Econometric Theory, **0**, 2025, 1–29.
doi:10.1017/S0266466624000409

A NONPARAMETRIC TEST FOR INSTANTANEOUS CAUSALITY WITH TIME-VARYING VARIANCES

Jilin Wu
Xiamen University

Ruike Wu
Shanghai University of Finance and Economics

Zhijie Xiao
Boston College

This paper proposes a consistent nonparametric test with good sampling properties to detect instantaneous causality between vector autoregressive (VAR) variables with time-varying variances. The new test takes the form of the U-statistic, and has a limiting standard normal distribution under the null. We further show that the test is consistent against any fixed alternatives, and has nontrivial asymptotic power against a class of local alternatives with a rate slower than $T^{-1/2}$. We also propose a wild bootstrap procedure to better approximate the finite sample null distribution of the test statistic. Monte Carlo experiments are conducted to highlight the merits of the proposed test relative to other popular tests in finite samples. Finally, we apply the new test to investigate the instantaneous causality relationship between money supply and inflation rates in the USA.

1. INTRODUCTION

The concept of causality, as introduced by Granger (1969), plays a key role in analyzing dynamic relationships between time series. In studying Granger causality, predictability is the central issue, which is of great importance to economists, policymakers, and investors. Testing for Granger-causality has become a standard procedure to assess whether changes in one variable can help explain movements in another variable. Most existing literature focuses on Granger causality in conditional means. For example, Sims (1972) and Geweke (1982, 1984) investigate

Jilin Wu acknowledges the supports from the National Natural Science Foundation of China (Grant No. 72371213) and from the NSFC Basic Science Center Project for Econometric Modeling and Economic Policy Studies (Grant No. 71988101). We are most grateful to the Editor, Peter Phillips, for carefully going through an earlier version of the paper and providing many pertinent comments and constructive suggestions that markedly improved the paper. We also would like to thank the Co-Editor (Xu Cheng), and three referees for their constructive comments and suggestions. The authors contributed equally to this paper and are credited in alphabetical order. Address correspondence to Jilin Wu, Department of Finance, School of Economics, Gregory and Paula Chow Institute for studies in Economics, and MOE Key Laboratory of Econometrics, Xiamen University, Xiamen, China, e-mail: rainforest1061@126.com.

linear Granger causality based on VAR models. Dufour and Taamouti (2010) extend Geweke's work and design a test that can study short-run and long-run Granger causality separately. Li et al. (2012) and Song and Taamouti (2018) develop new test statistics that can be used to detect nonlinear Granger causality in conditional means. There is also other work considering testing for Granger causality in conditional distributions, see Taamouti, Bouezmarni, and El Ghouch (2014), Hu and Liang (2014), and Song and Taamouti (2021).

Instantaneous causality is also one type of causality proposed by Granger (1969). We say that there exists instantaneous causality between Y_{1t} and Y_{2t} if the prediction of Y_{2t}, given the past values of $Y_t = \left(Y'_{1t}, Y'_{2t}\right)'$, can be improved by adding the current information in Y_{1t} (see Lütkepohl, 2005, p. 42). If $\{Y_t\}_{t=1}^{T}$ take the form of a VAR representation, then testing instantaneous causality is equivalent to examining whether the off-diagonal elements of the error covariance matrix are equal to zero (see Lütkepohl, 2005, pp. 46–47). In this case, the standard Wald test and the heteroskedasticity-robust Wald test can serve this purpose. The importance of caring about instantaneous causality can be summarized as follows: First, it is quite common that the releases of macroeconomic data are asynchronous, particularly for those collected at a low frequency (e.g., monthly, quarterly, and yearly). For example, two sequences of time series data may be available monthly and the actual causal lag in one sequence is only a couple of days. Temporal aggregation is also a similar reason for resulting in possible instantaneous causality. In this case, testing for instantaneous causality can help reveal causal relationships among the considered variables. Second, as pointed out by Hyvärinen et al. (2010) and Faes and Nollo (2010), whether instantaneous causality is considered or not in VAR models can significantly impact the estimated values of time-lagged coefficients. Neglecting instantaneous influences may result in misleading interpretations of causal effects measured by the time-lagged coefficients. Therefore, it is crucial to test for instantaneous causality first, enabling the selection of an appropriate model under the specified setups for accurate estimation of lagged coefficients. Third, instantaneous causality can sometimes occur because of omitted variables. For instance, this may happen when there are actually three variables $\{Y_{1t}, Y_{2t}, W_{t-1}\}$, and the third variable, W_{t-1}, influences the first two, which are independent within the group of three variables. However, if the third variable is missing and thus marginalized, observed instantaneous causality may occur. For detailed discussion on this, see Granger (1988). Clearly, in this case, testing instantaneous causality can indirectly help prejudge whether there exists such an omitted variable.

Up to now, there are many theoretical and empirical studies concerning instantaneous causality. Pierce and Haugh (1977) review and compare several recent methodologies proposed for empirically examining causal relationships between variables, which also include the detection of instantaneous causality. Granger (1988) discusses three possible explanations for the presence of instantaneous causality. Breitung and Swanson (2002) investigate the impact of temporal aggregation on detecting instantaneous causality. Hafner (2009) discusses the effects of temporal aggregation on causality and forecasting in multivariate GARCH processes, and finds that instantaneous causality also occurs in volatility. In addition, Faes and Nollo (2010) and Faes et al. (2013) employ instantaneous causality to study either fast (within-sample) physiologically meaningful interactions or non-physiological effects in physiological data. Raïssi (2011) extends the Wald tests for instantaneous causality to a more generalized framework that allows for nonlinear dynamics of unknown forms in error terms.

However, the above-mentioned tests are constructed under the assumption of constant unconditional variances. In fact, more and more evidence shows that time variation in unconditional variances is a common feature in macroeconomic and financial data. For example, Sensier and Van Dijk (2004) report that about 80% of 214 U.S. macroeconomic time series displayed breaks in variances during the period 1959–1999. Kim and Nelson (1999) and Justiniano and Primiceri (2008) demonstrate that the volatilities of U.S. major macroeconomic variables, especially GDP, have declined since the 1980s. Clark (2011) provides empirical evidence strongly suggesting that the volatilities of U.S. macroeconomic variables rose sharply during the severe recession of 2007–2009. Similarly, Andreou and Ghysels (2002) discover that the Asian and Russian financial crises have caused obvious structural breaks in the volatility dynamics of international financial markets. Mikosch and Stărică (2004) and Liu and Maheu (2008) also find strong evidence of structural change relating to shifts in the variance of S&P 500 returns.

The presence of time-varying variances could invalidate conventional statistical inference and hypothesis testing. Lamoureux and Lastrapes (1990) and Granger and Hyung (2004) prove that structural change in unconditional variance causes spurious persistence and long memory effects in volatility dynamics. Hamori and Tokihisa (1997), Kim, Leybourne, and Newbold (2002), and Cavaliere (2005) show that ignoring the effect of a volatility shift results in significant over-sized distortion in unit root tests. Hansen (1995), Xu and Phillips (2008), and Linton and Xiao (2019) point out that time-varying variances lead to inefficient estimation and unreliable inference in parametric and nonparametric models. Hammoudeh and Li (2008) and Groen, Paap, and Ravazzolo (2013) conclude that the macroeconomic predictors not allowing for structural breaks in variances result in very poor point and density forecasts. Of course, the presence of time-varying variances also damages Granger causality testing. Vilasuso (2001) shows that unconditional heteroskedasticity leads to an erroneous claim that a statistically significant causal relation exists in conditional means. Van Dijk, Osborn, and Sensier (2005) also find that structural breaks in unconditional variances will distort testing size and power when examining Granger causality in second moments. Patilea and Raïssi (2012) prove that the standard Wald test for Granger causality in VAR models is invalid in the presence of time-varying variances, and thus suggest adaptive Wald tests to improve the testing reliability. Similarly, the Wald-type tests for instantaneous causality do not provide suitable critical values under time-varying variances. Moreover, they may suffer from severe power loss when the integration of time-varying covariance is close to zero. Hence it is also of great necessity to develop powerful tests for instantaneous causality that are robust to time-varying variances.

In order to avoid power loss incurred by the Wald-type tests, Gianetto and Raïssi (2015) propose a cumulative-sum (cusum) test for instantaneous causality. Because the new test has an asymptotically nonstandard null distribution and depends on unknown time-varying variance–covariance structures, a wild bootstrap method is therefore suggested to improve its testing robustness. Monte Carlo shows that the new test of Gianetto and Raïssi (2015) controls Type I error reasonably well and achieves decent power gains, clearly outperforming those traditional tests. However, it is well-known that the cusum-type tests mainly consider one-time shift as the alternative, making them the most appropriate (in terms of the most powerful) when the parameters of the underlying statistical model only contain one single changepoint, and they may not have good power against multiple breaks or smooth change, see Chen and Hong (2012) and Wu and Xiao (2018b). The test of Gianetto and Raïssi (2015) is no exception. It is powerful to test for instantaneous causality with a single structural break in variances, but does not perform satisfactorily when the unconditional variances exhibit multiple structural breaks or smooth structural change. This observation is also consistent with the findings in our Monte Carlo experiments of Section 4. In practice, more and more evidence demonstrates that financial and macroeconomic data are better characterized by multiple breaks or smooth change, see Bai and Perron (1998), Hansen (2001), and Fu, Hong, and Wang (2023). Hence, it is desirable to develop an omnibus test to diagnose instantaneous causality that is robust to various forms of structural changes in unconditional variances.

The purpose of this article is to propose a consistent nonparametric test for instantaneous causality in the presence of time-varying variances. The idea is to first estimate the instantaneous causality nonparametrically and then develop an L_2-type test statistic by comparing the norms of the nonparametric estimators with zero. After expanding and modifying the expression, as well as dropping the negligible terms, our final test statistic can be formulated as a convenient U-statistic form. The appealing advantages of our proposed test statistic are summarized as follows: First, the proposed test is consistent against various alternatives that deviate from the null hypothesis, and no prior information about the alternatives is required. Specifically, our test allows for smooth structural changes and sudden single or multiple structural breaks with unknown breakdates or unknown number of breaks in the unconditional covariance matrix. Second, unlike many tests for Granger causality in the literature, which often have nonstandard asymptotic null distributions, the new test is automatically centered and, after appropriate standardization, has an asymptotically standard normal distribution under the null hypothesis of no instantaneous causality. Third, a wild bootstrap procedure is proposed to better approximate the finite sample null distribution of the test statistic.

The structure of this article is organized as follows: Section 2 presents the basic model, describes the hypotheses of interest, and constructs the nonparametric test.

