variance specifications have been considered in the literature. For instance Francq and Gaudiot (2004) studied linear models allowing for unconditional changes in the variance driven by a Markov chain. Engle and Rangel (2008) and Hafner and Linton (2010) proposed deterministic changes of the unconditional variance combined with a GARCH structure. Kokoszka and Leipus (2000) and Dahlhaus and Rao (2006) studied ARCH processes with time-varying parameters. Hansen (1995) assumed that the volatility is driven by a near integrated process. In the above contributions, the nonconstant unconditional variance is subject to stochastic effects. Reference can be also made to Azrak and Melard (2006) who investigated the estimation of ARMA models with possible nonconstant deterministic unconditional variance. Zhao and Li (2013) considered a deterministic general form of nonconstant variance which allows for nonlocally stationary processes. In this article the rescaling device of Dahlhaus (1997) is used for the deterministic covariance structure (see Assumption A1). This kind of time-varying covariance specification is widely used in the literature and encompasses the piecewise constant case (see Hamori and Tokihisa 1997; Pesaran and Timmermann 2004; or Xu and Phillips 2008 and references therein). In this...
framework, it is highlighted that the standard Wald test for instantaneous causality does not provide suitable critical values. To remove such undesirable effects of heteroscedasticity, it is common to consider a White correction or apply a wild bootstrap procedure on the standard test statistic. However, it emerges that the tests with White or HAC corrections can suffer from a severe loss of power. It also appears that applying a wild bootstrap procedure directly to the standard statistic does not solve the problem. Noticing that the previous approaches fail to handle data with nonconstant covariance, we propose a new test for investigating instantaneous causality based on cumulative sums. We find that the asymptotic distribution of the corresponding statistic is nonstandard involving the unknown nonconstant covariance structure in a functional form. Therefore, we develop a wild bootstrap procedure to propose accurate critical values for our task (see, e.g., Wu 1986; Gonçalvez and Kilian 2004; Horowitz et al. 2006; Inoue and Kilian 2002 for the wild bootstrap method). We establish through theoretical and empirical results that the modified test is preferable to the tests based on the spurious assumption of stationary time series.

The plan of the article is as follows. In Section 2 the VAR model with nonconstant covariance is introduced. The properties of the tests based on the assumption of constant unconditional covariance are given. It emerges from this part that this kind of tests should be avoided in our nonstandard framework. As a consequence a test based on the wild bootstrap procedure taking into consideration nonconstant covariance is proposed in Section 3. The finite sample properties of the tests are investigated in Section 4 by Monte Carlo experiments. Outcomes of U.S. macroeconomic datasets illustrate our findings. In Section 5, we draw up a conclusion on our results.

The following general notations will be used. The convergence in distribution will be denoted by \( \Rightarrow \), while \( \rightarrow \) denotes the convergence in probability. The usual Kronecker product between matrices is denoted by \( \otimes \). The usual operator which stacks the columns of a given matrix into a vector is signified by \( \text{vec}(.)/\| . \| \) is the Euclidean norm. A \( d \times d \) identity matrix is denoted by \( I_d \).

2. VAR MODEL WITH NONCONSTANT COVARIANCE

Let \( X_t = (X_{1t}', X_{2t}')' \in \mathbb{R}^d \), with \( X_{it} \) of dimension \( d_i \), \( i \in \{ 1, 2 \} \), satisfy

\[
X_t = A_0X_{t-1} + \ldots + A_0X_{t-p} + \varepsilon_t, \quad u_t = H_t\varepsilon_t,
\]

(2.1)

where the matrices \( A_0 \) are such that det \( A(z) \neq 0 \) for all \( |z| \leq 1 \), and \( A(z) = I_d - \sum_{i=1}^{p} A_0 z^i \). It is assumed that the sequence \( X_{-p+1}, \ldots, X_0, X_1, \ldots, X_T \) is observed. The true autoregressive order \( p \) is supposed to be known to state the asymptotic results and to assess their finite sample reliability, eliminating the effects of a misspecification of \( p \) (Sections 2, 3, and 4.1). Of course, in practice, the lag length is unknown. For instance, we select using a goodness-of-fit test taking into account the time-varying causality in the real data study (Section 4.2). The case where the lag length is adjusted before testing for instantaneous causality is investigated in the online supplementary material through Monte Carlo experiments. The nonconstant covariance induced by \( H_t \) is specified in A1:

**Assumption A1.** (i) The \( d \times d \) matrices \( H_t \) are lower triangular nonsingular matrices with positive diagonal elements and satisfy \( H_t = G(t)/T \), where the components of the matrix \( G(r) := (g_{kl}(r)) \) are measurable deterministic functions on the interval \( (0, 1] \), such that \( \sup_{r \in (0, 1]} |g_{kl}(r)| < \infty \), and each \( g_{kl} \) satisfies a Lipschitz condition piecewise on a finite number of subintervals partitioning \( (0, 1] \). The matrices \( \Sigma(r) = G(r)G(r)' \) are assumed positive definite for all \( r \in (0, 1] \).