In Section 3.1, we study the asymptotic distribution of the proposed test statistic under the null hypothesis. Section 3.2 investigates the asymptotic power properties of our test statistic under a fixed alternative and a sequence of local alternatives. Section 3.3 discusses lag length identification and bandwidth selection in testing procedures, and Section 3.4 proposes a wild bootstrap to improve the testing accuracy under the null. We report the results of Monte Carlo simulations to assess the finite sample performance of our test statistic compared with other popular tests in Section 4. Section 5 applies the proposed test statistic to examine the instantaneous causality between money supply and inflation in the United States. Section 6 considers one possible extension, and the conclusion is given in Section 7. Proofs and other simulation results are relegated to Supplementary Material.

2. THE BASIC MODEL AND THE TEST

Let $Y_t = \left(Y'_{1t}, Y'_{2t}\right)'$ be a d-dimensional vector of multivariate time series with Y_{it} of dimension $d_i, i = 1,2$. It is assumed that Y_t is generated by the following stable vector autoregressive model of lag order p:

$$\begin{cases} Y_t = \sum_{j=1}^{p} A_j Y_{t-j} + u_t, \\ u_t = G_t \varepsilon_t, t = 1,2,\ldots,T, \end{cases} \tag{1}$$

where the $d \times d$ parameter matrices A_j, $j \in \{1,\ldots,p\}$, are such that $\det\left(I_d - \sum_{j=1}^{p} A_j z^j\right) \neq 0$ for all $|z| \leq 1$. The error term u_t is the product of two components, the long-run component G_t and the short-run component ε_t. The $d \times d$ matrix G_t is a lower triangular nonsingular matrix with positive diagonal elements, and ε_t is a $d \times 1$ martingale difference sequence (m.d.s.) satisfying $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$ and $E(\varepsilon_t \varepsilon'_t|\mathcal{F}_{t-1}) = I_d$, where $\mathcal{F}_{t-1}$ is the information set at time $t-1$. Here G_t is assumed to be a nonstochastic matrix. $\Sigma_t = G_t G'_t$ is assumed to be a symmetric and positive definite covariance matrix and it represents possibly time-varying unconditional multivariate heteroskedasticity, see, e.g., Hafner and Linton (2010), Xu (2012), and Patilea and Raïssi (2013), for more discussion.

Let $u_t = \left(u'_{1t}, u'_{2t}\right)'$, where $u_{it} = G_{it}\varepsilon_t$ has dimension d_i, $i = 1,2$, with $G_t = \left(G'_{1t}, G'_{2t}\right)'$, and define also Σ_t^{12} the upper right block of $\Sigma_t = \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{21} & \Sigma_t^{22} \end{pmatrix}$, where $\Sigma_t^{21} = \left(\Sigma_t^{12}\right)'$. We are interested in testing whether or not there is instantaneous causality between Y_{1t} and Y_{2t}. In VAR models, it is well known that there is no instantaneous causality between Y_{1t} and Y_{2t} if and only if all elements of Σ_t^{12} are equal to zero. Therefore, testing for instantaneous linear causality between Y_{1t} and Y_{2t} amounts to testing the following hypothesis

$$\mathbb{H}_0 : \Sigma_t^{12} = \mathbf{0}_{d_1 \times d_2} \text{ for all } t, \tag{2}$$

where $\mathbf{0}_{d_1 \times d_2}$ represents a $d_1 \times d_2$ zero matrix. The null indicates no instantaneous causality between Y_{1t} and Y_{2t} in the observation period (see Lütkepohl, 2005, pp. 46–47). The alternative hypothesis is

$$\mathbb{H}_A : \Sigma_t^{12} \neq \mathbf{0}_{d_1 \times d_2}. \tag{3}$$

Under the alternatives, Σ_t^{12} can be a nonzero constant matrix, be smoothly changing over time or contain unknown finite breakpoints. Following Robinson (1989) and Cai, Wang, and Wang (2015), we assume that $G_t = G(r_t)$, where $r_t = t/T$, is a deterministic function of sample-size scaled time index, which also implies that $\Sigma_t = \Sigma(r_t)$. Then, the null in (2) and the alternative in (3) become

$$\mathbb{H}_0 : \Sigma^{12}(r) = \mathbf{0}_{d_1 \times d_2}, \text{ for all } r \in [0,1], \tag{4}$$

versus

$$\mathbb{H}_A : \Sigma^{12}(r) \neq \mathbf{0}_{d_1 \times d_2} \text{ for } r \in S \subseteq [0,1],$$

where S is a set of positive Lebesgue measure. **(5)**

If $\Sigma^{12}(r)$ were observable, it is therefore natural to measure the discrepancy between $\Sigma^{12}(r)$ and the zero matrix $\mathbf{0}_{d_1 \times d_2}$ for all $0 \leq r \leq 1$, then an L_2-type distance can be designed as follows:

$$\lambda = \int_0^1 \phi(r) \left\| vec\left(\Sigma^{12}(r)\right)\right\|^2 dr, \tag{6}$$

where $\phi(\cdot)$ is a positive weighting function over [0,1], $\|\cdot\|$ represents the usual Euclidean norm, and $vec(\cdot)$ is the operator that stacks the columns of matrix as a vector. Under the null, λ should be close to zero, and under the alternative, λ should be a strictly positive constant. The L_2-type function has been widely employed to test model specifications, see Härdle and Mammen (1993) and Chen and Hong (2012).

Construction of a formal test based on (6) requires estimation of $\Sigma^{12}(\cdot)$. Because $E(u_{1t}u_{2t}'|\mathcal{F}_{t-1}) = \Sigma^{12}(r_t)$, we consider the following regression model

$$u_{1t}u_{2t}' = \Sigma^{12}(r_t) + e_t, \tag{7}$$

where e_t is an m.d.s. over time. If $\left\{u_{1t}u_{2t}'\right\}_{t=1}^T$ were observable, we could estimate $\Sigma^{12}(\cdot)$ nonparametrically based on (7). However, the innovations $\{u_t\}_{t=1}^T$ are unobservable in practice, but they can be replaced by the OLS residuals estimated from the first equation in (1). Denote $\Pi = \left((vec(A_1))', \ldots, \left(vec\left(A_p\right)\right)'\right)'$ and $X_{t-1} = \left(Y_{t-1}', Y_{t-2}', \ldots, Y_{t-p}'\right)'$. The VAR(p) model in (1) can be rewritten as

$$Y_t = \left(X_{t-1}' \otimes I_d\right)\Pi + u_t, \tag{8}$$

where $\otimes$ is the Kronecker product. Then, the OLS estimator for Π is given by

$$\hat{\Pi} = \left(\sum_{t=p+1}^T X_{t-1}X_{t-1}' \otimes I_d\right)^{-1} vec\left(\sum_{t=p+1}^T Y_t X_{t-1}'\right), \tag{9}$$

and the estimated residual is given by $\hat{u}_t = Y_t - \left(X_{t-1}' \otimes I_d\right)\hat{\Pi}$. Based on $\hat{u}_t = \left(\hat{u}_{1t}', \hat{u}_{2t}'\right)'$, we have the following Nadaraya–Watson estimator for $\Sigma^{12}(r)$:

$$\hat{\Sigma}^{12}(r) = \frac{\sum_{t=p+1}^{T} k_h(r_t - r)\,\hat{u}_{1t}\hat{u}_{2t}'}{\sum_{t=p+1}^{T} k_h(r_t - r)}, \tag{10}$$

where $k_h(\cdot) = k(\cdot/h)\big/h$, $k(\cdot)$ is a kernel function defined on $[-1,1]$, and h is a bandwidth parameter satisfying $h \to 0$ and $Th \to \infty$ as $T \to \infty$. With $\hat{\Sigma}^{12}(r)$ at hand, the L_2-type test statistic corresponding to (6) can be constructed via the following quadratic form

$$\hat{\lambda}_T = \int_0^1 \phi(r) \left\| vec\left(\hat{\Sigma}^{12}(r)\right)\right\|^2 dr. \tag{11}$$

Remark 1. It is possible to extend the testing procedure proposed in this paper to test the null hypothesis of constant instantaneous causality over time, that is, $\mathbb{H}_0 : \Sigma_t^{12} = \bar{\Sigma}^{12}$ for all t, where $\bar{\Sigma}^{12}$ is a $d_1 \times d_2$ nonzero constant matrix.

Remark 2. Note that the nonparametric estimator $\hat{\Sigma}^{12}(r)$ is not consistent in the neighboring regions around the breakpoints in $\Sigma^{12}(r)$. However, the number of breakpoints is assumed to be finite (see Assumption 3 below), so the set of inconsistent estimators has zero measure and is negligible in integration.

Under $\mathbb{H}_0$, $\hat{\Sigma}^{12}(r)$ is a consistent estimator of zero matrix, so the test statistic $\hat{\lambda}_T$ converges to 0; under $\mathbb{H}_A$, $\hat{\Sigma}^{12}(r)$ is a consistent estimator of $\Sigma^{12}(r)$ that is significantly different from the zero matrix except in the neighboring regions of breakpoints. Consequently the statistic $\hat{\lambda}_T$ converges to a strictly positive constant. Thus, any significant departure of $\hat{\lambda}_T$ from 0 is evidence of instantaneous causality between Y_{1t} and Y_{2t}. Notice that (10) can be rewritten as

$$\sum_{t=p+1}^{T} k_h(r_t - r)\, vec\left(\hat{\Sigma}^{12}(r)\right) = \sum_{t=p+1}^{T} k_h(r_t - r)\, vec\left(\hat{u}_{1t}\hat{u}_{2t}'\right). \tag{12}$$

By taking the weighting function $\phi(r) = \left(\frac{1}{T}\sum_{t=p+1}^{T} k_h(r_t - r)\right)^2$, we rewrite (11) as

$$\hat{\lambda}_T = \frac{1}{T^2 h} \sum_{t=p+1}^{T} \sum_{s=p+1}^{T} a_{t,s} \hat{m}_t' \hat{m}_s, \tag{13}$$

where $a_{t,s} = h\int_0^1 k_h(r_t - r)\, k_h(r_s - r)\, dr$ and $\hat{m}_t = vec(\hat{u}_{1t}\hat{u}_{2t}')$. We notice that $a_{t,s}$ is actually a convolution kernel which converges to $\int_{-1}^{1} k\left(u + \frac{r_s - r_t}{h}\right) k(u)\, du$ as $T \to \infty$. As pointed out by Cai et al. (2015), one does not even have to use the convolution kernel. Simply replacing it with $k_{s,t} \equiv k\left(\frac{s-t}{Th}\right)$ will not affect the essence of the test statistic since the local weight property is preserved. Finally, we use a leave-one-out estimator to remove a nonzero center term from the test statistic and obtain the following kernel-smoothed test statistic:

$$\tilde{\lambda}_T = \frac{1}{T^2 h} \sum_{t=p+1}^{T} \sum_{s \neq t} k_{s,t} \hat{m}_t' \hat{m}_s. \tag{14}$$

If the null hypothesis is true, $\tilde{\lambda}_T$ should be close to zero, and has an asymptotic $\mathbb{N}(0,1)$ distribution after appropriate standardization. Under the alternative, $\tilde{\lambda}_T$ will be distant away from zero. Juhl and Xiao (2005) and Wu and Xiao (2018a) have employed similar U-statistic tests to examine parametric model specifications in univariate time series. Thus, our proposed test can be regarded as an important extension to diagnose instantaneous causality in multivariate time series.

3. ASYMPTOTIC RESULTS

In what follows we first establish the asymptotically normal distribution of $\tilde{\lambda}_T$ under the null, then proceed to study its asymptotic power properties under a fixed alternative and a sequence of local alternatives.

3.1. Asymptotic Null Distribution

To facilitate asymptotic analysis of the proposed test under the null, we introduce the following regularity conditions.

Assumption 1. The autoregressive matrix parameters $\{A_i\}_{i=1}^{p}$ are such that $\det[A(z)] \neq 0$ for any $|z| \leq 1$, where $A(z) = I_d - \sum_{i=1}^{p} A_i z^i$.

Assumption 2. (i) $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}) = I_d$, and $\Upsilon = E\left(\varepsilon_t\varepsilon_t' \otimes \varepsilon_t\varepsilon_t'\right)$, where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$. (ii) $\{u_t\}_{t=1}^{T}$ are β-mixing processes with mixing coefficient $\beta(j) = \sup_s E\left\{\sup_{A \in \mathcal{G}_{s+j}^{\infty}} \left|\Pr\left(A|\mathcal{G}_{-\infty}^{s}\right) - \Pr(A)\right|\right\} \to 0$, as $j \to \infty$, where $\mathcal{G}_s^t$ is the σ-field generated by $\{u_n, n = s, \ldots, t\}$, and $\{\beta(j)\}_{j=1}^{\infty}$ satisfy $\sum_{j=1}^{\infty} j^2 \beta(j)^{\delta/(1+\delta)} < C$ for some $0 < \delta < 1$. (iii) $\sup_t E\|u_{it}\|^8 < C$, and $\sup_{t,s,t',s'} E\left|u_{it}'u_{is}u_{it'}'u_{is'}\right|^{4(1+\delta)} < C$, $i = 1,2$, where at least two time indices of $\left\{t,s,t',s'\right\}$ are not equal.

Assumption 3. $G(\cdot)$ is a $d \times d$ lower triangular nonsingular matrix with positive diagonal elements, where the components of the matrix $G(\cdot) := \left\{g_{ij}(\cdot)\right\}, i,j = 1,\ldots,d$, are bounded deterministic functions and have continuous second derivatives except for a finite number of points on $[0,1]$. Moreover, $\Sigma(\cdot) = G(\cdot)G(\cdot)'$ is symmetric and positive definite on $[0,1]$.

Assumption 4. $k(\cdot) : [-1,1] \to R^+$ is a symmetric and bounded probability density.

Assumption 5. As $T \to \infty$, $h \to 0$ and $Th^2 \to \infty$.

Assumption 1 is the usual stationarity condition for VAR models, and ensures that the eigenvalues of the matrix $A(z)$ all lie outside the unit root. As a result, the unit root or explosive processes are ruled out in our model. Additionally, the true autoregressive order p in VAR models is supposed to be known. In practice, the lag length p can be determined by some testing procedures, which will be discussed in Section 3.3. Under Assumption 2(ii), the temporal dependence of $\{u_t\}_{t=1}^T$ decreases fast as the time distance increases and thus is asymptotically negligible. Assumption 2(iii) imposes some moment conditions on $\{u_{it}\}_{t=1}^T$, $i = 1,2$, to help develop a martingale central limit theorem for the new test statistic $\tilde{\lambda}_T$. A similar moment assumption is also given by Hsiao and Li (2001), who propose a U-statistic test for conditional heteroskedasticity. Assumption 3 allows for smooth structural change and finite abrupt breaks with known or unknown breakpoints in $\Sigma(\cdot)$. For abrupt breaks, the break size is bounded. The literature imposing similar conditions on $\Sigma(\cdot)$ can be referred to Hafner and Linton (2010), Xu (2012), and Patiliea and Raïssi (2013). Assumption 4 implies that $\int_{-1}^1 k(u)du = 1$, $\int_{-1}^1 uk(u)du = 0$ and $\int_{-1}^1 u^2k(u)du < \infty$. Assumption 5 is a standard assumption in kernel regression literature.

In order to obtain a valid asymptotic test, we standardize $\tilde{\lambda}_T$ by $Th^{1/2}$ and a variance estimator, say, $\hat{\sigma}_T^2$, to achieve a standard normal limit. Formally, we construct the final test statistic:

$$J_T = \frac{Th^{1/2}\tilde{\lambda}_T}{\hat{\sigma}_T}, \tag{15}$$

where

$$\hat{\sigma}_T^2 = \frac{2}{T^2h}\sum_{t=p+1}^{T}\sum_{s\neq t} k_{s,t}^2\left(\hat{m}_t'\hat{m}_s\right)^2 \tag{16}$$

is a consistent estimate of

$$\sigma^2 = 2\left(vec(\Upsilon)\right)'\left(\int_0^1 \left(\Omega(r)\otimes\Omega(r)\right)dr\right)vec(\Upsilon)\int_{-1}^1 k^2(u)du \tag{17}$$

under the null and the alternatives, where $\Omega(\cdot) = G_2'(\cdot)G_2(\cdot)\otimes G_1'(\cdot)G_1(\cdot)$.

We now state the asymptotic distribution of J_T under the null.

THEOREM 1. *Suppose that Assumptions 1-5 hold, then under the null of* $\Sigma_t^{12} = \mathbf{0}_{d_1\times d_2}$, *as* $T\to\infty$, *we have*

$$J_T \xrightarrow{d} \mathbb{N}(0,1). \tag{18}$$

Remark 3. The proof of Theorem 1 uses the fact that J_T is asymptotically equivalent to a degenerate U-statistic. However, the kernel function solely involves deterministic components of sample-size scaled time index and $\{u_t\}_{t=1}^T$ is not an independently and identically distributed sequence. Thus, we cannot directly apply the result of Zheng (1996) and Hsiao and Li (2001) to establish asymptotic normality. To obtain the asymptotic null distribution of J_T, we need to resort to the martingale central limit theorem of Brown (1971). Additionally, it is well known that the asymptotic null distributions of nonparametric kernel-based tests like J_T may not be well approximated by $\mathbb{N}(0,1)$ in finite samples, see, for example, Li and Wang (1998) and Lee and Ullah (2007). More accurate critical values of the finite sample distribution of J_T under the null can be obtained via a wild bootstrap procedure, see Section 3.4 for more discussion.

3.2. Asymptotic Power

Now, we study the power properties of the test under the alternatives. We first consider the power of the test under a fixed alternative. The consistency property of the test rejecting $\mathbb{H}_0$ for large values of J_T is stated in the following theorem.

THEOREM 2. *Suppose that Assumptions 1-5 hold, then under $\mathbb{H}_A$, for any sequence of nonstochastic constants $\left\{C_T = o\left(T\sqrt{h}\right)\right\}$, as $T \to \infty$, we have*

$$\Pr(J_T > C_T) \to 1. \tag{19}$$

Remark 4. From the proof of Theorem 2 in the Supplementary Material, it is shown that J_T diverges to positive infinity at the nonparametric rate $Th^{1/2}$ as $T \to \infty$. This implies that the proposed test is a one-sided test, and has an asymptotic unit power against any fixed alternatives.

In order to gain some insights into the local power of the test, the next theorem shows the behavior of J_T under the local alternatives given by

$$\mathbb{H}_{LA} : \Sigma^{12}(r) = j_T \pi(r), \tag{20}$$

where $\pi(\cdot): [0,1] \to R^{d_1 \times d_2}$ is a twice continuously differentiable function except for a finite number of discontinuity points. The term $j_T \pi(r)$ characterizes the potential departure of $\Sigma^{12}(r)$ from the null at the time $\lfloor Tr \rfloor$, where $\lfloor x \rfloor$ is the integer part of x. Here $j_T = j(T)$ governs the rate at which the local alternatives deviate from the null. For notational simplicity, we have suppressed the dependence of $\Sigma^{12}(r)$ on T. In the special case of $\pi(r) = \mathbf{0}_{d_1 \times d_2}$, for any $r \in [0,1]$, we obtain the null model.

THEOREM 3. *Suppose that Assumptions 1-5 hold, then under $\mathbb{H}_{LA}$ with $j_T = T^{-1/2}h^{-1/4}$, as $T \to \infty$, we have*

$$J_T \xrightarrow{d} \mathbb{N}(\mu, 1), \tag{21}$$

where $\mu = tr\left(\int_0^1 \pi(r)\pi(r)' dr\right) \Big/ \sigma$.

Remark 5. The "non-centrality parameter" μ represents the shift in charge of asymptotic local power against $\mathbb{H}_{LA}$. Clearly, when $\pi(r) = \mathbf{0}_{d_1 \times d_2}$ in (20), Theorem 3 is reduced to Theorem 1. Theorem 3 means that J_T has nontrivial power against the given alternatives (20) that diverge from the null at the rate of $T^{-1/2}h^{-1/4}$. Note that $T^{-1/2}h^{-1/4}$, which is the typical rate for nonparametric kernel tests, is slower than the parametric rate $T^{-1/2}$. Therefore, unlike the test proposed by Gianetto and Raïssi (2015), J_T cannot detect Pitman local alternatives that approach the null at the rate of $T^{-1/2}$. In contrast, the test of Gianetto and Raïssi (2015) enjoys nontrivial power against $\mathbb{H}_{LA}$ with the rate $T^{-1/2}$. However, their test is constructed only utilizing the information of the null hypothesis, and does not consider any information of the alternatives. As a result, the test of Gianetto and Raïssi (2015) is mainly good at detecting the instantaneous causality with a single breakpoint, although it also enjoys some power against the instantaneous causality with multiple breakpoints or smooth structural change. In contrast, our proposed nonparametric test is constructed by comparing the discrepancy between the nonparametric estimators of $\Sigma^{12}(r)$ under the alternatives and the zero matrix $\mathbf{0}_{d_1\times d_2}$ under the null, and any deviation from the null can be immediately captured by J_T. Hence, our new test has all-round power against various kinds of alternatives, including smooth structural change and abrupt breaks, which will be demonstrated in subsequent Monte Carlo experiments.