(ii) \( (\varepsilon_t) \) is \( \alpha \)-mixing, such that \( E(\varepsilon_t | F_{t-1}) = 0, E(\varepsilon_t\varepsilon_t' | F_{t-1}) = I_d \), where \( F_t \) is the \( \sigma \)-field generated by \( \{ \varepsilon_k : k \leq t \} \), and \( \sup_t \| \varepsilon_t \|< \infty \) for some \( \mu > 1 \).

Since the rescaling approach of Dahlhaus (1997) is used, the process \( (X_t) \) should be formally written in a triangular form. However, the double subscript is suppressed for notational simplicity.

If the errors covariance is constant \( \Sigma_t = \Sigma_0 \) for all \( t \), the usual hypotheses are tested for instantaneous causality:

\[
H_0^u : \Sigma^{12} = 0 \quad \text{vs} \quad H_1^u : \Sigma^{12} \neq 0,
\]

where \( \Sigma^{12} \) is the upper right block of \( \Sigma_0 \) (see Lütkepohl 2005, pp. 46–47). Let \( u_t = (u_{1t}', u_{2t}')' \) and define the (OLS) residuals \( \hat{u}_t = (\hat{u}_{1t}', \hat{u}_{2t}')' \). Define also \( \Sigma^{12}(r) \) the upper right block of \( \Sigma(r) \). Actually we have no instantaneous causality between \( X_{1t} \) and \( X_{2t} \) if and only if \( \Sigma^{12}(r) = 0 \). The block \( \Sigma^{12} \) is usually estimated by \( T^{-1}\sum_{t=1}^{T} \hat{u}_{1t}\hat{u}_{2t}' \) which converges in probability to \( J_0^1 \Sigma^{12}(r)dr \) under A1. Hence, such hypothesis testing can only be interpreted as a global zero restriction testing of the covariance structure, that is, testing \( J_0^1 \Sigma^{12}(r)dr = 0 \) versus \( J_0^1 \Sigma^{12}(r)dr \neq 0 \), and are inappropriate in our case.

In this part the effects of nonconstant covariance on the instantaneous causality tests based on the assumption of a stationary process are analyzed. Let \( \delta := T^{-1}\sum_{t=1}^{T} \text{vec}(\hat{u}_{1t}\hat{u}_{2t}') = T^{-1}\sum_{t=1}^{T} \hat{u}_{1t}\hat{u}_{2t}' \otimes \hat{u}_{1t} \). The standard test statistic is given by \( S_t := \delta_1'\hat{\Omega}_t^{-1}\delta_1 \), with \( \hat{\Omega}_t := (T^{-1}\sum_{t=1}^{T} \hat{u}_{2t}\hat{u}_{2t}') \otimes (T^{-1}\sum_{t=1}^{T} \hat{u}_{1t}\hat{u}_{1t}') \). If the practitioner assumes that the error process is iid but with \( u_{1t} \) and \( u_{2t} \) dependent, the statistic with White correction should be used: \( S_w = \delta_1'\hat{\Omega}_w^{-1}\delta_1 \), where \( \hat{\Omega}_w := T^{-1}\sum_{t=1}^{T} \hat{u}_{2t}\hat{u}_{2t}' \otimes \hat{u}_{1t}\hat{u}_{1t}' \).

In our framework of deterministic time-varying covariance, the same conclusions for White and HAC corrections hold and we only focus on the White correction (see Den Haan and Levin 1997 for the HAC correction). Usually the \( S_t \) and \( S_w \) statistics are compared to the critical values of the \( \chi^2_{d_1d_2} \) distribution. The resulting tests are denoted by \( S_t \) and \( S_w \).

The behaviors of the above tests are first given if \( \Sigma^{12}(r) = 0 \) holds (no instantaneous causality). Under A1 it can be shown that \( S_t = J_0^1 \Sigma^{22}(r)dr \otimes J_0^1 \Sigma^{11}(r)dr \) and \( \Omega = J_0^1 (G_2(r) \otimes G_1(r))M \) (\( G_2(r) \otimes G_1(r) \)) \( dr \), \( G(r) := (G_1(r), G_2(r))' \) and \( M = E(\varepsilon_t\varepsilon_t' \otimes \varepsilon_t\varepsilon_t') \). It follows that the standard test \( S_t \) is not able to control the Type I error since \( \Omega_t \neq \Omega \) in general. On the other hand, we also have \( S_w \Rightarrow \chi^2_{d_1d_2} \), so that the White-type test \( S_w \)
should have good size properties for large enough \( T \). However, it is well known that White correction can lead to important finite sample size distortions.