3.3. Lag Length Identification and Bandwidth Selection

The correct choice of lag length in the VAR model (1) is the first task. If the chosen lag order is too small, the test may produce misleading outputs; if a large lag order is chosen, then many unnecessary parameters are introduced into the model, which will cause power loss in the proposed test. The information criteria (see, e.g., Hannan and Quinn, 1979; Cavanaugh, 1997; Boubacar, 2012) and the portmanteau tests (see, e.g., Chitturi, 1974; Hosking, 1980) are the most recommended to help identify lag length in VAR models. However, the presence of time-varying variances invalidates these classic methods but the corrected portmanteau test developed by Patilea and Raïssi (2013) can be applied to situations where the time-varying variances are present. Specifically, the portmanteau statistic is given by

$$\tilde{Q}_m^{OLS} = T^2 \sum_{i=1}^{m} (T-i)^{-1}\, tr\left[\hat{\Gamma}'_{OLS}(i)\,\hat{\Gamma}^{-1}_{OLS}(0)\,\hat{\Gamma}_{OLS}(i)\,\hat{\Gamma}^{-1}_{OLS}(0)\right], \tag{22}$$

where $\hat{\Gamma}_{OLS}(i) = T^{-1}\sum_{t=i+1}^{T} \hat{u}_t \hat{u}'_{t-i}$ and $\hat{u}_t = Y_t - \left(X'_{t-1} \otimes I_d\right)\hat{\Pi}$. If the lag length p is correctly selected, Patilea and Raïssi (2013) prove that the statistic $\tilde{Q}_m^{OLS}$ converges to

$$U\left(\delta_m^{OLS}\right) = \sum_{i=1}^{d^2 m} \delta_i^{OLS} U_i^2$$

as $T \to \infty$, where $\boldsymbol{\delta}_m^{OLS} = \left(\delta_1^{OLS}, \ldots, \delta_{d^2m}^{OLS}\right)'$ is the vector of the eigenvalues of the matrix Δ_m^{OLS} (whose form is given in Theorem 4.1 of Patilea and Raïssi, 2013), and U_i is an independent $\mathbb{N}(0,1)$ variable. The limiting distribution of $U\left(\boldsymbol{\delta}_m^{OLS}\right)$ in the presence of time-varying variances is very different from a chi-square law, and its p-values cannot be obtained from the chi-square distribution. Therefore, a simulated algorithm is suggested to generate the critical values of $U\left(\boldsymbol{\delta}_m^{OLS}\right)$, which of course requires consistent estimation of the eigenvalue vector $\boldsymbol{\delta}_m^{OLS}$ in advance. The reader can be referred to Patilea and Raïssi (2013) and Imhof (1961) for more discussion on estimating $\boldsymbol{\delta}_m^{OLS}$ and computing the critical values of $U\left(\boldsymbol{\delta}_m^{OLS}\right)$.

The selection procedure is given as follows: We start with a VAR(1) model and test whether the choice of lag length is appropriate by comparing the statistic $\tilde{Q}_m^{OLS}$ with the simulated critical value of $U\left(\boldsymbol{\delta}_m^{OLS}\right)$ at the significance level α, If $\tilde{Q}_m^{OLS}$ is smaller than the given critical value, we choose $p=1$; otherwise, the next step is to estimate a VAR(2) model and perform the test again. This procedure is repeated until a non-rejection occurs. In practice, a maximum lag order $p_{\max}$ is commonly required to be prespecified due to limited sample sizes.

The proposed test statistic involves choosing the bandwidth parameter h, which also plays an important role in our testing procedure. Because the test statistic is derived from the nonparametric estimation of $\Sigma^{12}(\cdot)$, following Xia and Li (2002) and Li and Racine (2004) we consider the method of least-squares cross validation (CV) to choose the bandwidth h. Define a "leave-one-out" estimator

$$\hat{\Sigma}_{-t}^{12}=\frac{\sum_{s=p+1,s\neq t}^{T}k_h\left(r_s-r_t\right)\hat{u}_{1s}\hat{u}_{2s}'}{\sum_{s=p+1,s\neq t}^{T}k_h\left(r_s-r_t\right)}. \tag{23}$$

Then a data-driven choice of h based on (23) is given by

$$\hat{h}_{cv}=\underset{c_1T^{-1/5}\leq h\leq c_2T^{-1/5}}{\arg\min}\frac{1}{T-p}\sum_{t=p+1}^{T}\left\|\hat{u}_{1t}\hat{u}_{2t}'-\hat{\Sigma}_{-t}^{12}\right\|^2, \tag{24}$$

where c_1 and c_2 are two prespecified constants, and $T^{-1/5}$ is the optimal order of magnitude obtained by minimizing the asymptotic mean integrated squared error of $\hat{\Sigma}^{12}(\cdot)$. The data-driven bandwidth automatically adjusts to the data and leads to a test based on the trade-off between size and power. Under the null (Σ_t^{12} is irrelevant to t), $\hat{\Sigma}^{12}(\cdot)$ is unbiased and CV tends to minimize the integrated variance, thereby selecting a bandwidth close to the upper limit. Under the alternative, CV searches for a way of balancing bias and variance and settles down at an optimal (shrinking) bandwidth. The CV method does not affect the limiting distribution of our test statistic since it satisfies the requirement of Assumption 5 that $\hat{h}_{cv}\to 0$ and $T\hat{h}_{cv}^2\to\infty$ as $T\to\infty$. In finite samples, using the CV method may introduce some extra noise into size performance, but it can help achieve decent power[1]. In this article, we specify $h=1.03^{i-15}T^{-1/5}$, $1\leq i\leq 25$,

[1] Of course, the bandwidth produced by the CV method varies with the sample. Extension of our testing procedure allowing for a random bandwidth is possible along the lines of Li and Li (2010).

which implies that $c_1 = 0.661$ and $c_2 = 1.344$, and our Monte Carlo evidence demonstrates that the data-driven bandwidth selection procedure works well.

3.4. A Bootstrap Method

The test statistic J_T has an asymptotic standard normal distribution under the null hypothesis. However, empirical sizes of the kernel-based nonparametric test using asymptotic critical values may differ quite a bit from the nominal levels in finite sample applications, and can be sensitive to bandwidth selection. For this reason, a wild bootstrap procedure is proposed to improve the testing accuracy of J_T under the null. We show that the bootstrap-based test is asymptotically valid. The Monte Carlo results in Section 4 also show that the bootstrap-based test indeed has a much better size than the asymptotic test using normal critical values.

In the following, we outline the key steps in computing the bootstrapped test statistic.

(a) Determine the lag length p in the VAR(p) model, and obtain the OLS estimator $\hat{\Pi}$ by (9) as well as the corresponding OLS residuals $\hat{u}_t$. Select the bandwidth h and calculate the test statistic J_T by (15).

(b) Generate B bootstrap sets given by $\hat{m}_t^{(i)} = \xi_t^{(i)} vec(\hat{u}_{1t}\hat{u}_{2t}'), t \in \{p+1, \ldots, T\}$, and $i \in \{1, ..., B\}$, where the univariate random variable $\xi_t^{(i)}$ is taken from the i.i.d. standard Gaussian distribution, independent of the original samples $\{Y_t\}_{t=1}^T$.

(c) Calculate the test statistic $J_{T,(i)}^*$ in the same way as J_T with the bootstrapped $\hat{m}_t^{(i)}$ replacing the original $\hat{m}_t$, where $i \in \{1, ..., B\}$.

(d) Obtain the $(1-\alpha)th$ quantile $q_{1-\alpha}$ from the empirical distribution of $\left\{J_{T,(i)}^*\right\}_{i=1}^B$ as the cutoff value. Reject the null hypothesis at the level α if $J_T > q_{1-\alpha}$.

In the above procedure, we do not have to re-estimate the VAR model with the bootstrapped samples, and the OLS residuals are directly employed to generate the bootstrap residuals. This is motivated by the fact that the null is tested on the error covariance structure, so that we only consider the residuals in the test statistic. The wild bootstrap method is mainly designed to replicate the pattern of potential nonconstant covariance of the residuals. Alternative bootstrap methods could be considered but at the price of more computational burden.

Let $\xrightarrow{d^*}$ denote the convergence in distribution under the bootstrap law—conditional on the data and for almost all sample paths. The following theorem justifies asymptotic validity of the proposed bootstrap method.

THEOREM 4. *Suppose that Assumptions 1-5 hold, then under the null and the alternatives, as $T \to \infty$, we have*

$$J_T^* \xrightarrow{d^*} \mathbb{N}(0,1).$$

Remark 6. Theorem 4 shows that the bootstrapped test statistic J_T^* converges in distribution to $\mathbb{N}(0,1)$, thus providing an asymptotically valid procedure. By comparing J_T with the quantile $q_{1-\alpha}$ of J_T^*, we can judge whether the null is rejected or not. Monte Carlo experiments are conducted to evaluate the finite sample performance of the proposed bootstrap procedure, and to compare it with the asymptotic test using critical values from a standard normal. The readers are referred to an earlier version of this paper (in particular, Table 1 on page 15) for more simulation results, see Wu, Wu, and Xiao (2022). The Monte Carlo results show that the asymptotic test using normal critical values tends to over-reject the null, and the bootstrap-based test has a much-improved size property—thus improving the finite sample performance of the test.

4. MONTE CARLO SIMULATION

In this section, we conduct Monte Carlo experiments to investigate the finite-sample performance of the nonparametric test J_T, and compare it with three other tests S_{st}, S_w, and S_b considered by Gianetto and Raïssi (2015)[2]. For J_T, we use the Epanechnikov kernel $k(u) = \frac{3}{4}\left(1-u^2\right) I\left(|u| \le 1\right)$ and choose the bandwidth h based on the data-driven method given by (24). In addition, we also specify three different bandwidths $h = \gamma T^{-1/5}$, where $\gamma = 0.5, 0.75$, and 1.0, to demonstrate that our bootstrapped test seems to be less sensitive to bandwidth selection. The corresponding tests are denoted as $J_{T,cv}^B, J_{T1}^B, J_{T2}^B$, and J_{T3}^B, respectively. The first experiment design follows Gianetto and Raïssi (2015), and we consider the following bivariate VAR(2) model

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.2 \\ 0.3 & -0.3 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \begin{pmatrix} 0.1 & 0.3 \\ 0.1 & 0.4 \end{pmatrix} \begin{pmatrix} Y_{1t-2} \\ Y_{2t-2} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix},$$

where the innovations are Gaussian with the variance-covariance structure $\Sigma(\cdot)$ fulfilling Assumption 3. To investigate the size performance of the new test under the null, we consider no instantaneous causality relation between Y_{1t} and Y_{2t} by the following null

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

where $\Sigma^{11}(r) = 1.1 - \cos(11r)$, and $\Sigma^{22}(r) = 1.1 + \sin(11r)$ correspond to the nonconstant smooth variances of the innovations.

We generate 1000 data sets of $\{Y_t\}_{t=1}^T$ for each $T = 100, 200$, 500, and 800. For the tests J_{T1}^B, J_{T2}^B, J_{T3}^B, and $J_{T,cv}^B$, their testing results are obtained from the wild bootstrap method presented in Section 3.4. Similarly, the critical values of S_b are also obtained by the wild bootstrap method of Gianetto and Raïssi (2015). In all our experiments we use $B = 299$ bootstrap iterations for each simulated data set.

[2] S_{st} is the standard test statistic given by $S_{st} = \delta_1' \hat{\Omega}_{s,t}^{-1} \delta_1$ with $\delta_1 = T^{-1/2} \sum_{t=p+1}^T \hat{u}_{2t} \otimes \hat{u}_{1t}$ and $\hat{\Omega}_{s,t} = \left(T^{-1} \sum_{t=p+1}^T \hat{u}_{2t}\hat{u}_{2t}'\right) \otimes \left(T^{-1} \sum_{t=p+1}^T \hat{u}_{1t}\hat{u}_{1t}'\right)$, S_w is the White corrected test given by $S_w = \delta_1' \hat{\Omega}_w^{-1} \delta_1$ with $\hat{\Omega}_w = T^{-1} \sum_{t=p+1}^T \hat{u}_{2t}\hat{u}_{2t}' \otimes \hat{u}_{1t}\hat{u}_{1t}'$, and $S_b = \sup_{r \in [0,1]} \|\delta_r\|^2$ with $\delta_r = T^{-1/2} \sum_{t=p+1}^{\lfloor Tr \rfloor} \hat{u}_{2t} \otimes \hat{u}_{1t}$, $r \in [0,1]$, where $\left(\hat{u}_{1t}', \hat{u}_{2t}'\right)' = Y_t - \left(X_{t-1}' \otimes I_d\right) \hat{\Pi}$ is the OLS residual.

Table 1. Empirical sizes of the tests under *DGPS*.1

	J_{T1}^{B}	J_{T2}^{B}	J_{T3}^{B}	$J_{T,cv}^{B}$	S_b	S_w	S_{st}
			1% asymptotic nominal level				
$T=100$	0.019	0.018	0.018	0.017	0.016	0.010	0.014
$T=200$	0.008	0.008	0.009	0.008	0.006	0.008	0.007
$T=500$	0.012	0.008	0.008	0.007	0.007	0.010	0.006
$T=800$	0.007	0.006	0.007	0.009	0.002	0.005	0.003
			5% asymptotic nominal level				
$T=100$	0.059	0.057	0.058	0.056	0.057	0.064	0.069
$T=200$	0.057	0.052	0.043	0.049	0.038	0.039	0.037
$T=500$	0.043	0.044	0.045	0.044	0.043	0.045	0.043
$T=800$	0.056	0.052	0.047	0.046	0.046	0.046	0.039
			10% asymptotic nominal level				
$T=100$	0.105	0.103	0.103	0.111	0.126	0.118	0.121
$T=200$	0.096	0.092	0.091	0.091	0.079	0.086	0.078
$T=500$	0.101	0.088	0.086	0.089	0.099	0.099	0.099
$T=800$	0.096	0.099	0.096	0.102	0.094	0.096	0.091

Lastly, the testing results of S_{st} and S_w are based on the asymptotic chi-square critical values. Here the lag length in the VAR(p) model is selected by using the corrected portmanteau statistic $\tilde{Q}_m^{OLS}$ at the 5% significance level with $m=6$, and we specify the maximum lag $p_{\max}=6$. In the Supplementary Material, we also provide the simulation results for $p=2$, where the true lag length is assumed to be known.

Table 1 reports the rejection rates of all tests at the 1%, 5%, and 10% significance levels. We observe that the tests S_b, S_{st}, and S_w have empirical sizes close to the nominal ones in all cases. In addition, the new tests J_{T1}^{B}, J_{T2}^{B}, J_{T3}^{B}, and $J_{T,cv}^{B}$ are insensitive to bandwidth selection, and their estimated sizes are also quite close to their nominal ones. Thus, our wild bootstrap indeed approximates the finite sample null distribution of the test statistic very accurately.

In order to explore the testing power of the proposed test, we still follow Gianetto and Raïssi (2015) and consider the following alternative

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

where $\Sigma^{11}(r)$ and $\Sigma^{22}(r)$ are still defined under the null, and $\Sigma^{12}(r) = c\sin(2\pi r)$ is such that $\int_0^1 \Sigma^{12}(r)\,dr = 0$ with $\Sigma^{12}(r) \neq 0$ almost everywhere for $r \in [0,1]$. For the time being, we let $c=0.5$. Later, we shall treat the testing power of J_T as a function of c, and try to investigate whether J_T enjoys monotonic power when the deviation from the null is increased.

TABLE 2. Empirical power of the tests under *DGPP*.1

	J_{T1}^B	J_{T2}^B	J_{T3}^B	$J_{T,cv}^B$	S_b	S_w	S_{st}
			1% asymptotic nominal level				
$T = 100$	0.311	0.333	0.305	0.304	0.017	0.010	0.016
$T = 200$	0.659	0.718	0.711	0.708	0.072	0.005	0.010
$T = 500$	0.989	0.994	0.994	0.992	0.488	0.008	0.024
$T = 800$	1.000	1.000	1.000	1.000	0.892	0.012	0.017
			5% asymptotic nominal level				
$T = 100$	0.538	0.579	0.567	0.564	0.127	0.048	0.061
$T = 200$	0.849	0.872	0.874	0.871	0.312	0.062	0.076
$T = 500$	0.999	1.000	1.000	1.000	0.841	0.055	0.085
$T = 800$	1.000	1.000	1.000	1.000	0.994	0.051	0.059
			10% asymptotic nominal level				
$T = 100$	0.662	0.690	0.699	0.679	0.260	0.081	0.108
$T = 200$	0.909	0.924	0.926	0.925	0.499	0.096	0.128
$T = 500$	1.000	1.000	1.000	1.000	0.947	0.100	0.117
$T = 800$	1.000	1.000	1.000	1.000	0.999	0.090	0.115

We still generate 1000 data sets of $\{Y_t\}_{t=1}^T$ for each $T = 100, 200$, 500, and 800. Table 2 reports the rejection rates of all the tests at the 1%, 5%, and 10% levels, and shows that the bootstrapped tests J_{T1}^B, J_{T2}^B, J_{T3}^B, and $J_{T,cv}^B$ are much more powerful than the three tests S_b, S_{st}, and S_w considered by Gianetto and Raïssi (2015) for all cases. Although the choice of bandwidth has almost no impact on the estimated sizes of the proposed test, it does on its testing power to some extent, which seems to ameliorate by increasing the sample size T. In addition, we also note that the tests S_w and S_{st} have almost no power even though the sample size T is increased to 800, and the test S_b exhibits some power but is always inferior to our four bootstrapped tests.

In order to illustrate the ability of the proposed test to detect the alternative with monotonic power, we assume the empirical power of all tests to be functions of c. When $c = 0$, we are back to the null hypothesis. Here we keep all specifications unchanged as in the power experiment, except that we let $T = 200$ and choose the 5% nominal level. Since the bootstrapped tests J_{T1}^B, J_{T2}^B, J_{T3}^B, and $J_{T,cv}^B$ display similar power performance, here we only compare $J_{T,cv}^B$ with the tests S_b, S_{st}, and S_w. We clearly observe from Figure 1 that the tests S_w and S_{st} have almost no power as we gradually increase c values, while the other two tests $J_{T,cv}^B$ and S_b exhibit monotonic power. Comparatively speaking, the test $J_{T,cv}^B$ is much more powerful than the test S_b, and it has a much faster climbing rate. When the deviation c is increased to a larger extent, the power function of $J_{T,cv}^B$ approaches unity.

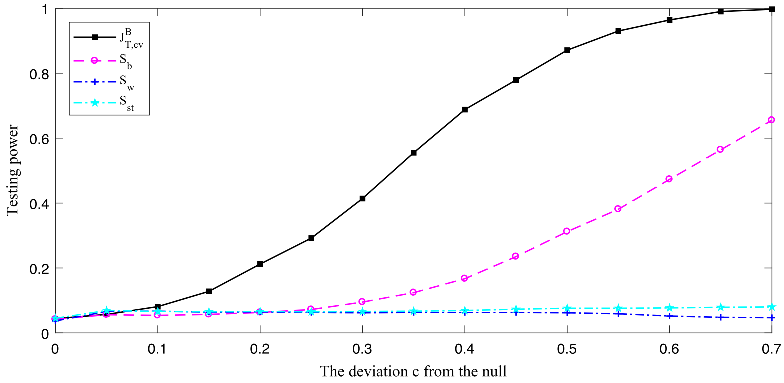

FIGURE 1. The testing power curves of the tests under *DGPP*.1.

The above Monte Carlo simulation only considers smooth structural change in $\Sigma(\cdot)$. Next, we study two cases of abrupt structural breaks in $\Sigma(\cdot)$. The entire designs of the Monte Carlo experiment are the same as before, except that we replace the smooth change with the structural breaks. First, to investigate the testing sizes under the null, we consider the following specifications for $\Sigma(\cdot)$:

$$\Sigma(r) = \begin{pmatrix} \Sigma^{11}(r) & 0 \\ 0 & \Sigma^{22}(r) \end{pmatrix}, \forall\, r \in [0,1], \tag{25}$$

where the first case is given by

$$DGPS.2 : \Sigma^{11}(r) = \begin{cases} 1 & r \le 0.25 \\ 5 & r > 0.25 \end{cases}, \text{ and } \Sigma^{22}(r) = \begin{cases} 3 & r \le 0.75 \\ 1 & r > 0.75 \end{cases}, \tag{26}$$

which corresponds to a single structural break in the unconditional variances, and the second case is given by

$$DGPS.3 : \Sigma^{11}(r) = \begin{cases} 2 & r < 0.2 \\ 5 & 0.2 \le r \le 0.8 \\ 2 & r > 0.8 \end{cases}, \text{ and } \Sigma^{22}(r) = \begin{cases} 3 & r < 0.4 \\ 4 & 0.4 \le r \le 0.6 \\ 2 & r > 0.6 \end{cases}, \tag{27}$$

which contains two structural breaks in the unconditional variances.

The results for *DGPS*.2 and *DGPS*.3 are reported in Tables 3 and 4. We find that the size performance of all the tests in both cases is similar to that of the smoothly changing case in Table 1, and their estimated sizes are close to the nominal ones at the three different significance levels.