Now, we study the ability of the standard \( W_{st} \) and White \( W_w \) tests to detect instantaneous causal relationships \( \Sigma^{12}(r) \neq 0 \). In this eventually, we have \( \hat{\Omega}_w = O_p(1) \) and \( \hat{\Omega}_{st} = O_p(1) \) using standard arguments. If \( \int_0^T \Sigma^{12}(r) dr \neq 0 \), it can be shown under A1 that the \( W_w \) test achieve a gain in power in the Bahadur sense (Bahadur 1960) when compared to the \( W_{st} \) test. If \( \Sigma^{12}(r) \neq 0 \) and \( \int_0^T \Sigma^{12}(r) dr = 0 \), the noncentrality term is \( T^{-\frac{1}{2}} \sum_{i=1}^T \hat{\delta}_t = o(T^{\frac{1}{2}}) \). Therefore, we have \( S_t = o_p(T) \), with \( i = st, w \), and it is clear that the classical tests can suffer from a severe loss of power. From the outputs of this section it also clearly emerges that applying directly a wild bootstrap procedure to the standard statistic does not solve the above problem. As a consequence a bootstrap test based on a cumulative sum statistic is proposed in the next part.

3. A BOOTSTRAP TEST TAKING INTO ACCOUNT NONCONSTANT COVARIANCE

Recall that \( \Sigma^{12}(r) \) is the upper right block of \( \Sigma(r) \) given in Assumption A1. The following pair of hypotheses for instantaneous causality has to be tested in our framework:

\[
H_0: \Sigma^{12}(r) = 0 \quad \text{versus} \quad H_1: \Sigma^{12}(r) \neq 0
\]

for \( r \in [a, b] \subseteq [0, 1] \) with fixed \( a < b \).

To this aim introduce \( \delta_s = T^{-\frac{1}{2}} \sum_{s=1}^T \hat{\delta}_t \) for \( s \in (0, 1] \) and consider the following statistic:

\[
S_b = \sup_{s \in [0,1]} ||\delta_s||_2^2.
\]

Under A1 and if \( H_0 \) holds true we write:

\[
S_b \Rightarrow \sup_{s \in [0,1]} ||K(s)||_2^2, \tag{3.1}
\]

where \( K(s) = \int_0^s (G_2(r) \otimes G_1(r)) d B_{3d}(r) \) is a Brownian motion of covariance matrix \( M = E(\epsilon_1 \epsilon_2^\prime) - \text{vec}(I_d) \text{vec}(I_d)' \). The proof of (3.1) is given in the supplementary material. Under \( H_1 \) we obtain

\[
T^{-\frac{1}{2}} \delta_s = T^{-\frac{1}{2}} \sum_{t=1}^{[Ts]} \hat{\delta}_t \otimes \hat{\delta}_t + \text{vec}(\Sigma^{12}) + T^{-1} \sum_{t=1}^{[Ts]} \text{vec}(\Sigma^{12}). \tag{3.2}
\]

The first term in the right-hand side of (3.2) converges to zero in probability, while we have \( T^{-1} \sum_{t=1}^{[Ts]} \text{vec}(\Sigma^{12}) = \int_0^T \text{vec}(\Sigma^{12}(r)) dr + o(1) \) and \( \sup_{s \in [0,1]} ||\int_0^s \text{vec}(\Sigma^{12}(r)) dr||_2^2 = C > 0 \). Hence, we can deduce that in this situation \( S_b = CT + o_p(T) \).

From (3.1) we see that the asymptotic distribution of \( S_b \) under the null \( H_0 \) is nonstandard and depends on the unknown covariance structure and the fourth-order cumulants of the process \( \{\epsilon_t\} \) in a functional form. Thus the statistic \( S_b \) cannot directly be used to build a test and we consider a wild bootstrap procedure to provide reliable quantiles for testing the instantaneous causality.

In the literature such procedures have been used to investigate VAR model specification as in Inoue and Kilian (2002) among others. The reader is referred to Davidson and Flachaire (2008), Gonçalvez and Kilian (2004), Gonçalvez and Kilian (2007), and references therein for the wild bootstrap procedure method.

For resampling our test statistic we draw \( B \) bootstrap sets given by \( \xi_t^{(i)}(\hat{\omega}_2 \otimes \hat{\omega}_1), i \in \{1, \ldots, B\} \) and \( i \in \{1, \ldots, B\} \), where the univariate random variables \( \xi_t^{(i)} \) are taken iid standard Gaussian, independent from \( \{u_t\} \).

For a given \( i \in \{1, \ldots, B\} \) set \( \delta_s^{(i)}(T) = T^{-\frac{1}{2}} \sum_{t=1}^{[Ts]} \xi_t^{(i)}(\hat{\omega}_2 \otimes \hat{\omega}_1) \) and \( S_b^{(i)} = \sup_{s \in [0,1]} ||\delta_s^{(i)}||_2^2 \). In our procedure bootstrap counterparts of the \( \lambda_s \)’s are not generated and the residuals are directly used to generate the bootstrap residuals. This is motivated by the fact that zero restrictions are tested on the error’s covariance structure, so that we only consider the residuals in the test statistic. In addition it can be shown that the residuals and the errors are asymptotically equivalent for our purpose (see the proof of (3.1) in the online supplementary material).

The wild bootstrap method is designed to replicate the pattern of nonconstant covariance of the residuals in \( S_b^{(i)} \).

**Proposition 1.** Under A1 we have

\[
S_b^{(i)} \Rightarrow \sup_{s \in [0,1]} ||K(s)||_2^2, \tag{3.3}
\]

where \( \Rightarrow \) denotes the weak convergence in probability.

The proof of Proposition 1 is provided in the online supplementary material. The \( W_b \) test consists in rejecting \( H_0 \) if the statistic \( S_b \) exceeds the \((1 - \alpha)\) quantile of the bootstrap distribution. As pointed out by an anonymous referee the \( S_b \) statistic is unbalanced in the sense that the sup is more likely attained for large \( s \) under \( H_0 \). This can be solved by considering for instance the local sums

\[
\delta_s^{(i)} = \frac{1}{\sqrt{TF}} \sum_{t=TF(k-1)+1}^{TF(k-1)+1} \hat{\omega}_2 \otimes \hat{\omega}_1, \tag{3.4}
\]

where the bandwidth \( 0 < f_T < 1 \) decrease to zero at a suitable rate, the integers \( k \) are such that \( 1 \leq k \leq \lfloor T[TF] \rfloor =: k_{\text{max}} \) with \( k_{\text{max}} \to \infty \) as \( T \to \infty \), and \( \lfloor . \rfloor \) refers to the integer part. Define the statistic:

\[
\tilde{S}_b = \sup_{s \in [0,1]} \sum_{k=1}^{k_{\text{max}}} ||\delta_s^{(i)}||_2^2.
\]

Since it is assumed that we have a finite number of breaks and that \( \sup_{s \in [0,1]} |\hat{\omega}_1(r)\| < \infty \) in A1, it can be shown that \( \tilde{S}_b = O_p(1) \) using similar arguments to the proof of Remark 2.1 in Aue et al. (2009). To build a test for instantaneous causality one can use a bootstrap methodology as described above.

Under \( H_1 \) with \( \int_0^T \Sigma(r) dr \neq 0 \) we note that all the statistics considered in this article increase at the rate \( T \). However, when \( \Sigma(r) \neq 0 \) with \( \int_0^T \Sigma(r) dr \approx 0 \), we can expect that the \( W_b \) test is more powerful than the tests based on the assumption of constant unconditional covariance. In this situation the standard and White statistics are such that \( S_w = o_p(T) \), \( S_w = o_p(T) \) while again \( S_0 = O_p(T) \). If the unconditional covariance is constant and in the case of instantaneous causality, that is, \( \Sigma^{12}(r) = \Sigma^{12} \neq 0 \), note that \( S_0 = O_p(T) \), \( S_w = O_p(T) \) and \( S_b = O_p(T) \). Hence, we can expect no major loss of power for the \( W_b \) when compared to the standard \( W_{st} \) and White \( W_w \) tests if the underlying structure of the covariance is constant. In
general, since $S_b = O_p(T)$ and in view of Proposition 1, the $W_b$ test is consistent.

From the above results we can draw the conclusion that the $W_b$ test is preferable if the unconditional covariance is nonconstant for large enough sample sizes. In the next section the small sample properties of the tests are studied by means of Monte Carlo experiments.