In order to explore the testing power of our proposed test under structural breaks, we also consider the following specifications for $\Sigma(\cdot)$:

$$\Sigma(r) = \begin{pmatrix} \Sigma^{11}(r) & \Sigma^{12}(r) \\ \Sigma^{12}(r) & \Sigma^{22}(r) \end{pmatrix}, \forall\, r \in [0,1], \tag{28}$$

TABLE 3. Empirical sizes of the tests under *DGPS*.2

	J_{T1}^B	J_{T2}^B	J_{T3}^B	$J_{T,cv}^B$	S_b	S_w	S_{st}
			1% asymptotic nominal level				
$T=100$	0.012	0.012	0.014	0.014	0.009	0.009	0.014
$T=200$	0.018	0.015	0.014	0.012	0.008	0.009	0.010
$T=500$	0.015	0.016	0.012	0.015	0.011	0.006	0.008
$T=800$	0.011	0.008	0.007	0.008	0.007	0.005	0.004
			5% asymptotic nominal level				
$T=100$	0.057	0.055	0.058	0.057	0.065	0.060	0.064
$T=200$	0.039	0.038	0.036	0.039	0.043	0.039	0.032
$T=500$	0.049	0.047	0.044	0.048	0.046	0.052	0.048
$T=800$	0.052	0.045	0.044	0.046	0.045	0.042	0.032
			10% asymptotic nominal level				
$T=100$	0.102	0.107	0.104	0.105	0.109	0.110	0.106
$T=200$	0.078	0.073	0.075	0.075	0.086	0.085	0.077
$T=500$	0.090	0.097	0.104	0.092	0.099	0.113	0.107
$T=800$	0.089	0.091	0.099	0.098	0.100	0.101	0.090

TABLE 4. Empirical sizes of the tests under *DGPS*.3

	J_{T1}^B	J_{T2}^B	J_{T3}^B	$J_{T,cv}^B$	S_b	S_w	S_{st}
			1% asymptotic nominal level				
$T=100$	0.012	0.014	0.010	0.012	0.006	0.009	0.015
$T=200$	0.015	0.013	0.012	0.012	0.009	0.008	0.010
$T=500$	0.012	0.013	0.015	0.016	0.008	0.009	0.011
$T=800$	0.008	0.006	0.008	0.007	0.008	0.007	0.007
			5% asymptotic nominal level				
$T=100$	0.056	0.055	0.050	0.054	0.054	0.060	0.066
$T=200$	0.045	0.044	0.043	0.044	0.042	0.037	0.045
$T=500$	0.040	0.044	0.039	0.040	0.045	0.052	0.053
$T=800$	0.043	0.039	0.039	0.041	0.041	0.042	0.049
			10% asymptotic nominal level				
$T=100$	0.094	0.096	0.095	0.096	0.110	0.112	0.119
$T=200$	0.079	0.076	0.082	0.078	0.090	0.092	0.095
$T=500$	0.089	0.087	0.098	0.096	0.101	0.108	0.117
$T=800$	0.084	0.091	0.095	0.098	0.090	0.097	0.099

where, specifically, $\Sigma(\cdot)$ takes the following two data generating processes: $DGPP.2$: $\Sigma^{11}(\cdot)$ and $\Sigma^{22}(\cdot)$ are the same as in (26), and $\Sigma^{12}(\cdot)$ is specified as

$$\Sigma^{12}(r) = \begin{cases} 0 & r \leq 0.5 \\ c & r > 0.5 \end{cases}, c = 1. \tag{29}$$

$DGPP.3$: $\Sigma^{11}(\cdot)$ and $\Sigma^{22}(\cdot)$ are the same as in (27), and $\Sigma^{12}(\cdot)$ is specified as

$$\Sigma^{12}(r) = \begin{cases} -c & r < 0.3 \\ 0 & 0.3 \leq r \leq 0.7 \\ c & r > 0.7 \end{cases}, c = 1. \tag{30}$$

Tables 5 and 6 report the estimated power for $DGPP.2$ and $DGPP.3$. When $\Sigma(\cdot)$ exhibits one breakpoint, we find that all the tests exhibit high power. Comparatively speaking, our bootstrapped tests J_{T1}^B, J_{T2}^B, J_{T3}^B, and $J_{T,cv}^B$ always enjoy marginal advantages over the other three tests S_b, S_w, and S_{st} most of the time, and the test S_b does not dominate the tests S_{st} and S_w. In contrast, the latter two ones are always a little better than the former one. When we turn to the case of two breaks in $\Sigma(\cdot)$, we find that our bootstrapped tests obviously outperform the other three tests. The test S_b is the second best and enjoys some testing power. However, the tests S_w and S_{st} have almost no testing power in this case even though we increase the sample size T from 100 to 800.

TABLE 5. Empirical power of the tests under $DGPP.2$

	J_{T1}^B	J_{T2}^B	J_{T3}^B	$J_{T,cv}^B$	S_b	S_w	S_{st}
			1% asymptotic nominal level				
$T = 100$	0.175	0.205	0.204	0.204	0.127	0.186	0.167
$T = 200$	0.403	0.442	0.471	0.473	0.285	0.421	0.386
$T = 500$	0.899	0.925	0.941	0.934	0.755	0.840	0.853
$T = 800$	0.994	0.994	0.997	1.000	0.951	0.987	0.989
			5% asymptotic nominal level				
$T = 100$	0.377	0.408	0.418	0.417	0.329	0.344	0.344
$T = 200$	0.653	0.699	0.716	0.710	0.550	0.668	0.685
$T = 500$	0.975	0.982	0.987	0.986	0.910	0.934	0.941
$T = 800$	1.000	1.000	1.000	1.000	0.991	0.993	0.995
			10% asymptotic nominal level				
$T = 100$	0.494	0.530	0.539	0.533	0.433	0.492	0.500
$T = 200$	0.772	0.796	0.819	0.819	0.683	0.771	0.793
$T = 500$	0.992	0.995	0.994	0.993	0.951	0.968	0.971
$T = 800$	1.000	1.000	1.000	1.000	0.992	0.999	0.999

TABLE 6. Empirical power of the tests under *DGPP*.3.

	J_{T1}^B	J_{T2}^B	J_{T3}^B	$J_{T,cv}^B$	S_b	S_w	S_{st}
			1% asymptotic nominal level				
$T = 100$	0.124	0.122	0.104	0.108	0.023	0.014	0.014
$T = 200$	0.355	0.345	0.316	0.315	0.051	0.005	0.007
$T = 500$	0.934	0.935	0.923	0.928	0.267	0.014	0.018
$T = 800$	0.999	0.997	0.995	0.995	0.567	0.013	0.018
			5% asymptotic nominal level				
$T = 100$	0.312	0.315	0.287	0.262	0.106	0.052	0.061
$T = 200$	0.648	0.651	0.615	0.629	0.201	0.051	0.060
$T = 500$	0.988	0.988	0.987	0.983	0.643	0.051	0.064
$T = 800$	1.000	1.000	1.000	1.000	0.891	0.054	0.064
			10% asymptotic nominal level				
$T = 100$	0.447	0.435	0.406	0.398	0.211	0.093	0.119
$T = 200$	0.773	0.780	0.765	0.763	0.373	0.098	0.117
$T = 500$	0.995	0.997	0.997	0.996	0.821	0.089	0.111
$T = 800$	1.000	1.000	1.000	1.000	0.957	0.095	0.122

Similarly, to confirm that our test has monotonic testing power in the presence of structural breaks in $\Sigma(\cdot)$ when the deviation from the null is increased, we plot the empirical power of all the tests as functions of c (see Figures 2 and 3). The Monte Carlo experiment designs are the same as the smoothly changing case except that we let $\Sigma(\cdot)$ take the forms as generated by (28)–(30). We notice that all the tests exhibit monotonic power when $\Sigma(\cdot)$ contains a single structural break.

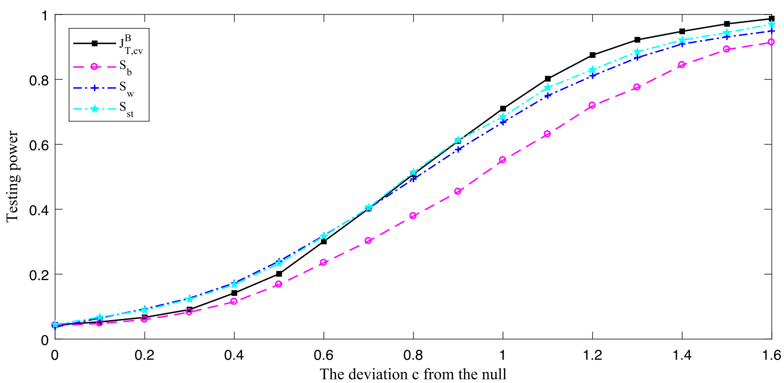

FIGURE 2. The testing power curves of the tests under *DGPP*.2.

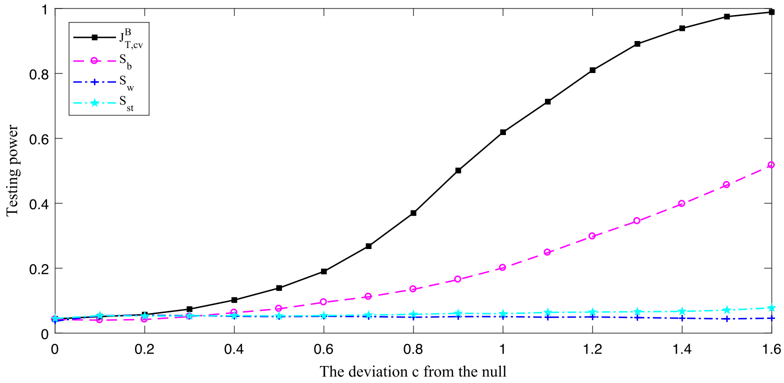

FIGURE 3. The testing power curves of the tests under *DGPP*.3.

Comparatively speaking, the test S_b always exhibits lower testing power than our test $J^B_{T,cv}$ as well as the tests S_w and S_{st} for all values that the deviation c takes. For two structural breaks in $\Sigma(\cdot)$, it is shown that the tests S_w and S_{st} have almost no power even though the deviation c is gradually increased to a large value. In contrast, our test $J^B_{T,cv}$ and the test S_b still exhibit monotonic power. Of course, $J^B_{T,cv}$ is climbing much faster than S_b.

To sum up, our new bootstrapped tests exhibit reasonable testing sizes at different given significance levels. Moreover, they are also more powerful than other popular tests in detecting various forms of instantaneous causality in the presence of time-varying variances.

5. EMPIRICAL APPLICATION TO MONEY SUPPLY AND INFLATION

The link between money supply and inflation is always an important issue in macroeconomics, and has been investigated by many economists in the United States. For example, Turnovsky and Wohar (1984) found that over the period 1923–1960, the inflation rate was independent of the money supply. In contrast, Benderly and Zwick (1985) provided some evidence of a relationship over the period 1955–1988. Gianetto and Raïssi (2015) also tested the relationship from 1979 to 1995 by proposing the test S_b that is robust to time-varying unconditional variances. They rejected no instantaneous causality at the 10% significance level but did not reject it at the 5% level.

In this section, we investigate the relationship between money supply and inflation in the United States by considering a longer period and including more proxy variables for both money supply and inflation, and compare the performance of the proposed test J_T with that of the three tests S_{st}, S_w, and S_b. The data series we employed are monthly frequencies taken from the OECD database (https://data.oecd.org/), covering the period of 01/1959-12/2019. For

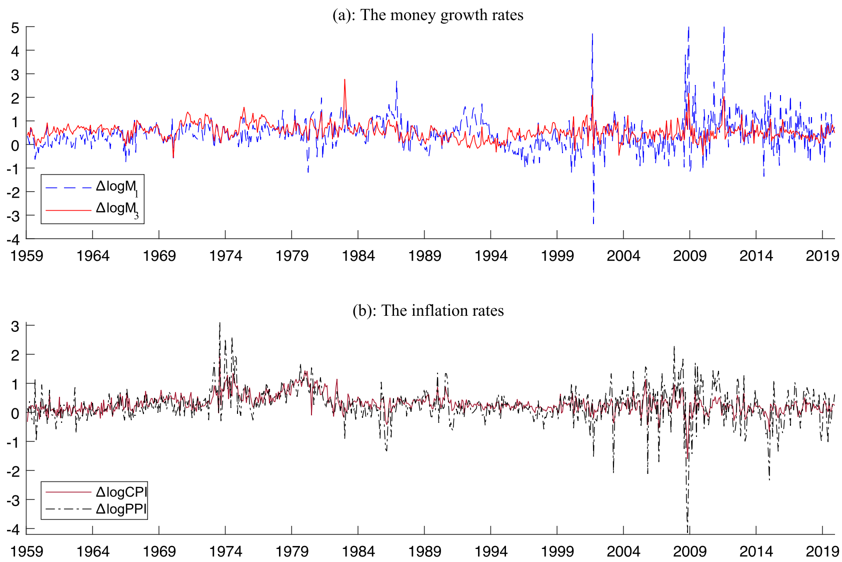

FIGURE 4. The time series of the money growth rates and inflation rates.

money supply, we choose the two money indices M_1 and M_3, representing the narrow money supply and the broad money supply, respectively. For inflation, we employ both the consumer price index (*CPI*) and the producer price index (*PPI*). Specifically, the *CPI* is the total one including food and energy, while the *PPI* is the one of manufacturing for the total market. For all four indices, the base year is specified as $2015 = 100$. In order to obtain money growth rates and inflation rates, we first make seasonal adjustments for the original four series via the X-12 method executed by the software OX. By taking log differences in the seasonally adjusted M_1, M_3, *CPI*, and *PPI* series, we then obtain 731 observations in total for each series, which are plotted in Figure 4(a) and 4(b). Visual inspection of plots reveals that the money growth rate $\Delta \log M_3$ is stable, and the rate $\Delta \log M_1$ is much more volatile, especially after the year 2000. In addition, the two inflation rates, $\Delta \log CPI$ and $\Delta \log PPI$, suffer great variations around the year 1974 and after the year 2000. Next, we apply the test of Wu and Xiao (2018b) to check whether the unconditional variances of the macroeconomic variables exhibit time-varying features. The testing procedure is implemented by using the estimated residuals of the $AR(p)$ models for $\Delta \log M_1$, $\Delta \log M_3$, $\Delta \log CPI$ and $\Delta \log PPI$ with $p = 0, 1, 2, 3$[3]. As the by-products of Wu and Xiao's (2018b) testing procedure, the nonparametric estimates of the unconditional variances are also plotted in Figure 5. From Table 7, we find that the testing results are robust to the different specifications for the conditional means, consistently rejecting the

[3]The MATLAB codes are provided by Wu and Xiao (2018b) with the "leave-q-out" bandwidth $q = \lfloor T^{1/2} \rfloor$.

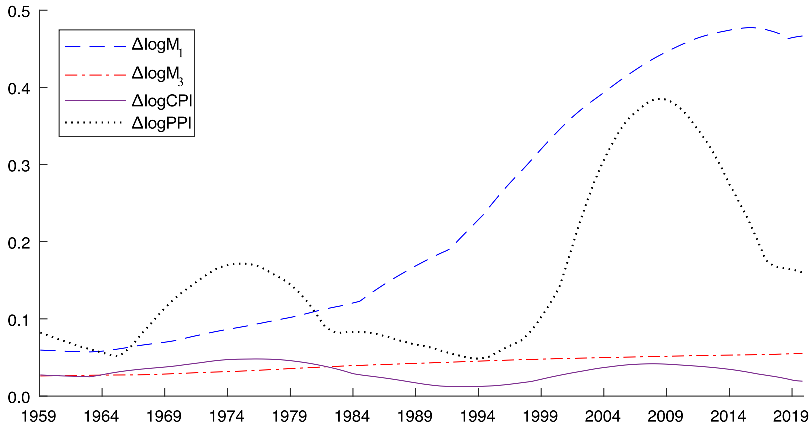

FIGURE 5. The estimated unconditional variances of the four macroeconomic variables.

TABLE 7. The p-values of testing for constant unconditional variances for the money growth rates and inflation rates

	AR(0)	AR(1)	AR(2)	AR(3)
$\Delta\log M_1$	0.000	0.000	0.000	0.000
$\Delta\log M_3$	0.994	0.389	0.510	0.523
$\Delta\log CPI$	0.000	0.018	0.003	0.004
$\Delta\log PPI$	0.000	0.000	0.000	0.000

nulls for $\Delta\log M_1$, $\Delta\log CPI$, and $\Delta\log PPI$ at least at the 2% significance level, and accepting the null for $\Delta\log M_3$ at very large p-values. So we conclude that the unconditional variances of $\Delta\log M_1$, $\Delta\log CPI$ and $\Delta\log PPI$ are statistically time-varying, while that of $\Delta\log M_3$ is constant over time for the given period, which can also be clearly seen from Figure 5.

Now we apply the VAR models to fit the bivariate series of the money growth rate and inflation rate, and the autoregressive lag order is chosen using the corrected portmanteau test $\tilde{Q}_m^{OLS}$. The testing results are given in Table 8, where $m = 3, 6, 12$ are considered for the test. We find that all the bivariate series are well captured by the VAR(1) model since all the p-values of the corrected portmanteau test are close to or equal to one.

Once the linear dynamics of the four bivariate series seem well fitted, we now turn to the analysis of instantaneous causality in the estimated residuals by employing the proposed test J_T and the three other tests S_{st}, S_w, and S_b. The testing procedure is the same as that in the Monte Carlo simulation part. Table 9 shows that no matter what the proxy variables for the money supply and inflation we choose, all the p-values of the bootstrapped tests $J_{T1}^B, J_{T2}^B, J_{T3}^B$, and $J_{T,cv}^B$ are very small, and most of them are equal to zero. Hence, the null of no instantaneous causality is

Table 8. The p-values of the corrected portmanteau test for checking adequacy of the VAR(1) model with time-varying variances

	Number of lags		
02/1959 – 12/2019	$m = 3$	$m = 6$	$m = 12$
$(\Delta \log M_1, \Delta \log CPI)$	0.997	1.000	1.000
$(\Delta \log M_1, \Delta \log PPI)$	0.989	0.992	0.999
$(\Delta \log M_3, \Delta \log CPI)$	0.998	1.000	1.000
$(\Delta \log M_3, \Delta \log PPI)$	0.974	0.983	0.995

Table 9. The p-values of the tests for instantaneous causality between the money supply and inflation

02/1959 – 12/2019	J_{T1}^B	J_{T2}^B	J_{T3}^B	$J_{T,cv}^B$	S_b	S_w	S_{st}
$(\Delta \log M_1, \Delta \log CPI)$	0.000	0.000	0.000	0.000	0.674	0.599	0.528
$(\Delta \log M_1, \Delta \log PPI)$	0.013	0.008	0.005	0.005	0.549	0.999	0.998
$(\Delta \log M_3, \Delta \log CPI)$	0.000	0.000	0.000	0.000	0.404	0.347	0.339
$(\Delta \log M_3, \Delta \log PPI)$	0.000	0.000	0.000	0.000	0.822	0.506	0.349

rejected consistently by the new test. On the contrary, all the p-values of S_{st}, S_w and S_b are very large, and the null cannot be rejected even at the 10% significance level. Considering the reliability of the new test J_T, we conclude that there exists instantaneous causality between money supply and inflation in the USA.

6. EXTENSION TO TIME-VARYING COEFFICIENT VAR MODELS

As pointed out by one of the reviewers, it is more interesting to consider the following VAR(p) model with time-varying coefficients,

$$\begin{cases} Y_t = \sum_{j=1}^{p} A_j(r_t) Y_{t-j} + u_t, \\ u_t = G(r_t)\varepsilon_t, t = 1, 2, \ldots, T, \end{cases} \tag{31}$$

where $r_t = t/T$, the eigenvalues of the matrix $I_d - \sum_{j=1}^{p} A_j(r) z^j$ all lie outside the unit circle uniformly in $r \in [0,1]$ and each element $A_j(r)$ is second order continuously differentiable on $[0,1]$ for $j = 1, \ldots, p$. Denote $\Pi_t = \left(vec(A_1(r_t))', \ldots, vec(A_p(r_t))'\right)'$ and $X_{t-1} = \left(Y'_{t-1}, Y'_{t-2}, \ldots, Y'_{t-p}\right)'$. Then the VAR(p) model in (31) can be rewritten as

$$Y_t = \left(X'_{t-1} \otimes I_d\right) \Pi_t + u_t. \tag{32}$$

Following Gao, Peng, and Yan (2024), the local linear estimator for Π_t is given by

$$\hat{\Pi}_t = [I_{d^2p}, \mathbf{0}_{d^2p}] \left(\sum_{s=p+1}^{T} L_b(r_s - r_t) Z_{s-1} Z'_{s-1} \right)^{-1} \sum_{s=p+1}^{T} L_b(r_s - r_t) Z_{s-1} Y_s, \tag{33}$$

where $Z_{s-1} = \left(X'_{s-1} \otimes I_d, \frac{r_s - r_t}{b} X_{s-1} \otimes I_d \right)'$, $\mathbf{0}_{d^2p}$ is a $d^2p \times d^2p$ zero matrix, $L_b(\cdot) = L(\cdot/b)/b$, b is the bandwidth parameter, and $L(\cdot)$ is a classic kernel function. Under some regularity conditions, Gao et al. (2024) have established a set of asymptotic properties for the estimator $\hat{\Pi}_t$, an information criterion to select the optimal lag, and a Wald-type test to determine the constant coefficients. Denote $\check{u}_t = Y_t - \left(X'_{t-1} \otimes I_d \right) \hat{\Pi}_t$, then we can construct the following U-statistic

$$\check{\lambda}_T = \frac{1}{T^2 h} \sum_{t=p+1}^{T} \sum_{s \neq t} k_{s,t} \check{m}'_t \check{m}_s, \tag{34}$$

where $\check{m}_t = vec(\check{u}_{1t} \check{u}'_{2t})$. In order to obtain a valid test, we standardize $\tilde{\lambda}_T$ by $Th^{1/2}$ and the variance estimator $\check{\sigma}_T^2 = \frac{2}{Th^2} \sum_{t=p+1}^{T} \sum_{s \neq t} k_{s,t}^2 (\check{m}'_t \check{m}_s)^2$. Now, we have the following test statistic:

$$\check{J}_T = \frac{Th^{1/2} \check{\lambda}_T}{\check{\sigma}_T}. \tag{35}$$

To ensure Theorems 1–4 still hold for the test statistic $\check{J}_T$, which implies that the nonparametric residuals $\{\check{u}_t\}$ and the errors $\{u_t\}$ are asymptotically equivalent for our purpose, we need to impose some additional assumptions, especially the relationship of the second bandwidth b with the first one h. We do not provide theoretical justification here, but our additional Monte Carlo simulations show that $\check{J}_T$ also works well for the time-varying VAR models, as shown in the Supplementary Material.