4. NUMERICAL ILLUSTRATIONS

In this section, the bootstrap $W_b$ test is compared to the standard $W_{st}$ and White $W_w$ tests. First the Type I errors and power properties of the three tests are compared using simulated bivariate VAR(1) processes with unconditional time-varying covariance. To focus on the properties of the studied tests it is assumed that the true lag length $p = 1$ is known. In the online supplementary material simulation outputs where the lag length is selected by considering goodness-of-fit tests are given. In Section 4.2 the tests for instantaneous causality are applied to two macroeconomic datasets with a lag length selected using a portmanteau test adapted to our framework.

4.1 Simulation Study

For our experiments we simulated simple bivariate VAR(1) processes where the autoregressive parameters are inspired from those estimated from the money supply and inflation in the U.S. data (see Section 4.2). The data-generating process is given by

$$\begin{pmatrix} X_{1,t} \\
X_{2,t} \end{pmatrix} = \begin{pmatrix} 0.64 & -1 \\
-0.01 & 0.44 \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\
X_{2,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\
u_{2,t} \end{pmatrix}, \tag{4.1}$$

where the innovations are Gaussian with variance structure $\Sigma(r)$ fulfilling the assumption A1. Two cases are considered for this structure:

- Case 1: Empirical size setting. It does not exist an instantaneous causality relation between $X_{1,t}$ and $X_{2,t}$:

$$\Sigma(r) = \begin{pmatrix} \Sigma^{11}(r) & 0 \\
0 & \Sigma^{22}(r) \end{pmatrix} \quad \forall r \in (0, 1],$$

where $\Sigma^{11}(r) = a - \cos(br)$ and $\Sigma^{22}(r) = a + \sin(br)$ correspond to the nonconstant variances of the innovations. We take $a > 1$ which represents the level of these variances and $b$ their angular frequency.

- Case 2: Empirical power setting. It exists an instantaneous causality relation between $X_{1,t}$ and $X_{2,t}$:

$$\Sigma(r) = \begin{pmatrix} \Sigma^{11}(r) & \Sigma^{12}(r) \\
\Sigma^{12}(r) & \Sigma^{22}(r) \end{pmatrix} \quad \forall r \in (0, 1],$$

where $\Sigma^{12}(r) = c \sin(2\pi r)$ is such that $\int_0^1 \Sigma^{12}(r) dr = 0$ with $\Sigma^{12}(r) \neq 0$ almost everywhere on $r \in (0, 1]$, and $\Sigma^{11}(\cdot), \Sigma^{22}(\cdot)$ are defined as in Case 1. In particular, the constant $c$ will allow to investigate the ability of our modified test for detecting such alternative when it gets closer to the null hypothesis.

Note that $a$, $b$ and $c$ have to be chosen to fulfill the positive definite condition on $\Sigma(r)$ for all $r \in (0, 1]$. For instance, this property is checked if $a = 1.1$, $b = 11$, and $\frac{3}{4} \geq c > 0$.

The finite sample properties of the tests are assessed by means of the following Monte Carlo experiments. For each sample size, 1000 time series following (4.1) are generated. The lag length is assumed known and the autoregressive parameters are estimated by using the commonly used OLS method. In all our experiments we use 299 bootstrap iterations for the $W_b$ test. Processes generated by Case 1 are considered to shed light on the control of the Type I errors of the studied tests. The results are reported in Table 1. On the other hand, processes generated by Case 2 are considered for the power study. The results are given in Table 2 and Figure 1. Note that in Table 1 and Table 2 we take $a = 1.1$, $b = 11$, and $c = 0.5$, while in Figure 1 we take several values for $c$ and $a = 1.1$, $b = 11$.

In our example the studied tests seem to control the Type I errors reasonably well (see Table 1). We can remark that the standard test provides satisfactory results. Nevertheless, this outcome does not have to be generalized since we have seen in Section 2 that the standard test is in general inadequate in presence of time-varying variance. Now if we turn to the alternative given by Case 2, Table 2 clearly shows that the standard $W_{st}$ and White $W_w$ tests have no power as the sample sizes increase on the contrary of the bootstrap $W_b$ test. For instance the $W_b$ test is almost always rejecting the null hypothesis $H_0$ for sample sizes $T = 1000$, while the $W_{st}$ and $W_w$ tests are completely not able to detect the alternative in this case. This confirms the theoretical results obtained when $\int_0^1 \Sigma^{12}(r) dr \approx 0$.

In the above power experiments the changes of $\Sigma^{12}(r)$ around zero were fixed by a constant $c$. In this way we illustrate the ability of the tests to detect departures from the null hypothesis $\Sigma^{12}(r) = 0$, while we again have $\int_0^1 \Sigma^{12}(r) dr = 0$ but taking several values for $c$ (Figure 1). Here the sample is fixed $T = 500$.

| Table 1. The empirical size for the studied tests with asymptotic nominal level 1%, 5%, 10% and $a = 1.1$, $b = 11$, $c = 0.5$ |
|---|---|---|---|---|---|---|---|---|
| | Asymptotic nominal level 1% | | Asymptotic nominal level 5% | | Asymptotic nominal level 10% | |
| | $W_{st}$ | $W_w$ | $W_b$ | $W_{st}$ | $W_w$ | $W_b$ | $W_{st}$ | $W_w$ | $W_b$ |
| Sample size | | | | | | | | | |
| 50 | 0.010 | 0.007 | 0.007 | 0.048 | 0.050 | 0.047 | 0.102 | 0.113 | 0.105 |
| 100 | 0.009 | 0.009 | 0.010 | 0.057 | 0.066 | 0.066 | 0.095 | 0.096 | 0.107 |
| 200 | 0.011 | 0.010 | 0.011 | 0.045 | 0.047 | 0.052 | 0.101 | 0.116 | 0.122 |
| 500 | 0.011 | 0.010 | 0.010 | 0.042 | 0.047 | 0.050 | 0.099 | 0.101 | 0.101 |
| 1000 | 0.008 | 0.009 | 0.010 | 0.056 | 0.051 | 0.051 | 0.088 | 0.097 | 0.103 |
Table 2. The empirical power for the studied tests based on asymptotic nominal levels 1%, 5%, 10% and \( a = 1.1, b = 11, c = 0.5 \)

| Sample size | \( W_s \) | \( W_w \) | \( W_b \) | \( W_s \) | \( W_w \) | \( W_b \) | \( W_s \) | \( W_w \) | \( W_b \) |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 50          | 0.014     | 0.005     | 0.003     | 0.056     | 0.040     | 0.045     | 0.110     | 0.085     | 0.132     |
| 100         | 0.012     | 0.005     | 0.012     | 0.056     | 0.038     | 0.102     | 0.098     | 0.079     | 0.213     |
| 200         | 0.017     | 0.010     | 0.076     | 0.063     | 0.048     | 0.305     | 0.106     | 0.093     | 0.512     |
| 500         | 0.011     | 0.006     | 0.486     | 0.050     | 0.038     | 0.837     | 0.105     | 0.080     | 0.930     |
| 1000        | 0.015     | 0.011     | 0.966     | 0.056     | 0.045     | 0.997     | 0.108     | 0.088     | 1.000     |

We clearly observe that the relative rejection frequencies of the \( W_b \) test increase when the covariance structure \( \Sigma^{12}(r) \neq 0 \) goes away from zero but with \( \int_0^1 \Sigma^{12}(r)dr = 0 \). On the other hand we again remark that the relative rejection frequencies of the tests based on the assumption of constant variance remain close to the asymptotic nominal level even when \( c \) takes large values.

4.2 Application to Macroeconomic Datasets

In this part we compare the \( W_s \) and \( W_w \) tests with the \( W_b \) test by investigating instantaneous causality relationships in U.S. macroeconomic datasets.

4.2.1 Money Supply and Inflation in the United States. The relationship between money supply and inflation is fundamental in the macroeconomic theories for explaining the influence of monetary policy on economy. For instance, the quantity theory of money assumes a proportional relationship between money supply and the price level. The reader is referred to Case, Fair, and Oster (2011) or Mankiw and Taylor (2006) concerning the theoretical links which can be made between money supply and inflation. Many studies investigate this relation from an empirical point of view. Their results are, however, ambiguous. For instance, Turnovsky and Wohar (1984) used a simple macro model to investigate the relationship and find that the rate of inflation is independent of the monetary growth rate in the United States over the period 1923–1960, while Benderly and Zwick (1985) or Jones and Uri (1987) gave some evidence of a relationship over the respective periods 1955–1982 and 1953–1984. Here we investigate the hypothesis of an instantaneous causal relationship between money supply and inflation in the United States over the period 1979–1995.

The data considered here are the M1 money stock (M1) and the Producer Price Index for all commodities (PPIACO). The M1 represents the money supply and PPIACO the inflation from the point of view of producers. The M1 index is provided by the Board of Governors of the Federal Reserve System while the PPIACO index is provided by the U.S. Department of Labor. The data are taken from 04/1979 to 12/1995 with a monthly frequency. The length of the series is \( T = 200 \). The data are available on the web site of the Federal Reserve Bank of St. Louis (Series ID: M1 and PPIACO). Their first differences denoted by \( \Delta M1 \) and \( \Delta PPIACO \) are plotted in Figure 2.