7. CONCLUSION

In this paper, we propose a nonparametric test for instantaneous causality in the presence of time-varying variances. Compared with the existing tests, our proposed test is intuitively appealing and straightforward to compute. It has a simple asymptotically standard normal distribution under the null. The only inputs required in the test are the OLS residuals from VAR(p) models. The test is consistent against various forms of alternatives that deviate from the null, and allows for smooth structural change and structural breaks with known or unknown breakdates in unconditional variances. To reduce the size distortion of the proposed test in finite sample applications, we also propose using a wild bootstrap method to improve its size performance. Monte Carlo simulations indicate that our new test implemented with the bootstrap p-values has both reasonable sizes and all-around monotonic power in finite samples. The new test is then applied to check the relationship between money supply and inflation rates in the USA, and significantly rejects the null of no instantaneous causality. Finally, we also consider one possible extension to test for instantaneous causality in the context of time-varying coefficient VAR models, which is left for future research.

SUPPLEMENTARY MATERIAL

Jilin Wu, Ruike Wu and Zhijie Xiao (February 23, 2024): Supplement to "A Nonparametric Test for Instantaneous Causality with Time-Varying Variances," Econometric Theory Supplementary Material. To view, please visit https://doi.org/10.1017/S0266466624000409.

REFERENCES

Andreou, E., & Ghysels, E. (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17, 579–600.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66, 47–78.

Benderly, J., & Zwick, B. (1985). Inflation, real balances, output, and real stock returns. *American Economic Review*, 75, 1115–1123.

Boubacar, M. Y. (2012). Selection of weak VARMA models by Akaike's information criteria. *Journal of Time Series Analysis*, 33, 121–130.

Breitung, J., & Swanson, N. R. (2002). Temporal aggregation and spurious instantaneous causality in multiple time series models. *Journal of Time Series Analysis*, 23, 651–665.

Brown, B. M. (1971). Martingale central limit theorems. *Annals of Mathematical Statistics*, 42, 59–66.

Cavanaugh, J. E. (1997). Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statistical and Probability Letters*, 33, 201–208.

Cai, Z. W., Wang, Y. F., & Wang, Y. G. (2015). Testing instability in a predictive regression model with nonstationary regressors. *Econometric Theory*, 31, 953–980.

Cavaliere, G. (2005). Unit root tests under time-varying variances. *Econometric Reviews*, 23, 259–292.

Chen, B., & Hong, Y. (2012). Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica*, 80, 1157–1183.

Chitturi, R. V. (1974). Distribution of residual autocorrelations in multiple autoregressive schemes. *Journal of the American Statistical Association*, 69, 928–934.

Clark, T. (2011). Real-time density forecasts from BVARs with stochastic volatility. *Journal of Business and Economic Statistics*, 29, 327–341.

Dufour, J. M., & Taamouti, A. (2010). Short and long run causality measures: Theory and inference. *Journal of Econometrics*, 154, 42–58.

Faes, L., & Nollo, G. (2010). Extended causal modelling to assess partial directed coherence in multiple time series with significant instantaneous interactions. *Biological Cybernetics*, 103, 387–400.

Faes, L., Erla, S., Porta, A., & Nollo, G. (2013). A framework for assessing frequency domain causality in physiological time series with instantaneous effects. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1997), 20110618.

Fu, Z. H., Hong, Y. M., & Wang, X. (2023). On multiple structural breaks in distribution: An empirical characteristic function approach. *Econometric Theory*, 39, 534–581.

Gao, J. T., Peng, B., & Yan, Y. Y. (2024). Estimation, inference, and empirical analysis for time-varying VAR models. *Journal of Business and Economic Statistics*. 42, 310–321
Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77, 304–313.
Geweke, J. (1984). Measures of conditional linear dependence and feedback between time series. *Journal of the American Statistical Association*, 79, 907–915.
Gianetto, Q. G., & Raïssi, H. (2015). Testing instantaneous causality in presence of nonconstant unconditional covariance. *Journal of Business and Economic Statistics*, 33, 46–53.
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 developments in a concept of causality. *Journal of Econometrics*, 39, 199–211.
Granger, C. W. J., & Hyung, N. (2004). Occasional structural breaks and long memory with an application to the SP 500 absolute stock returns. *Journal of Empirical Finance*, 11, 399–421.
GroenJ., Paap, R., & Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business and Economic Statistics*, 31, 29–44.
Hafner, C. (2009). Causality and forecasting in temporally aggregated multivariate GARCH processes. *The Econometrics Journal*, 12, 127–146.
Hafner, C., & Linton, O. (2010). Efficient estimation of a multivariate multiplicative volatility model. *Journal of Econometrics*, 159, 55–73.
Hammoudeh, S., & Li, H. (2008). s Sudden changes in volatility in emerging markets: The case of Gulf Arab stock markets. *International Review of Financial Analysis*, 17, 47–63.
Hamori, S., & Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift. *Economics Letters*, 57, 245–253.
Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, 41, 190–195.
Hansen, B. E. (1995). Regression with nonstationary volatility. *Econometrica*, 63, 1113–1132.
Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in US labor productivity. *Journal of Economic Perspectives* 15, 117–128.
Härdle, W., Mammen, E. (1993). Comparing nonparametric versus parametric regression fits efficiency. *The Annals of Statistics*, 41, 1926–1947.
Hyvärinen, A., Zhang, K., Shimizu, S., & Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non-Gaussianity. *Journal of Machine Learning Research*, 11, 1709–1731.
Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *Journal of the American Statistical Association* 75, 343–386.
Hsiao, Z., & Li, Q. (2001). A consistent test for conditional heteroskedasticity in time series regression models. *Econometric Theory*, 17, 188–221.
Hu, M., & Liang, H. (2014). A Copula approach to assessing Granger causality. *NeuroImage*, 100, 125–134.
Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419–426.
Juhl, T., & Xiao, Z. J. (2005). A nonparametric test for changing trends. *Journal of Econometrics*, 127, 179–199.
Justiniano, A., & Primiceri, G. E. (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review*, 98, 604–641.
Kim, C. J., & Nelson, C. R. (1999). Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *The Review of Economics and Statistics*, 81, 608–616.
Kim, T., Leybourne, H. S., & Newbold, P. (2002). Unit root tests with a break in innovation variance. *Journal of Econometrics*, 109, 365–387.
Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change, and the

GARCH model. *Journal of Business and Economic Statistics*, 8, 225–234.

Lee, T. H., & Ullah, A. (2007). Nonparametric bootstrap tests for neglected nonlinearity in time series regression models. *Journal of Nonparametric Statistics*, 13, 425–451.

Li, D., & Li, Q. (2010). Nonparametric semiparametric estimation and testing of econometric models with data dependent smoothing parameters. *Journal of Econometrics*, 157, 179–190.

Li, Q., & Racine, J. (2004). Cross validated local linear nonparametric regression. *Statistica Sinica*, 14, 485–512.

Li, Y., Wei, H. L., Billings, S. A., & Liao, X. F. (2012). Time-varying linear and nonlinear parametric model for Granger causality analysis. *Physical Review E*, 85, 41906.

Li, Q., & Wang, S. (1998). A simple consistent bootstrap test for a parametric regression function. *Journal of Econometrics*, 87, 145–165.

Linton, O., & Xiao, Z. J. (2019). Efficient estimation of nonparametric regression in the presence of dynamic heteroskedasticity. *Journal of Econometrics*, 213, 608–631.

Liu, C., & Maheu, J. M. (2008). Are there structural breaks in realized volatility?. *Journal of Financial Econometrics*, 6, 326–360.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Mikosch, T., & Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics*, 86, 378–390.

Patilea, V., & Raïssi, H. (2012). Adaptive estimation of vector autoregressive models with time-varying variance: Application to testing linear causality in mean. *Journal of Statistical Planning and Inference*, 142, 2891–2912.

Patilea, V., & Raïssi, H. (2013). Corrected Portmanteau tests for VAR models with time varying variance. *Journal of Multivariate Analysis*, 116, 190–207.

Pierce, D. A., & Haugh, L. D. (1977). Causality in temporal systems: Characterizations and a survey. *Journal of Econometrics*, 5, 265–293.

Raïssi, H. (2011). Testing instantaneous linear Granger causality in presence of nonlinear dynamics. *Comptes Rendus Mathematique*, 349, 1203–1206.

Robinson, P. M. (1989). Nonparametric estimation of time-varying parameters. In P. Hackl (Ed.), *Statistical analysis and forecasting of economic structural change* (pp. 253–264). Springer.

Sensier, M., & Van Dijk, D. (2004). Testing for volatility changes in U.S. macroeconomic time series. *The Review of Economics and Statistics*, 86, 833–839.

Sims, C. (1972). Money, income, and causality. *American Economic Review*, 62, 540–552.

Song, X., & Taamouti, A. (2018). Measuring nonlinear Granger causality in mean. *Journal of Business and Economics Statistics*, 36, 321–333.

Song, X., & Taamouti, A. (2021). Measuring Granger causality in quantiles. *Journal of Business and Economic Statistics*, 39, 937–952.

Taamouti, A., Bouezmarni, T., & El Ghouch, A. (2014). Nonparametric estimation and inference for Granger causality measures. *Journal of Econometrics*, 180, 251–264.

Turnovsky, S. J., & Wohar, M. E. (1984). Monetarism and the aggregate economy: Some longer-run evidence. *The Review of Economics and Statistics*, 66, 619–629.

Van Dijk, D., Osborn, D. R ., & Sensier, M. (2005). Testing for causality in variance in the presence of breaks. *Economics Letters*, 89, 193–199.

Vilasuso, J. (2001). Causality tests and conditional heteroskedasticity: Monte Carlo evidence. *Journal of Econometrics*, 101, 25–35.

Wu, J. L., Wu, R. K. , & Xiao, Z. J. (2022). A nonparametric test for instantaneous causality with time-varying variances. Working paper.

Wu, J. L., Xiao, Z. J. (2018a). A powerful test for changing trends in time series models. *Journal of Time Series Analysis*, 39, 488–501.

Wu, J. L., & Xiao, Z. J. (2018b). Testing for changing volatility. *The Econometrics Journal*, 21, 192–217.

Xia, Y., & Li, W. K. (2002). Asymptotic behavior of bandwidth selected by cross validation method for local polynomial fitting. *Journal of Multivariate Analysis*, 83, 265–287.

Xu, K. L., & Phillips, P. C. B. (2008). Adaptive estimation of autoregressive models with time-varying variances. *Journal of Econometrics*, 142, 265–280.

Xu, K. L. (2012). Robustifying multivariate trend tests to nonstationary volatility. *Journal of Econometrics*, 169, 147–154.

Zheng, J. X. (1996). A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics*, 75, 263–290.