We adjusted a VAR(1) model to the first differences of the series. The autoregressive order is chosen by using portmanteau tests adapted to the framework of VAR processes with time-varying covariance (see Patilea and Raïssi 2013 for details). The outcomes in Table 3 suggest that the model is well fitted. The estimation of the model by the OLS method is given in Table 4.

The residuals of the VAR(1) estimation are next recovered to test the null hypothesis of constant unconditional residual variance. A corrected Inclan-Tiao test proposed by Sansó, Aragó,
Figure 2. Evolution of the variations $\Delta M1$ and $\Delta PPI$.  

Table 3. The $p$-values of the Box–Pierce test adapted to our nonstandard framework for checking the adequacy of the VAR(1) model adjusted to the U.S. M1 and inflation data  

| Number of lags | 3       | 6       | 12      |
|----------------|---------|---------|---------|
| $BP_{OLS}$     | 0.4384  | 0.8169  | 0.8165  |
| $[1.1198]$     | $[1.9071]$ | $[3.2870]$ |

NOTE: The corresponding statistics are displayed into brackets. The $BP_{OLS}$ corresponds to the portmanteau test based on the OLS proxies of the study’s.  

Table 4. The OLS estimators of the matrix $A_{01}$ (see Equation (2.1)) for the VAR(1) model adjusted to the U.S. M1 and inflation data  

| $A_{01}$      | 0.643 [0.064] | -1.124 [0.360] | -0.009 [0.007] | 0.439 [0.102] |

NOTE: Standard deviations of the parameters are displayed in brackets.  

Table 5. The $p$-values of the $W_{st}$, $W_{w}$, and $W_{b}$ tests for testing instantaneous causality between the U.S. M1 and inflation data  

| Test  | $W_{st}$  | $W_{w}$  | $W_{b}$  |
|-------|-----------|----------|----------|
| $p$-values | 0.268 [1.225] | 0.201 [1.632] | 0.058 [10.54] |

NOTE: The corresponding test statistics are displayed in brackets.  

and Carrion (2004) is used for this purpose. It is based on the following statistic:

$$\kappa = \sup_k \frac{|C_k - k\hat{\sigma}^2|}{\sqrt{T(\hat{\eta}^4 - \hat{\sigma}^4)}},$$

where $C_k = \sum_{t=1}^{k}(u_t - \bar{u})^2$, $\hat{\sigma}^2 = T^{-1}C_T$ and $\hat{\eta}^4 = T^{-1}\sum_{t=1}^{k}(u_t - \bar{u})^4$ with $u_t$ the studied residuals. Under the null hypothesis of constant unconditional variance errors and if $u_t$ is iid with $E[(u_t - \bar{u})^4] < +\infty$, the asymptotic distribution of the test is given by

$$\kappa \Rightarrow \sup_r |W^*(r)|,$$

where $W^*(r) = W(r) - rW(1)$ is a Brownian bridge and $W(r)$ is a standard Brownian motion. The values of the $\kappa$ statistic are 1.219 and 1.214 for the two residuals components of the VAR(1) model, while the critical values at the significance level of 10% and 5% are, respectively, 1.167 and 1.3 (see Table 1 in Sansó, Aragó, and Carrion 2004). Let us recall that our bootstrap $W_b$ test is valid whatever the behavior of the unconditional residual variance contrary to the standard $W_{st}$ and White $W_w$ tests.

Next, the tests studied in this article are implemented. Note that we used 399 bootstrap iterations for the bootstrap $W_b$ test. From Table 5 we see that the $p$-value of the $W_b$ test is quite different from those of the standard $W_{st}$ and White $W_w$ tests. For instance the null hypothesis of no instantaneous causality is rejected by the $W_b$ test for a significance level of 10% on the contrary to the other tests. These observations can be explained by the covariance structure of the innovations. Indeed the nonparametric estimation of this covariance structure plotted in Figure 3 shows that $\Sigma^{12}(r)$ seems not null over the considered period while its seems that $\int_0^1 \Sigma^{12}(r)dr \approx 0$.  

4.2.2 Merchandise Trade Balance and Balance on Services in the United States. The merchandise trade balance and the balance on services can be seen as indicators of the economic health of a country. The U.S. merchandise trade balance is the account which redraws the value of the exported goods and the value of the imported goods. The U.S. balance on services is similarly the account which redraws the value of the exported services and the value of the imported services. Here, we search to quantify if it exists an instantaneous causality relation between these two macroeconomic indicators. The data are provided by the Bureau Analysis of the U.S. Department of Commerce and taken from 01/1960 to 01/2011 with quarterly frequency. The length of the series is $T = 204$. They are available on the web site of the Federal Reserve Bank of St. Louis (Series ID : BOPBM and BOPBSV).

Similarly to the first dataset, we consider the first differences of the data (see Figure 4). A VAR(2) model is adjusted to the data (estimation results not reported here). The adequacy
of the model is again checked using portmanteau tests which are valid in our framework. The portmanteau test suggests to choose a VAR(2) model. Indeed the $p$-value of the BPOLS test is 0.65 [5.29] for five autocorrelations in the portmanteau statistics (the portmanteau statistic is given into brackets). Next the corrected Inclan-Tiao test proposed by Sansó, Aragó, and Carrion (2004) is applied to the two residuals components of the VAR(2) estimation. The $\kappa$ statistics are 2.223 and 2.591 for the two residuals components while the corresponding critical value is 1.547 (see Table 1 in Sansó, Aragó, and Carrion 2004). This allows to reject the null hypothesis of constant variance with a significant level of 1%.

The $p$-values of the tests for instantaneous causality are next computed from the residuals as for the first dataset. The outcomes displayed in Table 6 show that the tests have quite different results. In view of the nonconstant variance of the studied series (see Figure 3), the result corresponding to the $W_b$ test is more reliable.

Table 6. The $p$-values of the $W_{st}$, $W_w$, and $W_b$ tests for testing the instantaneous causality between the U.S. balance data

|       | $W_{st}$ | $W_w$ | $W_b$ |
|-------|---------|-------|-------|
| $p$-value | 0.0498 [3.848] | 0.341 [0.907] | 0.441 [190.142] |

NOTE: The corresponding test statistics are displayed in brackets.

5. CONCLUSION

This article studies the problem of testing instantaneous causality in the important case where the unconditional covariance is time-varying. We found that the Wald tests based on the assumption of constant unconditional covariance may have no power in this nonstandard framework. As a consequence, we propose bootstrap test for testing the instantaneous causality hypothesis when the unconditional covariance structure is time-varying. In particular this test is shown to be consistent, and we illustrate these theoretical results through a set of numerical
experiments. We show that the tests based on the stationary assumption are outperformed by the bootstrap test. The outcomes obtained from macroeconomic datasets suggest that the classical Wald tests may deliver results which are quite different from the bootstrap test.

**SUPPLEMENTARY MATERIAL**

The supplementary material contains the proof of Equation (3.1) and Proposition 1. Some numerical experiments are also provided for illustrating the effect of lag length selection in VAR models prior testing instantaneous causality.

[Received March 2013. Revised February 2014.]

**REFERENCES**

Amendola, A., and Franço, C. (2009), “Concepts and Tools for Nonlinear Time Series Modelling,” in Handbook of Computational Econometrics, eds. D. A. Belsley, and E. J. Kontogiorghos, New York: Wiley, chapter 10. [46]

Andrews, B., Davis, R. A., and Jay Breidt, F. (2006), “Maximum Likelihood Estimation for All-Pass Time Series Models,” Journal of Multivariate Analysis, 97, 1678–1699. [46]

Ashkenfelter, O., and Card, D. (1997), “Testing for a Unit Root in the Presence of Nonconstant Unconditional Covariance,” Econometrica, 53, 715–741. [46]

Azrak, R., and Melard, G. (2006), “Asymptotic Properties of Quasi-Maximum Likelihood Estimators for Arma Models With Time-Dependent Coefficients,” Statistical Inference for Stochastic Processes, 9, 279–330. [46]

Bahadur, R. (1960), “Stochastic Comparison of Tests,” Annals of Mathematical Statistics, 31, 276–295. [46]

Bai, J. (2000), “Vector Autoregressive Models With Structural Changes in Regime Models And Models of the Labour Market,” The Review of Economic Studies, 49, 761–782. [46]

Bauwens, L., Laurent, S., and Rombouts, J. V. (2006), “Multivariate Garch Models: A Survey,” Journal of Applied Econometrics, 21, 79–109. [46]

Benderly, J., and Zwick, B. (1985), “Inflation, Real Balances, Output, and Stock Prices,” Journal of Applied Econometrics, 9, 279–330. [46]

Benderly, J., and Zwick, B. (1985), “Inflation, Real Balances, Output, and Stock Prices,” Journal of Applied Econometrics, 9, 279–330. [46]

Benderly, J., and Zwick, B. (1985), “Inflation, Real Balances, Output, and Stock Prices,” Journal of Applied Econometrics, 9, 279–330. [46]